Target Shooting? Benchmark-driven Earnings Management in Germany

D I S S E R T A T I O N
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Management

submitted by

Alexander Eisele

from

Germany

Approved on the application of

Prof. Dr. Peter Leibfried

and

Prof. Dr. Thomas Berndt

Dissertation no. 4009

Difo-Druck GmbH, Bamberg 2012
The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, May 11, 2012

The President:

Prof. Dr. Thomas Bieger
For my parents
# Table of Contents

Table of Contents ........................................... I

Abstract .................................................................. V

Abbreviations .................................................. VI

List of Figures .................................................... VIII

List of Tables .................................................... X

## 1 Introduction ..................................................... 1

## 2 Earnings Management ........................................ 3

2.1 Definition of Earnings Management ............................. 4

2.2 Motives for Earnings Management .............................. 6

2.2.1 Contracting Incentives ......................................... 6

2.2.2 Regulatory Incentives .......................................... 7

2.2.3 Capital Market Incentives ..................................... 8

2.3 Methods to Manage Earnings ................................. 12

2.3.1 Accrual-based Earnings Management ....................... 12

2.3.2 Real Earnings Management .................................. 13

2.3.3 Comparison of Earnings Management Techniques ........... 15

2.4 Expectations Management .................................... 17

## 3 Earnings and Expectations Management to Achieve Benchmarks ........................................... 18

3.1 Theoretical Explanations for Benchmark Beating ............ 18

3.1.1 Prospect Theory ............................................. 19

3.1.2 Transaction Costs ........................................... 23
## Table of Contents

### 3.2 Empirical Evidence of Benchmark Beating
- 3.2.1 Frequency of Earnings and Expectations Management
  - 3.2.1.1 Non-US Evidence and International Differences
  - 3.2.1.2 Temporal Shifts in Benchmark Beating
  - 3.2.1.3 Critical Studies
- 3.2.2 Earnings Management Techniques
  - 3.2.2.1 Accrual-based Earnings Management
    - 3.2.2.1.1 Aggregate Accruals Management
    - 3.2.2.1.2 Specific Accruals Management
  - 3.2.2.2 Real Earnings Management
    - 3.2.2.2.1 Boosting Revenues and Other Income
    - 3.2.2.2.2 Cutting Costs
  - 3.2.2.3 Expectations Management
  - 3.2.2.4 Competing Strategies
- 3.2.3 Incentives for Benchmark Beating
  - 3.2.3.1 Capital Market Incentives
    - 3.2.3.1.1 The Market Reaction to Benchmark Beating
    - 3.2.3.1.2 Investor Rationality, Opportunism, and Signaling
  - 3.2.3.2 Contracting Incentives
- 3.3 Surveys on Benchmark Beating

### 4 Research Questions, General Methodology, and Contribution
- 4.1 The German Economy as Institutional Setting
- 4.2 Research Questions and General Methodology
- 4.3 Contribution

### 5 The Market Reaction to Benchmark Beating
- 5.1 Empirical Methodology
- 5.2 Variable Definitions
  - 5.2.1 Dependent Variables
  - 5.2.2 Independent Variables
- 5.3 Sample Selection and Descriptive Statistics
  - 5.3.1 Sample Selection
  - 5.3.2 Descriptive Statistics
- 5.4 Empirical Results
  - 5.4.1 The Analyst Forecast Benchmark


<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>The Zero Earnings Benchmark</td>
<td>81</td>
</tr>
<tr>
<td>5.4</td>
<td>The Earnings Changes Benchmark</td>
<td>84</td>
</tr>
<tr>
<td>5.4</td>
<td>Robustness Tests</td>
<td>87</td>
</tr>
<tr>
<td>5.4.4</td>
<td>Non-linearities in the Returns/Earnings Relationship</td>
<td>87</td>
</tr>
<tr>
<td>5.4.4</td>
<td>Specification Issues</td>
<td>89</td>
</tr>
<tr>
<td>5.5</td>
<td>Summary</td>
<td>90</td>
</tr>
<tr>
<td>6</td>
<td>The Prevalence of Benchmark Beating</td>
<td>91</td>
</tr>
<tr>
<td>6.1</td>
<td>Empirical Methodology</td>
<td>92</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Visual Inspection of Earnings Distributions</td>
<td>93</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Statistical Tests of Distributional Discontinuity</td>
<td>95</td>
</tr>
<tr>
<td>6.1.2.1</td>
<td>The Burgstahler and Dichev Test Statistic</td>
<td>96</td>
</tr>
<tr>
<td>6.1.2.2</td>
<td>Test Statistic based on Kernel Density Estimates</td>
<td>99</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Critical Evaluation of the Distributional Approach</td>
<td>103</td>
</tr>
<tr>
<td>6.2</td>
<td>Variable Definitions</td>
<td>105</td>
</tr>
<tr>
<td>6.3</td>
<td>Sample Selection and Descriptive Statistics</td>
<td>109</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Sample Selection</td>
<td>109</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Descriptive Statistics</td>
<td>111</td>
</tr>
<tr>
<td>6.4</td>
<td>Empirical Results</td>
<td>113</td>
</tr>
<tr>
<td>6.4.1</td>
<td>The Analyst Forecast Benchmark</td>
<td>113</td>
</tr>
<tr>
<td>6.4.2</td>
<td>The Zero Earnings Benchmark</td>
<td>117</td>
</tr>
<tr>
<td>6.4.3</td>
<td>The Earnings Changes Benchmark</td>
<td>120</td>
</tr>
<tr>
<td>6.4.4</td>
<td>Robustness Tests</td>
<td>124</td>
</tr>
<tr>
<td>6.4.4.1</td>
<td>Bin- and Bandwidth Choice</td>
<td>125</td>
</tr>
<tr>
<td>6.4.4.2</td>
<td>Alternative Kernel Functions</td>
<td>126</td>
</tr>
<tr>
<td>6.4.4.3</td>
<td>Alternative Deflators</td>
<td>128</td>
</tr>
<tr>
<td>6.4.4.4</td>
<td>The Effect of Tax Asymmetries</td>
<td>134</td>
</tr>
<tr>
<td>6.5</td>
<td>Summary</td>
<td>137</td>
</tr>
<tr>
<td>7</td>
<td>Management’s Techniques to Achieve Benchmarks</td>
<td>138</td>
</tr>
<tr>
<td>7.1</td>
<td>Empirical Methodology</td>
<td>139</td>
</tr>
<tr>
<td>7.2</td>
<td>Variable Definitions</td>
<td>141</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Proxies for Accrual-based Earnings Management</td>
<td>142</td>
</tr>
<tr>
<td>7.2.1.1</td>
<td>Discretionary Accruals</td>
<td>142</td>
</tr>
<tr>
<td>7.2.1.1.1</td>
<td>The Jones Model</td>
<td>143</td>
</tr>
<tr>
<td>7.2.1.1.2</td>
<td>The Cross-Sectional Jones Model</td>
<td>145</td>
</tr>
<tr>
<td>Chapter</td>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>7.2.1</td>
<td>7.2.1.1.3</td>
<td>Adapted Versions of the Jones Model</td>
</tr>
<tr>
<td>7.2.1</td>
<td>7.2.1.1.4</td>
<td>Accuracy of Accrual Models and Model Choice</td>
</tr>
<tr>
<td>7.2.2</td>
<td>7.2.1.2</td>
<td>Discretionary Revenues</td>
</tr>
<tr>
<td>7.2.2</td>
<td>7.2.2</td>
<td>Proxies for Real Earnings Management</td>
</tr>
<tr>
<td>7.2.2</td>
<td>7.2.2.1</td>
<td>The Roychowdhury Model</td>
</tr>
<tr>
<td>7.2.2</td>
<td>7.2.2.2</td>
<td>The Gunny Model</td>
</tr>
<tr>
<td>7.2.3</td>
<td>7.2.3</td>
<td>Proxy for Expectations Management</td>
</tr>
<tr>
<td>7.3</td>
<td>7.3</td>
<td>Sample Selection and Descriptive Statistics</td>
</tr>
<tr>
<td>7.3.1</td>
<td>7.3.1</td>
<td>Sample Selection</td>
</tr>
<tr>
<td>7.3.2</td>
<td>7.3.2</td>
<td>Descriptive Statistics</td>
</tr>
<tr>
<td>7.4</td>
<td>7.4</td>
<td>Empirical Results</td>
</tr>
<tr>
<td>7.4.1</td>
<td>7.4.1</td>
<td>The Analyst Forecast Benchmark</td>
</tr>
<tr>
<td>7.4.2</td>
<td>7.4.2</td>
<td>The Zero Earnings Benchmark</td>
</tr>
<tr>
<td>7.4.3</td>
<td>7.4.3</td>
<td>The Earnings Changes Benchmark</td>
</tr>
<tr>
<td>7.4.4</td>
<td>7.4.4</td>
<td>Further Analyses and Robustness Tests</td>
</tr>
<tr>
<td>7.4.4.1</td>
<td>7.4.4.1</td>
<td>Alternative Expectations Management Measure</td>
</tr>
<tr>
<td>7.4.4.2</td>
<td>7.4.4.2</td>
<td>Alternative Interval Widths</td>
</tr>
<tr>
<td>7.4.4.3</td>
<td>7.4.4.3</td>
<td>Controlling for Firm Performance</td>
</tr>
<tr>
<td>7.4.4.4</td>
<td>7.4.4.4</td>
<td>The Distribution of Unmanaged EPS</td>
</tr>
<tr>
<td>7.4.4.5</td>
<td>7.4.4.5</td>
<td>Specification Issues</td>
</tr>
<tr>
<td>7.5</td>
<td>7.5</td>
<td>Summary</td>
</tr>
</tbody>
</table>

8 Conclusion | 189

Appendices | 193

A.1 Accounting Discretion under IFRS | 194
A.2 Earnings Surprise Calculation based on I/B/E/S Data | 195
A.3 Accrual Model Accuracy Test | 197
A.4 Rules of Thumb for Bin- and Bandwidth Calculation | 204

References | 207
Abstract

This study examines benchmark-driven earnings management in Germany. Prior research provides evidence that capital market pressures and increasing short-termism force managers to manipulate earnings numbers and market expectations to avoid missing earnings targets. These targets include the latest analyst consensus forecast, zero earnings, and last year’s earnings. While most studies investigate earnings management in the US or the UK, several more recent papers extend this line of research to both international and German firms. Adding to the recent literature on earnings management in an international context, this study addresses three interrelated research questions: 1) Do market forces provide incentives to beat earnings benchmarks?, 2) Is benchmark beating a prevalent phenomenon?, and 3) What techniques do managers apply to achieve earnings targets? The results suggest that, after controlling for the information content in earnings, investors reward achieving the latest analyst consensus forecast or last year’s earnings with a return premium. This finding is consistent with benchmark importance. A subsequent analysis of earnings distributions relative to earnings targets confirms prior research suggesting that managers engage in manipulations to meet earnings targets. This evidence, however, is alleviated when a set of ten earnings and expectations management measures is taken into account.

Three main conclusions can be drawn from the results: 1) Capital market forces are one but not the only explanation for benchmark importance, 2) Managers engage in accrual-based earnings management to meet the zero earnings benchmark and actively guide analysts down to beatable targets, and 3) Irregularities in the distributions of earnings metrics are only partly attributable to managerial discretion. Overall, the findings suggest that benchmark-driven earnings management in Germany is less intense than frequently assumed.
Abbreviations

BD    Burgstahler and Dichev (1997)
BP    Bollen and Pool (2009)
CAPM  Capital Asset Pricing Model
CDAX  Composite DAX
CEO   Chief Executive Officer
CFO   Operating Cash Flow/Chief Financial Officer
COGS  Cost of Goods Sold
DS    Datastream (Financial Database)
E/P   Earnings-to-Price
EPS   Earnings per Share
Eq.   Equation
ERC   Earnings Response Coefficient
ETR   Effective Tax Rate
EU    European Union
EUR   Euro
FRS   Financial Reporting Standards (UK)
GAAP  Generally Accepted Accounting Principles
H_A, H_0  Alternative Hypothesis/Null Hypothesis
I/B/E/S Institutional Brokers’ Estimate System (Financial Database)
IAS   International Accounting Standard(s)
IFRS  International Financial Reporting Standard(s)
IRO   Investor Relations Officer
ITC   US International Trade Commission
M/B   Market-to-Book
MEUR  Million Euro
OLS   Ordinary Least Squares
Par.  Paragraph
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;D</td>
<td>Research and Development</td>
</tr>
<tr>
<td>ROA</td>
<td>Return on Assets</td>
</tr>
<tr>
<td>RQ</td>
<td>Research Question</td>
</tr>
<tr>
<td>SEC</td>
<td>US Securities and Exchange Commission</td>
</tr>
<tr>
<td>SG&amp;A</td>
<td>Selling, General, and Administrative Expenses</td>
</tr>
<tr>
<td>SIC</td>
<td>Standard Industrial Classification</td>
</tr>
<tr>
<td>SOX</td>
<td>Sarbanes-Oxley Act</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>United States of America</td>
</tr>
<tr>
<td>US GAAP</td>
<td>US Generally Accepted Accounting Principles</td>
</tr>
<tr>
<td>VIF</td>
<td>Variance Inflation Factor</td>
</tr>
<tr>
<td>WS</td>
<td>Worldscope (Financial Database)</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Theoretical Links between Earnings and Stock Prices .......................... 9
3.1 Value and Weighting Functions under Prospect Theory .......................... 22
3.2 Survey Results: Important Earnings Benchmarks ............................... 45
3.3 Survey Results: Benchmark Beating Incentives .................................. 46
3.4 Survey Results: Actions to Achieve Benchmarks ................................. 47
5.1 Calculation of Abnormal Returns on a Time Line ................................. 63
5.2 Identification of Correct Earnings Announcement Dates ........................ 65
5.3 Linear and Non-linear Returns/Earnings Relationships .......................... 87
6.1 Pooled Distribution of Scaled US Earnings ........................................ 93
6.2 Effect of Binwidth Choice ................................................................. 95
6.3 Bias of the BD Test Statistic ............................................................... 99
6.4 Kernel Estimate of an Earnings Density Function .................................. 101
6.5 Distribution of Earnings Surprises per Share ...................................... 115
6.6 Earnings Surprise Histogram and Reference Distribution ....................... 117
6.7 Distribution of EPS ............................................................................. 119
6.8 EPS Histogram and Reference Distribution ......................................... 121
6.9 Distribution of EPS Changes ............................................................... 123
6.10 EPS Changes Histogram and Reference Distribution ............................ 125
6.11 Frequency Histograms of Deflated Earnings Surprises .......................... 131
6.12 Frequency Histograms of Deflated EPS ............................................. 132
6.13 Frequency Histograms of Deflated EPS Changes .................................. 133
6.14 Effect of Tax Asymmetries ................................................................. 135
7.1 Discretionary and Nondiscretionary Accruals ....................................... 143
7.2 Event Study Methodology in Earnings Management Research ................. 145
7.3 Accuracy of the Lagged Model ............................................................. 151
List of Figures

7.4 Forecast Errors and Earnings Surprises .................................................. 180
7.5 Distribution of Reported and Premanaged EPS ........................................ 186
A.2.1 Adjusted and Unadjusted I/B/E/S Data ............................................... 196
List of Tables

2.1 Accrual-based and Real Earnings Management .............................................. 15

3.1 Prospect Theory Experiment ........................................................................... 20

4.1 Traditional Institutional Differences between Germany and the US .............. 49

4.2 German Evidence on Benchmark Beating ....................................................... 51

5.1 Sample Selection Procedure ........................................................................... 70

5.2 Industry Composition ....................................................................................... 71

5.3 Variable Distributions ...................................................................................... 72

5.4 Benchmark Statistics ....................................................................................... 74

5.5 Variable Correlations ....................................................................................... 76

5.6 Market Rewards for Forecast Achievement—Univariate Results .................. 77

5.7 Market Rewards for Forecast Achievement—Regression Results .................. 80

5.8 Market Rewards for Avoiding a Loss—Univariate Results ............................. 81

5.9 Market Rewards for Avoiding a Loss—Regression Results ............................ 83

5.10 Market Rewards for Avoiding EPS Declines—Univariate Results ................. 84

5.11 Market Rewards for Avoiding EPS Declines—Regression Results ............... 86

5.12 Controlling for Non-linearities ........................................................................ 88

6.1 Deflator Choice Analysis ................................................................................... 108

6.2 Sample Selection Procedure ........................................................................... 110

6.3 Variable Distributions ...................................................................................... 112

6.4 Discontinuities in the Distribution of Earnings Surprises ............................... 114

6.5 Discontinuities in the Distribution of EPS ....................................................... 120

6.6 Discontinuities in the Distribution of EPS Changes ......................................... 122

6.7 Bin- and Bandwidth Analysis ........................................................................... 127

6.8 Alternative Kernel Functions ............................................................................ 128

6.9 Deflation Analysis ............................................................................................ 130
Chapter 1

Introduction

Dramatic corporate scandals shook the global financial markets at the beginning of the 21st century. Among the most prominent cases are Enron in the US, Parmalat in Italy, and FlowTex in Germany. The scandals fueled the discussion on managerial integrity and the reliability of financial reporting and led to major regulatory changes in corporate governance and enforcement mechanisms. Furthermore, they spawned tremendous interest in research on managerial discretion and inhibited the discipline of accounting research from being a solely academic exercise. Instead, knowledge about the motives, magnitude and frequency, methods, and resource allocation consequences of earnings management offers fruitful insights for the users of financial statements and regulatory bodies (Healy and Wahlen, 1999).

Managers manipulate earnings to maximize their own utility. In many cases, the manager’s utility is sensitive to earnings or stock price performance. Hence, managers face incentives to influence earnings and share prices by means of earnings management. Behavioral theory and empirical research provide evidence that these incentives increase when earnings are close to thresholds which are considered as focal points by investors and other stakeholders. As a result, managers are suspected to manipulate earnings upward to avoid falling short of important earnings thresholds. This study considers three benchmarks that have been documented to be important targets: Zero earnings per share (EPS), last year’s EPS, and the latest analyst EPS consensus forecast before the earnings announcement. A large fraction of previous research on these benchmarks is based on US data (for relevant reviews see, e.g., Schipper, 1989; Healy and Wahlen, 1999; Dechow and Schrand, 2004; Ronen and Yaari, 2008). European evidence, in contrast, is comparably scarce. This is surprising, given that an international study by Leuz et al. (2003) finds earnings management to be widespread in Europe and ranks Germany within
the “Top Ten” of earnings managing countries. The US, in contrast, exhibits the lowest earnings management score of all countries under investigation. The supposedly high intensity of earnings management in Germany provides rich grounds for my research and allows me to study the benchmark phenomenon in more detail.

This dissertation addresses three interrelated research questions to provide a comprehensive analysis of benchmark beating in Germany. First, I examine whether German investors reward managers for achieving earnings benchmarks. To do so, I analyze the relation of abnormal stock returns at the earnings announcement and the incidence of benchmark achievement. A positive relation suggests capital market incentives for benchmark related earnings management. Second, I focus on the prevalence of benchmark-driven manipulations. If benchmark beating is prevalent, I expect it to be reflected in the pooled frequency distributions of earnings metrics. Eventually, the third part of my research analyzes whether firms engage in specific types of manipulations to beat benchmarks. These include accrual-based earnings management, real earnings management, and forecast guidance.

My study contributes to the research on benchmark related earnings management in several aspects. First, it examines the capital market incentives for benchmark beating in Germany with a multivariate model that controls for the information content of current earnings and other factors known to affect the returns/earnings relationship. Second, it examines German earnings management prevalence under mandatory IFRS reporting with a novel non-parametric approach. In combination with rigid robustness tests, this technique minimizes the risk of erroneous inferences. Third, it is the first German study to provide evidence on firms’ manipulation techniques. Analyzing the incentives, prevalence, and techniques of benchmark beating, this study is of interest for executives, investors, and other stakeholders (e.g., governments or standard setters).

The remainder of this study is structured as follows: Chapter 2 introduces the term “earnings management”, considers possible explanations for managerial discretion, and describes potential manipulation techniques. Chapter 3 provides theoretical explanations for the importance of earnings benchmarks and summarizes prior research. Building on the preceding chapters, Chapter 4 develops research questions, introduces the general empirical methodology, and highlights the main contributions of my study. Empirical results are presented in Chapters 5 to 7. Chapter 5 analyzes capital market rewards for benchmark achievement. Chapter 6 examines the prevalence of benchmark beating. Chapter 7 digs deeper and explores potential techniques applied by managers to achieve earnings thresholds. Eventually, Chapter 8 summarizes, discusses the results, and proposes avenues for further research.
Chapter 2

Earnings Management

The special role of accounting earnings is today accepted in both accounting research and practice. Dechow et al. (1998, p. 133), for example, note:

Earnings occupy a central position in accounting. It is accounting’s summary measure of a firm’s performance. Despite theoretical models that value cash flows, accounting earnings is widely used in share valuation and to measure performance in management and debt contracts.

According to Graham et al. (2005), 51% of surveyed US financial executives name earnings as most important measure reported to the firm’s stakeholders. De Jong et al. (2009) find that—though lower—the relative majority (39%) of US analysts judge earnings as most important measure. Similar results are reported by Nöldeke (2007b) for a survey among financial executives (CFOs and IROs) in Austria, Germany, and Switzerland: with a fraction of 49%, most of the respondents rank earnings as the most important reported financial reporting figure. The same preference holds for financial analysts from Austria and Germany. Since earnings play an important role in valuation, contracting, and regulation, they are often the primary target of managerial discretion.

This first chapter introduces into the field of study by explaining the fundamentals of earnings management. Section 2.1 introduces the term “Earnings Management”. Section 2.2 summarizes potential motives for earnings management. Section 2.3 explains and compares common techniques of earnings management. Eventually, Section 2.4 compares earnings management and forecast guidance.
2.1 Definition of Earnings Management

Several definitions of earnings management have been suggested in the last decades. Though many of these provide proper descriptions of the phenomenon, the definition of Healy and Wahlen (1999) seems to prevail and is today most commonly cited in the literature:

Earnings management occurs when managers use judgment in financial reporting and in structuring transactions to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company or to influence contractual outcomes that depend on reported accounting numbers.

Healy and Wahlen (1999, p. 368)

Healy and Wahlen’s (1999) definition covers two important attributes of earnings management:

1. **Motives for Earnings Management.** Main motives for earnings management are rooted in contracting and/or communication with stakeholders. Popular examples for contracts that depend on accounting data are management bonus plans and/or debt contracts. Communication with stakeholders primarily refers to investors, but is not limited to that group. Earnings manipulation may be used to influence the expectations and decisions of, e.g., investors, financial intermediaries, or debt holders to maximize management’s utility. Motives for earnings management are discussed in detail in Section 2.2.

2. **Types of Earnings Management.** The definition distinguishes between judgment in financial reporting and structuring of transactions. With the first type of earnings management, judgment in financial reporting, Healy and Wahlen (1999) refer to deliberate interventions in the financial reporting process and the manipulation of real business activities. The second type of earnings management, structuring of transactions, refers to the customization of corporate transactions to allow a more favorable accounting treatment under GAAP.

In the course of this dissertation, I use the term earnings management in the sense of Healy and Wahlen (1999) with three important exceptions: First, I introduce two different categories for types of earnings management: 1) Discretion in the preparation of financial statements; and 2) Alteration of the firm’s underlying operations. The first category includes pure accrual management (e.g., discretionary adjustments of bad debt provisions) and classification shifting (e.g., shifting cost components to special items to manage core earnings). The second category covers
all changes in the normal business cycle that are aimed on a deliberate manipulation of net income. This includes real earnings management activities (e.g., boosting sales by means of granting temporary price discounts) and transaction structuring (e.g., customizing lease contracts to allow for operating or finance lease accounting). Second, Healy and Wahlen (1999) define earnings management as a means to deceive stakeholders about the truth. In other words, earnings management is pursued by opportunistic managers to maximize private utility, recklessly of negative consequences on (long-term) firm value and stakeholders’ interests. In a survey among US financial executives, Graham et al. (2006) report that 56% would knowingly defer positive net present value projects to increase short time earnings and avoid missing analyst estimates. Several authors, however, suggest that earnings management is not always opportunistic and suggest some earnings management to play a beneficial signaling role that makes financial reports more informative. \(^1\) Recent empirical studies by Gunny (2010) and Chen et al. (2010) highlight the role of earnings management in signaling management’s private information about the firm’s prospects. Third, the literature on earnings management covers another motive for earnings manipulation that is rooted in the “political cost hypothesis” (Watts and Zimmerman, 1986). This stream of research acknowledges the relation between political costs (e.g., taxes, costs/profits related to anti-trust or import relief investigations) and accounting earnings. Taking these exceptions into account, I define earnings management as follows:

<table>
<thead>
<tr>
<th>Earnings management occurs when managers deliberately</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. apply discretion in the preparation of financial statements (e.g., accrual management or classification shifting), and/or</td>
</tr>
<tr>
<td>2. alter the firm’s underlying operations (e.g., real activities management or transaction structuring)</td>
</tr>
<tr>
<td>to temporarily alter corporate earnings with the intention to</td>
</tr>
<tr>
<td>1. mislead stakeholders about the true economic performance of the firm,</td>
</tr>
<tr>
<td>2. signal management’s private information about the firm’s future prospects, or</td>
</tr>
<tr>
<td>3. influence contractual outcomes and/or political costs that depend on reported accounting numbers.</td>
</tr>
</tbody>
</table>

\(^1\)See Ronen and Yaari (2008, pp. 25–31) for a description of opportunistic and beneficial earnings management.
2.2 Motives for Earnings Management

Incentives for earnings management arise when altering financial results potentially increases the utility of the manipulating agents. As indicated in Section, 2.1 the main incentives for earnings management discussed in the literature include contracting, political costs and (capital market) communication.

2.2.1 Contracting Incentives

Building on Agency and Property Rights Theory, Jensen and Meckling (1976) examine the role of contracting when top management and firm owners both maximize personal utility and, hence, face diverging interests. Specifically, managers are suspected of consuming value-reducing private benefits to the disadvantage of the owners. In these situations, price-protection encourages managers to contract with the owners and align interests to reduce the owners’ risk of expropriation (Watts and Zimmerman, 1986, p. 199). Owners play a monitoring role and keep managements’ compliance with contracts under surveillance. Both, contracting and monitoring cause agency costs that reduce firm-value.

Watts and Zimmerman’s (1986) Positive Accounting Theory posits that accounting numbers facilitate the contracting process and thus reduce agency costs. The popularity of accounting measures in firm contracts is attributable to two key features: 1) Accounting regulation provides an existing framework for the calculation of accounting measures and adapting these rules avoids substantial costs of designing individual contracts with numerous performance measures and contingencies (Lambert, 2003), and 2) The auditing process and the public availability of accounting data makes accounting measures more reliable and transparent than some complicated performance measure “home-cooked” by the local controlling department. Prominent examples for accounting-based contracts are management compensation plans and debt contracts (see, e.g., Watts and Zimmerman, 1986, pp. 200–221). Management compensation plans are often linked to accounting related performance indicators and accounting is required for the calculation of management payouts. Debt covenants are based on financial ratios (such as interest coverage) derived from accounting figures and accounting data is required to monitor whether covenants were breached. In both cases, the allocation of the firm’s cash flows among participating parties depends on the accounting process. Consequently, such contracts motivate self-interested individuals to exercise (opportunistic) accounting discretion (i.e., earnings management) (Watts and
Empirical research has studied the contracting hypothesis of earnings management in great depth. Starting in the late 1970’s, earnings management in the context of managerial compensation raised tremendous interest within the accounting research community. Among the seminal studies is Healy’s (1985) paper on the linkage between management bonus contracts and accounting discretion. Healy (1985) examines the structure of bonus plans for senior corporate executives and reports a piecewise linear relationship between reported income and bonus rewards. More specifically, he shows that the linear relation is limited by a lower and upper bound, whereas the lower bound has to be met to receive a bonus and the upper bound caps the bonus payment. Applying an accrual model to measure the extent of earnings management, he finds that managers choose income decreasing accruals when their bonus plans’ upper or lower bounds are binding (i.e., upward managed earnings would not increase the bonus) and income increasing accruals when these bounds are not binding.\(^2\) An important implication of his finding is that bonus plans do not only trigger income increasing, but also income decreasing earnings management (e.g., big-bath accounting\(^3\)). Following Healy (1985), many others intensively studied the earnings management incentives of management compensation. These studies include, e.g., Holthausen et al. (1995) on bonus plans, Cheng and Warfield (2005) and Bergstresser and Philippon (2006) with a focus on equity compensation, and Kalyta (2009) on retirement benefits.

### 2.2.2 Regulatory Incentives

Another strand of research focuses on regulatory and political incentives for earnings management (see Watts and Zimmerman, 1986, pp. 222–243). Regulation induced earnings management arises in situations when regulatory constraints are implicitly or explicitly tied to accounting data. Although, the regulation explanation is similar to contracting, the latter comprises rules that were voluntarily imposed. In contrast, regulatory constraints do not result from bi- or multilateral contract negotiations, but are imposed by authorities (e.g., government or industry oversight bodies). These regulations may affect either certain industries or the whole economy. Promi-

---

\(^2\)Accrual models are used to separate the fraction of manipulated accruals from total accruals. The manipulated fraction of accruals is then used as proxy for accrual-based earnings management. See Section 7.2.1 for a detailed description of contemporary accrual models.

\(^3\)Big-bath accounting means that firms understate assets or overstate liabilities in the current period to build reserves that may be used to increase earnings in later periods.
nent examples of industry related constraints are regulatory capital requirements in the banking industry (e.g., Basel II), solvency requirements in the insurance industry, and rate regulation in the utilities sector (Healy and Wahlen, 1999). Economy wide regulations include, e.g., tax and anti-trust rules.

A broad body of literature addresses regulation related earnings management. Noteworthy in this context is Jones’s (1991) seminal work on earnings management during import relief investigations, since it was the first study to address regulation-driven earnings management with a powerful measure of discretionary accruals. Jones (1991) examines whether firms that are subject to import relief investigations by the US International Trade Commission (ITC) manage earnings to influence the authority’s decision. In her setting, the ITC grants relief (i.e., protects the firm from foreign competitors through tariff increases, quota reductions, or agreements on import limits) if firms suspected to suffer from foreign competition exhibit a deteriorating financial performance. Using a sample of 23 firms within five different industries, she finds that firms exercise income decreasing accounting discretion during ITC investigations to increase the chance of import reliefs. Industry related studies mainly concentrate on regulatory requirements of the banking (e.g., Collins et al., 1995; Kim and Kross, 1998; Ahmed et al., 1999; Shriives and Dahl, 2003) and insurance industry (e.g., Petroni, 1992; Adiel, 1996; Beaver et al., 2003). Other industry related studies include, e.g., Key (1997) who finds income decreasing earnings management in the cable industry during re-regulation scrutiny by the US Congress. Similarly, Cahan et al. (1997) show that chemical companies recorded income decreasing accruals precedent to an environmental-driven legislation change. Tax related studies include, e.g., Scholes et al. (1992) and Guenther (1994), who examine earnings management around corporate tax reforms.

2.2.3 Capital Market Incentives

It is nowadays widely accepted that accounting earnings (as one of many other information sources) convey information about current and expected firm performance and thus affect share prices. Beaver (1998, p. 86) uses three theoretical links to explain the relation between current accounting earnings and stock prices:

\[\text{For a detailed description of the Jones Model refer to Section 7.2.1.1.}\]
This theoretical concept is consistent with the residual income model under clean surplus accounting as formulated by Ohlson (1995) and Feltham and Ohlson (1995). Clear surplus accounting allows that dividends are expressed in terms of book values ($BV$) and net income ($NI$). Formally, dividends ($DI$) may be expressed as

$$DI_t = NI_t - (BV_t - BV_{t-1}),$$

and this allows to restate the classical dividend-discount valuation model as

$$MV_0 = BV_0 + \sum_{t=1}^{\infty} \frac{E(NI_t - r_E BV_t)}{(1 + r_E)^t},$$

where $MV_0$ is the estimated value of equity and $r_E$ is the company’s cost of capital. The term $E(NI_t - r_E BV_t)$ is referred to as expected “residual income” or “abnormal earnings”. The Ohlson-Model theoretically underpins the aforementioned link between current and expected earnings and share price. The importance of accounting earnings for asset valuation, however, critically hinges on their superiority when compared to other measures (e.g., current cash flows). Empirical evidence supports the role of accounting earnings in the valuation context: Dechow (1994)
shows that earnings are more persistent than cash flows and more strongly associated with stock returns. Dechow et al. (1998) report that earnings are superior in predicting future operating cash flows when compared to current operating cash flows. Penman and Sougiannis (1998) find that valuation errors are on average lower for the residual income model when compared with models based on cash flows. Overall, earnings seem to be a suitable alternative when it comes to asset valuation.

Starting with the seminal contributions of Ball and Brown (1968) and Beaver (1968), positive accounting research intensively examined the relation of accounting earnings and stock prices during the last decades. As Beaver (1998, p. 89) puts it, “few areas of empirical research in finance or accounting have received as much attention as the relation between stock prices and accounting earnings”. A detailed review of this broad body of empirical research is beyond the scope of this study. Comprehensive summaries can be found in, e.g., Watts and Zimmerman (1986), Bernard (1989), Lev (1989), Beaver (1998), and Kothari (2001). In a nutshell, the overall message is that earnings convey at least some information about future firm performance and may affect security pricing. Hence, managers face incentives to engage in earnings manipulations unless equity markets are informationally efficient and investors are able to “see through” earnings management activities (Stlowy and Breton, 2004). Ample evidence in contradiction of market efficiency increased the research community’s interest in capital market-driven earnings management. Kothari (2001, p. 208) describes this development as follows:

Evidence suggesting market inefficiency has also reshaped the nature of questions addressed in the earnings management literature. Specifically, the motivation for earnings management research has expanded from contracting and political process considerations in an efficient market to include earnings management designed to influence prices because investors and the market might be fixated on (or might over- or under-react to) reported financial statement numbers.

Accounting research has identified and investigated four main settings in which firms face capital market incentives to manage earnings:

- **Beating Benchmarks.** Firms manage earnings to achieve simple earnings targets. These targets include, e.g., zero earnings, last year’s earnings, and/or the latest earnings forecast provided by financial analysts. Benchmark related earnings management assumes that investors follow simple heuristics. Just missing the zero earnings frontier or falling closely short of the analyst forecast is expected to have significant deteriorating effects on share
prices (see, e.g., Skinner and Sloan, 2002). Consequently, firms have incentives to manage earnings upward to meet or even beat the respective target. Relevant literature is discussed in Section 3.2.

- **Issuing Stocks.** Stock issues include initial public offerings and seasoned equity offerings. For both types, firms have incentives to exercise income increasing discretion prior to the offer (see Dechow and Schrand, 2004, p. 48–50). Increased earnings supposedly affect investor perception of future performance and thus share price. A higher share price, in turn, increases the firm’s cash proceeds. Studies on earnings management around stock issues include, e.g., Rangan (1998), Teoh et al. (1998a), Teoh et al. (1998b), Teoh et al. (1998c), Shivakumar (2000), and Cohen and Zarowin (2010).

- **Mergers and Acquisitions.** Around mergers and acquisitions, incentives for earnings management may arise for both the acquiror and the acquiree (Dechow and Schrand, 2004, p. 50). In a cash-settled takeover, for instance, the acquiree’s owners may try to inflate share price prior to the transaction to either avoid a (hostile) takeover or maximize cash receipts. In stock-for-stock mergers, both the acquiror and the acquiree have incentives for income increasing earnings management, because maximizing share value reduces the number of exchange stock. In contrast to external acquisitions, management buyouts may stimulate income decreasing earnings management prior to the transaction, because lower pre-deal earnings are supposed to reduce share price (see Dechow and Schrand, 2004, p. 50). A lower share price, in turn, decreases the transaction price and increases the chance for higher abnormal post-buyout returns. Related evidence can be found in, e.g., Perry and Williams (1994), Wu (1997), Easterwood (1998), Erickson and Wang (1999), Louis (2004), and Evangelos et al. (2005).

- **Insider Trading.** Insiders have incentives to inflate (deflate) share price prior to selling (buying) transactions (see Dechow and Schrand, 2004, p. 51). To inflate (deflate) share price in the short run, earnings may be managed up (down) around anticipated events. Studies that analyze the relationship between insider trading and earnings management include, e.g., Elitzur and Yaari (1995), Beneish and Vargus (2002), Bergstresser and Philippon (2006), and Bhojraj et al. (2009).
2.3 Methods to Manage Earnings

Section 2.1 introduced two main categories of earnings management techniques: Accrual-based earnings management and real earnings management. This section discusses and compares these techniques in more detail.

2.3.1 Accrual-based Earnings Management

Accrual accounting forms the basis of most contemporary accounting systems around the world. As far as accruals are based on solid rules of revenue recognition and matching of income and expenses, earnings are a better measure of firm performance and more useful in predicting future cash flows than current cash flows (Dechow and Schrand, 2004, p. 11). However, accruals are the outcome of an estimation process with a high degree of uncertainty. As such, accruals may be affected by intentional (i.e., earnings management) or unintentional measurement errors.

Throughout this study, I refer to the deliberate manipulation of earnings without direct cash flow impact as accrual-based earnings management (Roychowdhury, 2004). Among many others, this definition covers the following managerial actions:

- **Property, Plant, and Equipment/Intangibles.** Areas of discretion include the recognition and subsequent valuation of assets, especially with regard to depreciation (e.g., methods, estimation of useful life and recoverable amounts) and impairment charges (e.g., identifying triggering events, fair value measurement). Another area of discretion arises from the classification of lease contracts and the valuation of finance lease assets and obligations.

- **Accounts Receivable and Other Assets.** Managers have some discretion to decide whether costs are capitalized as other assets or expensed immediately. Accounts receivable are affected by premature or delayed revenue recognition and the estimation of allowances for doubtful debt.

- **Inventories.** Areas of discretion in accounting for inventories arise upon initial recognition

---

5The term “accruals” does not refer to accrued balance sheet accounts (e.g., provisions or receivables). Instead, accruals represent the portion of net income resulting from changes in such accounts.

6A similar list of possible accrual manipulations can be found in, e.g., Dechow and Schrand (2004, pp. 40–41).
and in subsequent valuation. Discretion upon initial recognition may include the capital-
ization of certain costs (e.g., overheads, research and development, or interest expenses).
In subsequent periods, inventory allowances (e.g., estimating recoverable amounts) and
methods to determine the sequence of consumption (e.g., last-in, first-out) provide possi-
bilities to manage earnings. Furthermore, accounting for long-term construction contracts
offers leeway for earnings manipulations.

− **Provisions and Pension Accounting.** Defined as liabilities of uncertain timing or amount,
provisions require a high level of managerial judgment and thus offer a convenient op-
portunity for earnings management. Pension accounting, in specific, allows discretion in,
e.g., expected plan asset returns and actuarial assumptions (e.g., life expectancies, interest
rates).

− **Financial Instruments.** Accounting for financial instruments offers a wide range of ac-
counting choice and judgment. This includes the recognition and derecognition of financial
instruments, fair value measurement (e.g., assumptions underlying valuation models), im-
pairment considerations (e.g., identification of triggering events), and hedge accounting
(e.g., designation, effectiveness).

Appendix A.1 provides a comprehensive summary of earnings management opportunities
under IFRS.

### 2.3.2 Real Earnings Management

Historically, the discussion on earnings management primarily focused on accrual manipula-
tions. Strengthening enforcement regulation, investor protection, and executive responsibility
(especially following the Sarbanes-Oxley (SOX) regulation of 2002), however, increased the
attention to real earnings management activities (Cohen et al., 2008). In 2005, Graham et al.
published the results of a survey among more than 400 executives on the factors that drive re-
ported earnings and disclosure decisions. They find that managers would rather take economic
actions with negative long-term consequences than within-GAAP accounting choices to manage
earnings. Concerning the prevalence of real earnings management in practice, they report:

> [We] find strong evidence that managers take real economic actions to maintain
accounting appearances. In particular, 80% of survey participants report that they
would decrease discretionary spending on R&D, advertising, and maintenance to meet an earnings target. More than half (55.3\%) state that they would delay starting a new project to meet an earnings target, even if such a delay entailed a small sacrifice in value. (Graham et al., 2005, pp. 30–31)

Real earnings management does not focus on the immediate manipulation of financial accounts. Instead, it subsumes actions that change the timing and structuring of real transactions and thus alter financial results (Ewert and Wagenhofer, 2005). As summarized by Xu et al. (2007), such actions include:

- **Manipulation of Discretionary Expenditures.** Managers may exercise discretion in the timing of discretionary expenditures. Discretionary expenditures include payments that are immediately expensed under GAAP, such as costs for advertising, maintenance, or research. Postponing discretionary expenditures to the next reporting period increases current period earnings at the expense of future earnings, while an anticipation of future expenditures decreases current period earnings in favor of future earnings.

- **Manipulation of Production, Inventory, and Sales.** Managers may manipulate earnings by altering production volume. Since fixed production costs decrease on a per item level when production volume increases, firms may lower costs of goods sold (COGS) by means of overproduction. In contrast, underproduction increases COGS and relieves future earnings in expense of current earnings. Sales may be pushed in the current year by anticipation of future periods’ demand (channel stuffing). Offering customers large discounts at the end of the current year increases current profit at the expense of future earnings.

- **Sale of Non-Current Assets.** Managers may exercise discretion in the timing of fixed asset and investment sales. If the market value of long-term assets exceeds book value, selling these assets increases current period earnings. Conversely, selling long-term assets at a price below book value decreases current earnings.

- **Transaction Structuring.** Managers may structure real transactions in order to take advantage of more favorably rules under GAAP. This includes, e.g., building up special purpose entities excluded from consolidation scope to hide the outcome of certain transactions or risky assets and the contractual customization of leasing agreements to qualify for financial or operating lease accounting. Transaction structuring thus enables management to
manipulate earnings without explicitly breaching GAAP regulation and running the risk of adverse enforcement actions.

- **Altering Number of Shares Outstanding.** Managers may alter the number of outstanding shares to manipulate earnings on a per share basis. Reducing the number of shares (i.e., share buyback) increases EPS. Conversely, offering new shares results, all else equal, in a decrease of EPS.

### 2.3.3 Comparison of Earnings Management Techniques

Accrual-based and real earnings management differ in several aspects that determine their application in practice. Main differences are summarized in Table 2.1:

<table>
<thead>
<tr>
<th></th>
<th>Accrual-based Earnings Management</th>
<th>Real Earnings Management</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Timing</strong></td>
<td>after fiscal year-end</td>
<td>before fiscal year-end</td>
</tr>
<tr>
<td><strong>Costs</strong></td>
<td>low</td>
<td>high</td>
</tr>
<tr>
<td><strong>Visibility/Detection Risk</strong></td>
<td>moderate/high</td>
<td>low</td>
</tr>
<tr>
<td><strong>Affected Earnings Components</strong></td>
<td>Accruals</td>
<td>Accruals/Cash Flows</td>
</tr>
<tr>
<td><strong>Risk of Litigation</strong></td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td><strong>Constraints</strong></td>
<td>Prior Manipulations/GAAP/ Auditor &amp; Enforcement Bodies</td>
<td>Marginal Costs/Benefits</td>
</tr>
</tbody>
</table>

The table compares the characteristics of the two main types of earnings management examined in this study.

Timing relates to the question whether manipulation activities have to be initiated in advance or may be implemented on an ad-hoc basis. Accrual-based earnings management is generally regarded as a more flexible instrument in terms of manipulation timing, since it occurs during the preparation of the financial statements after the closing of the fiscal year (see, e.g., Roychowdhury, 2004; Gunny, 2010). In comparison, real earnings management needs to be initiated in advance. Price reductions to meet earnings or revenue targets, for example, have to be granted sufficiently early in the fiscal year to be efficient. Given the differences in timing, Roychowdhury (2004, p. 7) suggests a complementary strategy of both earnings management techniques: Real earnings management is used to steer towards relevant earnings targets during the year and, sub-
subsequently, accrual-based earnings management is applied to re-calibrate earnings after the fiscal year-end.

Costs of earnings management depend on the efforts management has to devote to discretionary activities and the dimension of negative consequences on current and future firm performance. Real earnings management is regarded as more costly than pure accrual manipulation (see, e.g., Roychowdhury, 2004; Graham et al., 2005; Ewert and Wagenhofer, 2005; Gunny, 2010). Altering real transactions for reasons other than increasing long-term firm value involves a deviation from “optimal” business practices (Ewert and Wagenhofer, 2005). Building up inventories, for instance, results in increased storage costs. Postponing R&D efforts, as another example, shifts the outcome of supposedly profitable projects into the future, reducing net present value of the firm’s investment set. Moreover, real earnings management involves more managerial effort than pure accrual manipulation. This includes, e.g., the planning of the respective transactions as well as communicating deviations from optimal business strategies.

Both types of earnings management face constraints that determine their intensity. The scrutiny of auditors and regulatory bodies is primarily aimed on the financial reporting process (i.e., accrual-based earnings management). Generally, GAAP defines the boundaries of “legal” earnings cosmetics and breaching these rules increases the risk of costly litigation and potential investor mistrust. Furthermore, accrual-based earnings management is limited by the level of previous manipulations (Barton and Simko, 2002): A prior overstatement of assets, for example, limits the opportunities to increase earnings by means of further overstatements. Real earnings management, in comparison, is less risky in terms of litigation (Gunny, 2010). Though managers can be sued for actively destroying long-term firm value, it is often very difficult to ascertain whether a specific real action is in fact attributable to strategic earnings considerations (Lev, 2003). A more important real earnings management constraint is thus given by the marginal costs and benefits of the respective action. Price reductions in the last quarter of the year, for example, are only effective if the additional sales contribute a positive margin to net income. Similarly, drastically deferring maintenance into the future can cause costly break-down of equipment.
2.4 Expectations Management

Expectations management is an alternative to earnings management when the company strives to beat market expectations. If analyst forecasts are above expected earnings, firms may either increase earnings by means of accrual-based or real earnings management or deliberately disclose information to guide analyst down to beatable targets. Voluntary disclosures include, for example, qualitative statements about the firm’s prospects or precise management forecasts (Nöldeke, 2007b). The trade off between expectations and earnings management is influenced by the respective institutional setting. Brown and Higgins (2005), for instance, argue that expectations management is more important than earnings management in countries with high investor protection, since strong regulation provides incentives to substitute costly earnings management with forecast guidance. Similarly, Koh et al. (2008) show that managers change from earnings to expectations management in times of increased auditor and regulator scrutiny or more rigorous enforcement. Given the drastic increase in regulatory constraints on earnings management, a growing body of empirical literature documents the importance of expectations management in financial reporting today.

7Amongst others, the most influential regulatory change addressing managerial discretion was SOX in the US, which was enacted as a response to Enron and other major accounting and corporate scandals such as Tyco International and WorldCom at the beginning of the 21st century. Section 3.2.2.4 provides a summary of studies that examine the influence of SOX on managerial discretion and earnings management.
Chapter 3

Earnings and Expectations Management to Achieve Benchmarks

The previous chapter explains the importance of earnings numbers in various kinds of economic transactions as well as the motivations and techniques of earnings management. In this chapter, I turn to focus on earnings management to meet or beat important earnings benchmarks. Section 3.1 provides theoretical explanations for benchmark beating and explains why managers or stakeholders may be fixated on earnings targets. In Section 3.2, I summarize previous empirical evidence of benchmark beating and earnings management. Supplementing this summary of empirical research, Section 3.3 presents survey results on benchmark beating and earnings management conducted among financial executives and analysts in the US and Europe. Providing the theoretical background and a review of prior evidence, this chapter builds the basis for developing a research outline in Chapter 4.

3.1 Theoretical Explanations for Benchmark Beating

Ample evidence in the accounting and finance literature supports that earnings are an important indicator of a firm’s current financial position and a key determinant in asset pricing. This relevance of earnings itself, however, does not explain the salience of earnings benchmarks. In their seminal contribution on benchmark-driven earnings management, Burgstahler and Dichev (1997) provide two explanations for benchmark importance, namely Kahneman and Tversky’s (1979) Prospect Theory and the role of transaction costs. Until today, most studies on benchmark
beating name these two as the most important theoretical explanations for the benchmark beating phenomenon.

### 3.1.1 Prospect Theory

Kahneman and Tversky’s (1979) Prospect Theory is a descriptive model of decision making under risk that addresses violations of classical expected utility theory. Under expected utility theory, the utility $U$ of two possible outcomes $x$ and $y$ is formally defined as (Kahneman and Tversky, 1979, pp. 263–264)

$$U(w + x, p; w + y, q) = pu(w + x) + qu(w + y),$$

(3.1.1)

where $p$ and $q$ denote the probabilities of outcome $x$ and $y$, respectively, and $w$ is the individual’s asset position. The function $u$ describes the concave utility function of a risk averse individual (i.e., $u'' < 0$).

Kahneman and Tversky (1979) show that expected utility theory is not capable of describing more complicated patterns of human decision making under risk and provide several experiments to lay the foundations of an alternative theoretical concept. As an example, consider the responses of 82 individuals to two simple choice problems summarized in Table 3.1. In every experiment, each individual may choose between two options (A or B and C or D). Under classical expected utility theory with $u(0) = 0$, the utility of options A, B, C, and D is given, respectively, as

$$U(A) = 0.33 \times u(2,500) + 0.66 \times u(2,400),$$

(3.1.2)

$$U(B) = u(2,400),$$

(3.1.3)

$$U(C) = 0.33 \times u(2,500),$$

(3.1.4)

and

$$U(D) = 0.34 \times u(2,400).$$

(3.1.5)

---

1For integrating insights from psychological research into economic science, Daniel Kahneman received the Nobel Memorial Prize in Economics in 2002.

2For details refer to Kahneman and Tversky (1979, pp. 265–266).
Given relative frequencies from both experiments, the individuals significantly prefer B to A (i.e., $U(B) > U(A)$) and C to D (i.e., $U(C) > U(D)$). Rearranging Eq. (3.1.2), the former relation becomes

$$0.34 \times u(2,400) > 0.33 \times u(2,500),$$

and the latter

$$0.34 \times u(2,400) < 0.33 \times u(2,500).$$

Since the inequality in Eq. (3.1.6) is the reverse of Eq. (3.1.7), expected utility theory is violated in this case. Kahneman and Tversky (1979) provide several additional examples for other situations not captured by expected utility theory and incorporate these experimental findings into a new theory.

### Table 3.1

**Prospect Theory Experiment**

<table>
<thead>
<tr>
<th>Experiment 1:</th>
<th></th>
<th>Experiment 2:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Option A:</strong></td>
<td>2,500 with a Chance of 0.33</td>
<td><strong>Option B:</strong></td>
<td>2,400 with Certainty</td>
</tr>
<tr>
<td></td>
<td>2,400 with a Chance of 0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 with a Chance of 0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Relative Frequency</strong></td>
<td>18%</td>
<td><strong>Relative Frequency</strong></td>
<td>82%</td>
</tr>
</tbody>
</table>

| **Option C:**       | 2,500 with a Chance of 0.33 | **Option D:**       | 2,400 with a Chance of 0.34 |
|                     | 0 with a Chance of 0.67     |                     | 0 with a Chance of 0.66     |
| **Relative Frequency** | 83%  | **Relative Frequency** | 17%  |

The table provides the results from two experiments on decision making under risk as documented in Kahneman and Tversky (1979). Total number of participants is 82. Each individual may choose between options A and B in Experiment 1 and C and D in Experiment 2. Relative frequencies denote how often a respective option was chosen. For details refer to Kahneman and Tversky (1979, pp. 265–266).

Under Prospect Theory, the utility function $u(w + x)$ is replaced by a value function $\nu(x)$, which reflects the subjective value of probable outcomes. Since individuals may weight probabilities differently, the probability $p$ of the classical model is replaced by a weighting function $\pi(p)$. Furthermore, $\pi(0) = 0$, $\nu(0) = 0$, $\pi(1) = 1$, and $\nu(1) = 1$ by definition. For two possible outcomes $x$ and $y$ with probabilities $p$ and $q$, expected utility $V$ is now given as
V(x, p; y, q) = \pi(p) \upsilon(x) + \pi(q) \upsilon(y) \tag{3.1.8}

if x is positive and y is negative or vice versa (i.e., x ≥ 0 ≥ y or x ≤ 0 ≤ y), and

V(x, p; y, q) = \upsilon(y) + \pi(p) \left[ \upsilon(x) - \upsilon(y) \right] \tag{3.1.9}

if both outcomes are either positive or negative (i.e., x > y > 0 or x < y < 0). The modified expected utility functions V incorporate the following deviations from classical expected utility theory (Kahneman and Tversky, 1979):

- **Reference Points.** Expected utility theory usually treats with levels of assets or wealth instead of their changes (i.e., the utility function covers wealth w and outcome x, y). Prospect Theory acknowledges that changes in assets are more important to individuals than levels and assumes the asset level as reference point from which changes in assets (i.e., gains and losses) are measured (i.e., the value function only considers outcomes x and y: \upsilon(x) and \upsilon(y)).

- **Risk avoidance vs. Risk seeking.** Expected utility theory assumes individuals to be risk averse. Consequently, the utility function u is concave for all x and y (i.e., u'' < 0). Prospect Theory partly relaxes the assumption of concavity and presumes that investors are risk averse when they face a gain and risk seeking when they face a loss. As a consequence, the value function is concave for all possible outcomes above (u'' < 0) and convex for all possible outcomes below a reference point (u'' > 0).

- **Loss Aversion.** In contrast to expected utility theory, Prospect Theory assumes that losing a certain amount of money is valued higher than gaining the identical amount and this effect is supposed to increase with the amount of the bet. As a result, a purely symmetrical bet (e.g., win 50 or lose 50 with a probability of \frac{1}{2}) is perceived as unfavorable for most people. Prospect Theory incorporates this behavior by allowing the value function to be steeper for losses than for gains.

- **Weighting Probabilities.** Expected utility theory does not incorporate subjective weightings of probabilities. Kahneman and Tversky (1979), however, suggest that individuals overweight small and underweight larger probabilities. Prospect Theory acknowledges
this assumption by introducing a weighting function $\pi(p)$ that overweights small probabilities (i.e., $\pi(p) > p$ for small $p$) and underweights large probabilities (i.e., $\pi(p) < p$ for large $p$).

Typical value and utility functions under Prospect Theory are depicted in Figure 3.1. Indicating asymmetric risk preference above and below the reference point, the value function in Panel A is clearly concave for gains and convex for losses. Additionally, loss aversion induces a steeper curve for losses than for gains. The utility function plot in Panel B illustrates that small probabilities are overweighted, while large probabilities are underweighted.

![Value and Weighting Functions under Prospect Theory](image)

**FIG. 3.1.—Value and Weighting Functions under Prospect Theory.** The figure illustrates a typical value (Panel A) and utility weighting function (Panel B) under the axioms of Prospect Theory. Similar illustrations can be found in Kahneman and Tversky (1979, pp. 279 and 283).

The basic principles of Prospect Theory explain the importance of simple earnings benchmarks. In the earnings management context, benchmarks are considered as reference points of managers’ and/or investors’ value functions. The steepest point of the value function is directly at the benchmark, where a marginal increase in earnings yields the maximum increase in value. Consequently, the incentives for earnings management are supposed to be strongest around the respective threshold. If last year’s earnings, for instance, are considered as an important refer-

---

3For a detailed description of Prospect Theory in the context of earnings benchmarks, see Burgstahler and Dichev (1997, pp.121–124).
ence point for investors and/or executives, managing earnings from barely below to just above the target will yield the largest gain in terms of utility. Prospect Theory thus provides a reasonable theoretical framework for the importance of earnings benchmarks.

### 3.1.2 Transaction Costs

Burgstahler and Dichev (1997) name transaction costs as a second explanation for the importance of earnings benchmarks. Transactions thereby encompass all implicit and explicit types of transactions with the stakeholders of the firm. Beside investors, these include, e.g., customers, suppliers, debt-holders, and employees. The transaction cost explanation grounds on two assumptions:

- **The relation of earnings and transaction terms.** Higher earnings are generally regarded as being positively related to favorable transactions terms. In other words, firms with higher earnings enjoy more favorable transaction terms than those with lower earnings. Relevant examples are reported in, e.g., Bowen et al. (1995) and Burgstahler and Dichev (1997): Customers pay higher prices for products because they trust more in warranty and service commitments, suppliers offer better prices due to repeated and larger orders in the future and/or lower default risks, lenders offer better contract terms because loan defaults or delayed payments occur less likely, and/or important employees are more likely to stay with the company when it is profitable.

- **Information Cost.** Retrieving, storing, and processing information is so expensive that at least some stakeholders base terms of transactions on simple heuristics, such as last year’s or positive earnings. Consider, for instance, an industrial supplier of many different production firms. Instead of analyzing every customer’s financial report in detail, executives can base contractual terms on simple heuristics. Sustainable earnings growth or constantly positive earnings are often considered as signal for financial health and rewarded with more favorable prices and/or credit terms in this context.

Taken together, the transaction cost explanation assumes that stakeholders evaluate financial health and prospects on the basis of simple heuristics. These include, e.g., reporting profits,

---

4For a detailed description of their transaction cost explanation, see Burgstahler and Dichev (1997), specifically pp. 121–124.
3.2 Empirical Evidence of Benchmark Beating

Burgstahler and Dichev (1997) were among the first to explicitly analyze the relation between simple earnings benchmarks and earnings management. Their seminal contribution raised tremendous interest in earnings management to meet benchmarks and spawned a huge body of empirical literature constantly growing until today. This section provides a comprehensive review of previous studies in the field. The review starts with Burgstahler and Dichev's (1997) landmark paper, summarizes main findings from related subsequent research, and discusses the growing criticism of the Burgstahler and Dichev (1997)-type studies in Section 3.2.1. Section 3.2.2 reviews papers that examine specific types of manipulation at relevant earnings thresholds. These studies directly address different types of accrual-based and real earnings management techniques, the trade-off between these techniques, and how benchmark beating and earnings management affect subsequent firm performance. Eventually, Section 3.2.3 summarizes evidence from studies that empirically address incentives for earnings management at benchmarks.

3.2.1 Frequency of Earnings and Expectations Management

In her study on the information content of losses, Hayn (1995) analyzes the distribution of pricescaled EPS. Groundbreaking for intensive subsequent research on benchmark related earnings management, she remarks (Hayn, 1995, p. 132):

Interestingly, there is a point of discontinuity around zero. Specifically, there is a concentration of cases just above zero, while there are fewer than expected cases (assuming the above normal distribution) of small losses (i.e., just below zero). The frequency of observations in both the region just above and that just below zero departs significantly from the expected frequency under the normal distribution at the 1% significance level using the binomial test. These results suggest that firms whose earnings are expected to fall just below the zero earnings point engage in earnings manipulations to help them cross the “red line” for the year.
Two years later, Burgstahler and Dichev (1997) pick up Hayn’s (1995) finding and hypothesize that managers face incentives to avoid losses and earnings declines and, as a result, engage in earnings management to avoid missing these earnings benchmarks. To test for earnings management and benchmark relevance, Burgstahler and Dichev (1997) examine the distribution of earnings levels and changes in earnings, both scaled by lagged market value of equity. They document a significant drop of observations immediately below and an unusual pile-up directly above the respective thresholds (i.e., zero earnings and zero earnings changes) and interpret these irregularities in otherwise smooth distributions as evidence of earnings management. Degeorge et al. (1999) extend their analysis with analyst forecasts as additional benchmark that triggers threshold related earnings management. Instead of price-scaled annual earnings, however, they analyze unscaled per share data on a quarterly basis. Their results confirm Burgstahler and Dichev’s (1997) findings. In addition, they detect a similar discontinuity in the distribution of earnings surprises (i.e., actual EPS minus the latest EPS consensus estimate); a finding pointing to earnings or expectations management to achieve analysts’ earnings estimates.

The distributional approach applied in Burgstahler and Dichev (1997) and Degeorge et al. (1999) paved the way for a large body of research on earnings management and earnings benchmarks. In the remainder of this section, I summarize Burgstahler and Dichev (1997)-type studies of benchmark related earnings management. The review considers different geographical regions and legal environments in Section 3.2.1.1 and the evolution of benchmark beating over time in Section 3.2.1.2. Since the growing popularity of earnings management research around benchmarks increased skepticism towards interpreting distributional irregularities as evidence of earnings management, critical studies are discussed in Section 3.2.1.3.

### 3.2.1.1 Non-US Evidence and International Differences

Starting with Ball et al. (2000) and Ball et al. (2003), research in accounting has emphasized that albeit the ongoing harmonization of international accounting rules, financial reporting practice is influenced by the respective institutional setting. This strand of research indicates that differences...
in actual reporting behavior are endogenous and determined by real economic and political factors that differ among countries (Ball, 2006). Those factors include the legal environment (e.g., legal system, legal enforcement), capital market importance and structure (e.g., size, ownership, investor rights), and financial reporting regulation (e.g., standard setting process, tax alignment, level of enforcement). The influence of economic and political factors on reporting practice provoked several authors to test the validity of Burgstahler and Dichev’s (1997) and Degeorge et al.’s (1999) results in different institutional settings. Among others, studies of non-US data include Holland and Ramsay (2003) (Australia), Holzapfel (2004) (Germany), Suda and Shuto (2005) (Japan), Nöldeke (2007a) (Germany), and Charoenwong and Jiraporn (2009) (Thailand and Singapore). Overall, these studies document irregularities in the distributions of earnings, earnings changes, and earnings surprises and generally suggest benchmark related earnings management in non-US settings.

Though earnings management seems to be an international phenomenon, incentives to achieve targets and the intensity of earnings management may vary. Several authors use the context of benchmark-driven earnings management to shed light on the influence of institutional factors on financial reporting and compare benchmark-driven earnings management in different international settings. In one of the first studies, Brown and Higgins (2001) contrast benchmark beating in the US with 12 other countries (including Germany). Drawing on differences in investor protection and corporate governance, they hypothesize that US managers engage more intensively in earnings or expectations management to avoid missing analyst estimates than their counterparts from non-US countries. A comparison of small positive and small negative earnings surprises confirms that US managers are indeed more suspicious of manipulating earnings or analysts than managers from other countries. Leuz et al. (2003) conduct a similar analysis for the zero earnings benchmark with a set of 31 countries. Their results suggest that earnings management in general and loss avoidance in particular are less pronounced in the US than in Europe or Asia. The results of Brown and Higgins (2001) and Leuz et al. (2003) are diverging on the first sight: The former suggest a higher and the latter a lower earnings management intensity in the US. In a subsequent study, however, Brown and Higgins (2005) show that US managers primarily engage in analyst guidance. US investor protection and corporate governance mechanisms thus seem to curb earnings management but not earnings guidance. Glaum et al. (2004) examine the distri-

---

6For a detailed discussion of corporate governance mechanisms, investor protection, and financial markets in an international context see, e.g., Shleifer and Vishny (1997) and La Porta et al. (2000).

7Among other proxies, Leuz et al. (2003) use loss avoidance as one measure of overall earnings management activity. Their loss avoidance measure is calculated as ratio of small positive to small negative earnings.
butions of earnings, earnings changes, and earnings surprises to test for differences in earnings management activity in the US and Germany. Contradicting with Leuz et al. (2003), they do not find significant differences in earnings management to avoid losses and earnings decreases. For the analyst forecast benchmark, however, they report that US managers are more concerned about meeting analyst expectations than German executives; a result consistent with previous results in Brown and Higgins (2001). Glaum et al. (2004) attribute this difference to stronger capital market forces and higher equity compensation for US managers. In a subsequent study, Daske et al. (2006) compare the distributional patterns around earnings benchmarks in the US and Europe. In support of the earnings management hypothesis, they document significant irregularities in the distributions of (scaled) earnings, earnings changes, and earnings surprises for all 13 EU member states under consideration. Furthermore, these irregularities seem to be stronger for EU countries when compared with their US counterparts, especially for the zero earnings and the earnings changes benchmark. Eventually, Burgstahler et al. (2006) examine the influence of legal systems and enforcement mechanisms on earnings management of public and private firms in 13 European countries. Their findings suggest that earnings management (among other factors measured as the ratio of small profits to small losses) in private and public firms decreases with stricter legal systems, power of enforcement, and (especially for private firms) weak book-tax alignment.

3.2.1.2 Temporal Shifts in Benchmark Beating

The evidence in Section 3.2.1.1 suggests that earnings management intensity depends on the underlying institutional setting. If characteristics of institutional settings shape financial reporting practices, then important changes in these characteristics may influence earnings management activity. Two research questions are common in this context: 1) Did the relative importance of earnings benchmarks change over time?; and 2) Do regulatory changes affect earnings management intensity?

Several studies examine the importance of one or several earnings benchmarks over time. Brown (2001), for instance, analyzes the importance of the analyst forecast benchmark for a large sample of firms from 1984 to 1999 and documents a considerable increase of observations with zero or positive earnings surprises over the sample period. Similarly, Matsumoto (2002) reports a significant increase of observations that beat quarterly earnings forecasts from 1985 to 1997. Both findings suggest rising interest in meeting or beating analysts’ expectations. Dechow et al. (2003) examine the benchmark hierarchy over time. For the period from 1989 to 2001,
they find a gradual decrease in the propensity to meet the zero earnings and earnings changes benchmark, while the propensity to achieve analyst forecasts increases significantly. Brown and Caylor (2005) use the Burgstahler and Dichev (1997) methodology to identify possible shifts in the relative importance of earnings benchmarks from 1985 to 2002. They document that loss avoidance is more relevant than avoiding earnings declines; a relation remaining relatively constant over the whole sample period. The importance of analyst forecasts, in contrast, increases significantly over time: Being the least important benchmark in the late 1980s and early 1990s, managers’ increasing attention to analyst estimates makes them the most important benchmark from the mid-1990s on. The authors attribute this shift to higher analyst coverage, increased media attention to analysts, improved accuracy and precision of estimates, and (as a result) a stronger market reaction to benchmark achievement in later years.

The second research question directly addresses specific regulatory changes that are assumed to curb earnings management. Altamuro et al. (2005), for example, analyze the distribution of earnings and earnings changes to test whether the clarification of revenue recognition principles by the Securities and Exchange Commission (SEC) in 1999 affected the level of earnings management. Confirming the influence of regulatory change, they show that stricter revenue recognition principles decreased earnings management to avoid losses and earnings declines. Koh et al. (2008) examine whether the introduction of SOX in 2002, a direct reaction of the US Congress to previous accounting scandals at Enron, WorldCom, and other US firms, changed earnings management behavior. They show that the fraction of small positive earnings surprises decreased significantly with the introduction of SOX. This decrease is consistent with more skeptical investors and lower market premiums for benchmark achievement in the post-scandals period. Jeanjean and Stolowy (2008) explore whether the introduction of IFRS affected the level of earnings management and analyze the distribution of earnings for the years 2002 to 2006 in three countries, namely UK, France, and Australia. They provide evidence that the introduction of IFRS did not generally decrease the level of earnings management and conclude that a unique set of accounting rules is necessary, but not sufficient to harmonize international financial reporting.

### 3.2.1.3 Critical Studies

The rapid spread of Burgstahler and Dichev (1997)-type studies raised skepticism in the research community and provoked scholars to “dig deeper” behind the factors that cause kinks earnings distributions. Dechow et al. (2003), for example, provide several alternative explanations. These
include real performance increases when earnings get close to important targets, sample selection bias due to listing requirements of public firms, distortion of the earnings distribution due to scaling, and accounting conservatism. Durtschi and Easton (2005) focus on the consequences of scaling and sample selection and reinforce some of Dechow et al.’s (2003) arguments with detailed empirical evidence. Specifically, they show that deflating earnings with lagged market value distorts the distribution of earnings, since profit firms are valued differently than loss firms. This deflator asymmetry induces a pile-up of profit observations near the threshold and causes a kink in the distribution of earnings. As documented in a later paper (Durtschi and Easton, 2009), this distortive effect is not eliminated when lagged market value is replaced by other scaling variables (e.g., sales, total assets, or equity). Another critical argument addressed in Durtschi and Easton (2005, 2009) is sample selection. They show that additional data requirements (e.g., lagged market value) cause an irregularity in the earnings distribution, since data availability is higher for profit than for loss firms. Beaver et al. (2007) provide another alternative explanation and show that systematic differences in effective tax rates and special items for profit and loss firms amplify the kink in the distribution of earnings. Specifically, higher effective tax rates of profit firms and more frequent negative special items of loss firms move profit observations closer to the benchmark at zero earnings and loss observations into the distribution’s left tail. Overall, these critical studies caution researchers to attribute discontinuities in earnings distributions to earnings management and emphasize that “[...] other possible explanations must be ruled out before researchers can confidently make the claim that the shapes of these distributions [...] are evidence of earnings management” (Durtschi and Easton, 2005, p. 561).

3.2.2 Earnings Management Techniques

A drawback of the Burgstahler and Dichev (1997)-type studies on threshold related earnings management is that they neither measure the magnitude of earnings management, nor provide details about manipulation methods (Healy and Wahlen, 1999). Moreover, the distributional approach may capture other effects than earnings management that contribute to a kink in the earnings distribution (see Section 3.2.1.3).

If distributional irregularities are caused by discretion in the financial reporting process or

---

8Noteworthy in this context, Burgstahler and Dichev (1997) report that earnings are manipulated by means of working capital adjustments and change in CFO. However, they do not provide any tests based on solid measures of abnormal accruals and/or real earnings management.
real activities manipulation, an adequate proxy of these techniques should capture this effect. Accounting research has developed a large number of alternative earnings management proxies for different manipulation techniques. These include measures of aggregate and specific discretionary accruals (e.g., Jones, 1991; Beatty et al., 2002), real earnings management (e.g., Roychowdhury, 2006; Gunny, 2010), and analyst guidance (e.g., Matsumoto, 2002; Bartov et al., 2002). This section summarizes the body of literature that uses these proxies to avoid caveats of the distributional methodology and examines the application specific earnings management techniques.

3.2.2.1 Accrual-based Earnings Management

The part of total accruals suspected to result from earnings management activity is generally referred to as discretionary accruals. Several measures of discretionary accruals have been developed in the field of earnings management research. Basically, these can be divided in two categories: Measures of aggregate discretionary accruals (Section 3.2.2.1.1) and measures of specific discretionary accruals (Section 3.2.2.1.2).

3.2.2.1.1 Aggregate Accruals Management

Following early aggregate abnormal accrual models (e.g., Healy, 1985; DeAngelo, 1986), unexpected aggregate accruals are nowadays estimated as residuals from cross-sectional Jones (1991)-type regressions of total or working capital accruals on firm-specific factors expected to affect the level of accruals (e.g., sales, level of property, plant, and equipment).⁹ Dechow et al. (2003) explicitly challenge the claim that earnings management is responsible for the discontinuity in the distribution of earnings and examine the level of abnormal accruals for firms in the intervals adjacent to the zero earnings benchmark. Consistent with earnings management, firms that barely beat the zero earnings benchmark report higher positive discretionary accruals than all remaining observations. Small loss firms, however, exhibit a similar level of discretionary accruals as small profit firms; a finding in strong contradiction with the earnings management hypothesis. Coulton et al. (2005) report qualitatively similar results for a sample of Australian firms: Though larger than the rest of the sample, positive discretionary accruals of benchmark beaters

⁹A detailed description of Jones (1991)-type regression models to measure discretionary accruals can be found in the empirical part of the dissertation (Section 7.2.1.1).
are not significantly larger than those of benchmark missers. Matsumoto (2002) uses a logistic regression approach to test whether discretionary accruals increase the probability of meeting or beating analyst expectations, but her results do not indicate income-increasing accruals management. Athanasakou et al. (2009) use a similar approach for a UK sample. Consistent with Matsumoto (2002), however, they fail to detect a positive relation between the probability of forecast achievement and discretionary working capital accruals. A different approach is chosen by Ayers et al. (2006), who use a “pseudo target”-analysis to test whether the distributional kink is attributable to accrual-based earnings management. If managers engage in earnings management to achieve benchmarks, then the positive relation of discretionary accruals and benchmark beating should either not hold or be considerably weaker for other (random) points in the distribution. Conflicting with the earnings management hypothesis, they find a positive relation of “pseudo targets” and discretionary accruals in the distributions of earnings and earnings changes. These relations, however, intensify at the “actual targets”. Overall, Ayers et al. (2006) are not able to rule out the possibility that the irregularities in the distributions of earnings and earnings changes are attributable to other factors than earnings management. In contrast to these studies, Peasnell et al. (2000a) report different results for a UK sample. They document that managers increase discretionary accruals to achieve target when premanaged earnings fall slightly short of the zero earnings and earnings changes benchmark. Based on logistic regressions, Lin et al. (2006) also report supporting evidence for accrual-based earnings management to achieve the consensus estimate. Specifically, they document that positive discretionary accruals significantly increase the chance of achieving the analyst forecast benchmark.

The summarized results indicate that the relation between earnings management and benchmark beating is not as clear-cut as postulated in academic research and the financial press. Generally, the results may be ambiguous for two reasons: Either the applied accrual models are too weak to measure earnings management activity, or accrual-based earnings management is a myth.

### 3.2.2.1.2 Specific Accruals Management

Healy and Wahlen (1999) suspect aggregate accrual proxies to lack accuracy and encourage the research community to focus on specific discretionary accruals. A growing number of authors followed their advise and developed models that capture specific discretionary accruals.

One stream of research suggests that tax accounts provide a convenient opportunity for “last
3. Earnings and Expectations Management to Achieve Benchmarks

minute"-earnings management. Frank and Rego (2006), for example, report that public firms use the deferred tax asset valuation allowance to achieve the consensus forecast. The same result does, however, not hold for the zero earnings and the earnings changes threshold. In a similar vein but focusing on tax expenses as a whole, Dhaliwal et al. (2004) find that managers decrease tax expenses if they would otherwise just miss the analyst consensus estimate. Eventually, Myers et al. (2007) document that firms use total tax expense to achieve the earnings changes benchmark.

Another stream of research suspects receivables, loans, and their related profit and loss accounts as potential earnings management candidates. Caylor (2010), for example, focuses on revenue management to meet or beat earnings targets and develops a model for abnormal trade receivables and deferred revenues. He reports that managers use both accounts to avoid negative earnings surprises. Beatty et al. (2002) focus on the banking sector and find that public banks use loan loss provisions and security gain realizations to avoid small earnings decreases. Beaver et al. (2003) consider allowance accounts in the property-casualty insurance industry and find that firms significantly understate the loss reserve to avoid reporting negative net income. In a more recent study, Jackson and Liu (2010) show that managers reverse prior overstatements of the bad debt allowance account to increase earnings and avoid falling short of analyst expectations.

Studies such as Marquardt and Wiedman (2004) provide a more comprehensive analysis and test for abnormal accounts receivable, inventories, accounts payable, accrued liabilities, depreciation, and special items. Their results reveal that managers discretionarily adjust special items to avoid falling short of prior year’s earnings, while none of the other accruals exhibit the typical patterns expected under the earnings management hypothesis. Plummer and Mest (2001) also conduct an analysis for a large set of specific accruals. In the spirit of Burgstahler and Dichev (1997), the authors plot the distributions of analyst forecast errors in terms of sales, operating, non-operating, and depreciation expenses for suspect firms that just barely meet the EPS consensus forecast. They find significant irregularities in the distribution of sales and operating expenses and conclude that firms use these components to manage earnings above target.

3.2.2.2 Real Earnings Management

Another potential explanation for the ambiguous results concerning accrual-based earnings management is that managers engage in real earnings management to achieve earnings targets (Roychowdhury, 2006). Real earnings management studies consider either income generation and
acceleration techniques (e.g., channel stuffing, fixed asset sales) or cost cutting techniques (e.g., cutting R&D or SG&A expenses).

3.2.2.2.1 Boosting Revenues and Other Income

Revenues can be subject to accrual-based and/or real earnings management activity. Accrual-based revenue management includes, for example, the premature recognition of revenues (see, e.g., Stubben, 2010). Real revenue management, in contrast, encompasses all activities that generate additional revenues (e.g., by granting discounts or more lenient credit terms) or accelerate the recognition of unrealized revenues (e.g., increased delivery, discretion in contract design) (Roychowdhury, 2006; Lev, 2003). Although real revenue management is suggested to be a common tool for earnings manipulation, most studies focus on “accounting maneuvers” that pull sales forward in time (Das et al., 2011b). In one of the few studies on real revenue management, Roychowdhury (2006) assumes that managing revenues with real activities yields, relative to sales, abnormally low cash flow from operations and abnormally high production cost (i.e., COGS plus change in inventories).10 His results indicate that barely meeting or beating the zero earnings or analyst forecast benchmark is, as expected, related with abnormal low CFO and abnormal high production costs (relative to sales). He therefore concludes that managers seem to engage in revenue management to beat these earnings targets.

Timing of asset sales is another instrument to push earnings towards targets. Bartov (1993) is among the first studies that explicitly address earnings management by means of asset sales. He finds that managers manipulate reported earnings through the timing of asset sales to minimize intertemporal earnings changes. That is, firms sell additional fixed assets when current year earnings are below, and delay planned asset sales until the next period when current year earnings are above prior year’s earnings. Furthermore, managing earnings upward is more pronounced than managing downward. This finding explicitly supports the hypothesis that managers use the timing of asset sales to avoid falling short of last year’s earnings. Black et al. (1998) extend Bartov’s (1993) study by comparing sales timing in countries with different accounting treatment of asset sales (Australia/New Zealand and pre-FRS 3 UK). Specifically, they hypothesize that earnings smoothing by means of asset sales decreases when GAAP allows upward asset reval-

10Offering discounts to increase sales lowers the profit margin of the product and thereby increases production costs relative to sales. More lenient credit terms lead to additional sales on credit and thereby lower levels of CFO relative to sales. For further details on the proxies for real revenue manipulation, see Roychowdhury (2006, p. 340).
uations. They find that income smoothing via sales timing is not existent in regimes that allow revaluation and common in regimes that stick to historical cost accounting. In a related study of Singaporean firms, Poitras et al. (2002) report that managers sell fixed assets to avoid earnings declines. Eventually, Gunny (2010) provides a model to measure the unexpected portion of income from fixed asset and investment sales. Her results do, however, not support the earnings management hypothesis: Meeting the zero earnings and/or earnings changes benchmark is not associated with abnormally high gains from asset sales.

### 3.2.2.2 Cutting Costs

As an alternative to real income management, managers can deliberately defer current expenses to future periods. Typical examples in this context are cuts in R&D and SG&A expenditures or overproduction as a means of lowering COGS.

In an early study, Baber et al. (1991) examine whether firms cut R&D expenses to avoid missing earnings benchmarks. They show that managers forfeit positive net present value R&D investments to avoid reporting a loss or earnings declines. Perry and Grinaker (1994) extend the Baber et al. (1991) study and explore whether falling short of earnings expectations affects R&D investment decisions. While Baber et al. (1991) assume that R&D expenses follow a random walk, Perry and Williams (1994) apply a more sophisticated model to measure (un)expected R&D activity that controls for several firm-specific and macro-economic factors. Their results confirm that managers deliberately cut R&D expenditures to meet earnings expectations. Subsequent studies on R&D investments and earnings targets generally confirm the results of Baber et al. (1991) and Perry and Grinaker (1994). Bange and De Bondt (1998), for example, find that firms adjust R&D expenses to decrease the gap between reported and expected earnings especially when stock prices are volatile and trading volume is high. Incentives for discretion in R&D spending, however, vanishes when upper management and/or institutional investors hold a large fraction of the company. In a similar vein, Bushee (1998) examines whether large institutional investor holdings curb discretionary R&D spending to avoid earnings declines. He finds R&D expense cutting to avoid earnings declines to be less prevalent in the presence of institutional investors. However, this effect is weakened when institutional investors have a high portfolio turnover or engage in momentum trading. Cheng (2004) considers R&D cutting to avoid losses or earnings declines with respect to CEO compensation and finds that compensation committees acknowledge the negative consequences of discretionary R&D cutting and design compensation contracts accordingly. More recently, Gunny (2010) provides a new measure of unexpected
Overproduction is supposed to be a convenient real earnings management tool to decrease COGS (Roychowdhury, 2006). Increasing the level of production decreases, all else equal, the fraction of fixed costs allocated to each produced unit and thus total COGS. Noteworthy, increasing production is effective only if the savings from lower fixed cost per unit is not fully absorbed by increased inventory costs (Roychowdhury, 2006). In his study on real earnings management, Roychowdhury (2006) measures overproduction as the unexpected portion of total production costs (i.e., COGS plus change in inventories) and provides evidence that managers increase production levels to achieve the zero earnings benchmark.\(^{11}\) Lin et al. (2006) analyze whether managers engage in overproduction to achieve analyst forecasts. They show that positive abnormal production decreases the probability of meeting or beating forecasts. Hence, managers seem to be rather cautious about using overproduction as a means of managing earnings towards analysts’ estimates. Cook et al. (2009) consider the case of production management for firms that would otherwise miss benchmarks and firms that significantly exceed benchmarks without manipulation. They find that managers increase production to lower COGS and avoid missing quarterly consensus forecasts. Firms that genuinely exceed analyst forecasts, however, seem to decrease inventories to build “cookie jars” and facilitate earnings management in later periods. Eventually, Gunny (2010) also considers overproduction as a real earnings management device and refines Roychowdhury’s (2006) model to measure (un)expected production levels. In support of Roychowdhury (2006), she confirms that just beating the zero earnings and/or earnings changes benchmark is positively related with abnormal levels of production costs.

Several authors focus on advertising expenses, SG&A costs, or aggregated expense accounts to test for real earnings management by means of cost-cutting. Roychowdhury (2006) provides a model to estimate the unexpected portion of aggregated R&D, advertising, and SG&A expenses. He finds that managers reduce costs to avoid reporting a loss and missing the latest analyst consensus. Gunny (2010) considers SG&A without aggregation. She finds that managers decrease SG&A costs to beat the zero earnings or earnings changes benchmark. Lin et al. (2006) follow the approach proposed in an earlier version of Gunny (2010) to measure SG&A cutting. In support of real earnings management, they find that cutting SG&A expenses increases the probability of meeting or beating analyst forecasts. Cohen et al. (2010a) draw on a unique database

\(^{11}\)Using the sum of COGS and change in inventories avoids capturing effects from inventory related accrual-based earnings management (e.g., delayed inventory writeoffs). See Roychowdhury (2006, p. 339) for details.
of monthly advertising expenditures to capture discretionary expense cutting activities around quarterly benchmarks. They report that managers cut advertising spending to avoid missing the zero earnings or earnings changes benchmark. Moreover, some firms seem to follow different advertising strategies with respect to earnings benchmarks: While more mature firms increase advertising to inflate short-term sales, their younger counterparts raise advertising activities in the third quarter to lower the quarterly earnings changes threshold at year-end.

3.2.2.3 Expectations Management

Expectations management provides an adequate alternative to earnings management when the threshold of interest is analyst forecasts. Matsumoto (2002) provides a unique model to measure the unexpected portion of analyst forecasts. If managers guide analysts towards beatable targets, then the unexpected portion of forecasts is expected to be negative (i.e., analyst forecasts are lower than expected). Her results reveal that the probability of meeting expectations is significantly higher in firm-quarters with negative unexpected forecasts. Hence, managers effectively use analyst guidance to avoid investor disappointment. Bartov et al. (2002) apply a different methodology and examine forecast revisions during the year to identify expectations management. Their results support Matsumoto (2002) and suggest that managers guide analysts from higher initial estimates down to beatable targets. Burgstahler and Eames (2006) combine both methodologies and examine unexpected forecasts (Matsumoto, 2002) and forecast revisions (Bartov et al., 2002) for ten narrow earnings surprise intervals surrounding the analyst forecast benchmark. Their findings are somewhat ambiguous: While forecast revisions suggest earnings guidance to achieve the analyst consensus, the analysis of unexpected forecasts does not. Eventually, Nöldeke (2007b) provides results for three European countries: Austria, Germany, and Switzerland. Analyzing the effect of management forecasts on analyst expectations, she shows that managers use forecasting to lower the analyst consensus estimate. Results from a survey among 85 CFOs/IROs and 82 financial analysts underpin her empirical results.

3.2.2.4 Competing Strategies

Following a large number of studies on accrual-based earnings management, real earnings management, and expectations management, a growing number of recent studies addresses the trade off between different techniques. Lin et al. (2006), for example, examine which tools are used by managers to achieve analyst forecasts. Their comprehensive set of manipulation tools covers
Earnings and Expectations Management to Achieve Benchmarks

Expectations management, discretionary accruals, classification shifting, and real activity manipulation. The latter includes SG&A expense cutting, overproduction, and temporary increases in sales. In summary, the results suggest that managers use expectations management, discretionary accruals, classification shifting, and cuts in SG&A expenses to avoid missing analysts’ earnings forecasts. Athanasakou et al. (2009) examine whether UK firms engage in earnings or expectations management to meet analyst expectations. Specifically, they consider two earnings management mechanisms: discretionary working capital accruals and classification shifting. They find that UK firms actively guide analysts to beatable targets. Furthermore, larger firms seem to shift core- to non-recurring expenses to avoid falling short of analysts’ estimates. Ewert and Wagenhofer (2005) show analytically that tighter accounting standards increase earnings quality and thus the marginal benefit from earnings management activities. As a result, firms are assumed to increase more costly real earnings management activity when accounting standards are tightened. Studying the trade off between accrual-based and real earnings management pre- and post-SOX, Cohen et al. (2008) underpin Ewert and Wagenhofer’s (2005) argument empirically: The use of discretionary accruals to achieve benchmarks dropped significantly, while the use of real earnings management increased with the introduction of SOX. As suggested by Ewert and Wagenhofer (2005), tightening accounting regulation does not prevent firms from managing earnings. Rather, the introduction of SOX induced a shift from cheaper and more visible accruals management to real earnings management, which is more costly but harder to detect. In a similar vein, Koh et al. (2008) analyze how the introduction of SOX affected the trade off between earnings and expectations management. They find that SOX generally decreased managerial actions to meet analysts’ earnings estimates. Furthermore, managers seem to substitute earnings management with less risky analyst guidance; a result most likely attributable to stricter enforcement and increased risk of litigation in the post-scandals period. A recent study by McInnis and Collins (2011) tests whether managers shift from accrual-based to real earnings management when analysts provide operating cash flow forecasts for the firm. Their reasoning is straightforward: Since cash flow and earnings forecasts implicitly provide accrual forecasts, accrual manipulations become more visible. Hence, managers replace accrual-based earnings management with more costly, but less transparent real earnings management. Their analysis confirms that the presence of cash flow estimates improves the quality of accruals and decreases the probability of achieving analyst expectations. However, at least some accrual manipulation is substituted with real earnings management or earnings guidance. Eventually, Das et al. (2011a) specifically address the costs of earnings management and analyst guidance and examine whether these tools are used as substitutes or complementary. The results suggest that both types of manipulation are used complementary. Increasing cost of earnings management, however, causes managers to
substitute earnings management with expectations management.

Overall, the summarized studies indicate that managers weigh advantages and disadvantages of possible manipulation techniques. Moreover and important for governments and standard setting bodies, drastic regulatory changes may significantly affect the degree and type of manipulation.

### 3.2.3 Incentives for Benchmark Beating

A huge body of empirical literature examines the prevalence and types of earnings and expectations management to achieve benchmarks. A comparably small fraction of these studies, however, directly address the question why managers seek to beat earnings benchmarks. Possible explanations for earnings management are capital market incentives, contracting, and political costs (see Section 3.2.3). This section summarizes their key findings.

#### 3.2.3.1 Capital Market Incentives

The literature on benchmark beating mainly considers capital market incentives as explanations for benchmark beating. This section reviews the evidence of stock market rewards (penalties) for beating (missing) earnings benchmarks and discusses studies on investor rationality, managerial opportunism, and signaling.

#### 3.2.3.1.1 The Market Reaction to Benchmark Beating

Among the first studies addressing market rewards for benchmark achievement were Lopez and Rees (2002), Bartov et al. (2002), and Kasznik and McNichols (2002). Lopez and Rees (2002) analyze whether, controlling for the information in current earnings, meeting or beating analyst expectations yields positive abnormal returns in a short window of three days around the earnings announcement date.12 They report that—irrespective of the earnings surprise level—beating the

---

12They control for the information content of earnings by including earnings surprises in the regression of abnormal returns on dummy variables for benchmark achievement. Earnings surprise is calculated as the difference between actual EPS and the latest analyst consensus EPS forecast issued prior to the earnings release of the respective quarter and scaled by beginning of the quarter stock price.
earnings benchmark is rewarded with a premium of 0.7%, while missing the threshold is penalized with an incremental abnormal return of −0.8%. Furthermore, they find that the marginal market reaction to one additional unit of earnings surprise (i.e., the earnings response coefficient, ERC) to be asymmetrical for observations below and above the threshold. Specifically, the ERCs for benchmark beaters are significantly larger than those of benchmark missers. Bartov et al. (2002) generally confirm Lopez and Rees’s (2002) findings for a longer return window. Irrespective of the earnings surprise level, beating market expectations yields an additional abnormal return of 2.3% over the current quarter. Furthermore, the ERCs of benchmark beaters are significantly greater than those of benchmark missers. Kasznik and McNichols (2002) examine abnormal returns for meeting analyst forecasts over a one-year return period. Controlling for the information in earnings during the year (i.e., reported EPS − beginning of the year forecast), benchmark beaters earn an incremental abnormal return of 8.0%. This incremental premium, however, decreases when expectations are met consecutively: Achieving the benchmark for the second and third time in a row yields a premium of 5% and 3%, respectively. These results are consistent with Lopez and Rees (2002), who argue that the market recognizes and partially discounts the systematic or recurring portion of positive earnings surprises.

Several subsequent studies extend these early papers. Lin et al. (2006), for example, study the short-term economic consequences of expectations and earnings management to achieve analyst forecasts. They find that expectations guidance, boosting sales, cutting SG&A, and classification shifting decrease the cumulated three-day premium associated benchmark achievement. Discretionary accruals and overproduction, in contrast, do not affect the benchmark premium. For a UK sample, Athanasakou et al. (2007) adopt the methodology of Kasznik and McNichols (2002) and report that meeting analyst expectations yields an additional cumulated abnormal return of 7.9% over the year. Extending Kasznik and McNichols (2002), they further show that investors discount the benchmark premium when targets are met by means of accrual-based earnings management, specific types of real earnings management, and/or classification shifting and do not reward benchmark achievement at all in the presence of earnings guidance. Gleason and Mills (2008) examine how the visibility of earnings management affects the rewards for meeting expectations. They suggest that more visible earnings management yields lower rewards than beating genuinely or with less transparent earnings management. As expected, the premium for meeting expectations is significantly discounted when analyst expectations are met by means of highly visible tax adjustments in the fourth quarter of the year. Investors thus seem to disentangle discretionary earnings decreases and discount their lack of persistence accordingly. Eventually, Das et al. (2011a) examine differences in the stock market reaction to meeting analyst expecta-
tions when the target is met genuinely and by means of earnings and expectations management. They find that the reward for beating expectations decreases significantly when managers engage in earnings or expectations management. This effect even intensifies when managers use both instruments complementary.

The aforementioned results confirm that meeting or beating the analyst forecast benchmark has significant valuation consequences. The rewards for benchmark achievement may, however, decrease or even vanish when managers engage in specific types of earnings management or analyst guidance.

3.2.3.1.2 Investor Rationality, Opportunism, and Signaling

The documented capital market reactions to benchmark beating raise questions about investors’ rationality and explanations for earnings and expectations management: 1) Do rewards for benchmark beating reflect positive future firm performance?; and 2) Are manipulations used to signal management’s positive expectations about future performance or to deceive investors about the firm’s true prospects? One way to address these questions is to compare the subsequent performance of benchmark beaters and missers and to examine differences in subsequent performance when benchmarks are met by means of earnings management or forecast guidance.

Kasznik and McNichols (2002), for instance, analyze whether the premium to meeting expectations is solely due to expected positive future earnings or reflects a distinct market premium. Their findings suggest that the benchmark premium is grounded in the predictive power of beating analyst forecasts. In other words, beating the benchmark is an indication of positive future performance and the market rationally incorporates this information into share prices. Results presented in Bartov et al. (2002) generally confirm Kasznik and McNichols (2002): Firms that achieve the analyst benchmark exhibit better future earnings performance than those that fail to do so. Though attenuated, this finding still holds when firms engage in earnings management to achieve the benchmark. They conclude that managers use earnings management to signal their expectations about firm performance and investors interpret these signals correctly. Similar results are reported by Gunny (2010) and Chen et al. (2010). Gunny (2010) finds a significantly higher industry-adjusted ROA for firms that previously engaged in real earnings management to meet analyst expectations and concludes that real earnings management is rather used to signal future performance than to deceive shareholders. According to Chen et al. (2010), firms that use accrual-based earnings management to beat analyst forecasts subsequently underperform firms
that achieve the benchmark genuinely or by means of real earnings management. Applying real earnings management, however, does not affect subsequent performance when compared to firms that genuinely fulfill earnings expectations. The results suggest that managers use real earnings management to signal positive future performance. Leggett et al. (2009) and Bhojraj et al. (2009) provide results that contradict the signaling hypothesis for real earnings management. Leggett et al. (2009) find that real earnings management to avoid reporting a loss adversely affects future operating performance and attribute this to the high economic costs associated with the manipulation of real business activities. Bhojraj et al. (2009) compare the short- and long-term consequences of real and accrual-based earnings management to meet earnings expectations for a narrow region of earnings surprises around the analyst forecast threshold. Their results contrast with the signaling hypothesis of earnings management and suggest myopic activities to improve short-term stock performance.

Overall, the results provide convincing evidence that the benchmark premium is attributable to higher subsequent performance of benchmark beaters. Considering the consequences of earnings and expectations management, however, yields ambiguous results. While some evidence supports the signaling hypothesis, others suggest myopic behavior to deliberately deceive shareholders about the firm’s true prospects.

3.2.3.2 Contracting Incentives

Empirical evidence of contracting incentives for benchmark beating mainly considers executive compensation. Richardson et al. (2004), for instance, show that strong equity incentives increase the probability of forecast guidance to avoid falling short of analysts’ earnings estimates. With a similar approach, Cheng and Warfield (2005) find that strong equity incentives (e.g., high volume of stock option grants or stock holdings) stimulate executives to engage in income increasing earnings management to beat analyst forecasts. Both studies suggest opportunistic actions to achieve important earnings targets and maximize proceeds from option exercises or stock sales. Other studies consider compensation contracts and the role of benchmark beating without directly referring to equity-based compensation. Matsunaga and Park (2001), for instance, analyze how missing quarterly earnings benchmarks affects the CEO’s annual bonus. After controlling for a “normal” penalty for poor performance, their results suggest that missing quarterly analyst forecasts or earnings of the same quarter last year at least two times during the year decreases bonus payments and provides incentives to manage earnings. Mergenthaler et al. (2011) extend Matsunaga and Park (2001) and examine whether missing quarterly analyst forecasts is asso-
associated with career penalties for CEOs and CFOs. In addition to decreases in bonus payments, career penalties include equity grant reductions and forced dismissals. Their results suggest that CFOs and CEOs are rewarded when the benchmark is achieved and penalized otherwise.

### 3.3 Surveys on Benchmark Beating

The previous section summarizes the large body of literature that empirically addresses the manipulation of financial reports or guidance of market intermediaries to beat important earnings benchmarks. A major drawback of all empirical studies is, however, that they test the joint hypothesis of earnings management existence and the adequacy of the underlying empirical model used to detect earnings management. In other words, failure to reject the null of no earnings management does not necessarily mean that earnings management is not existent. That is, earnings management could have taken place but the empirical model fails to detect it. Similarly, rejecting the null of no earnings management could be attributable to an oversensitive earnings management measure that detects earnings management although it did not not take place.\(^\text{13}\)

This section presents results on earnings management that avoid misleading inferences due to empirical measurement errors and summarizes several surveys on benchmark beating.

Graham et al. (2005) and Graham et al. (2006) present results from surveys and interviews among more than 400 US executives on the importance of earnings benchmarks and the relevance of earnings management activities in financial reporting practice. With a focus on financial statement users, De Jong et al. (2009) replicate the Graham et al. (2005) study and survey 306 US financial analysts. Nöldeke (2007a) transfers the Graham et al. (2005) methodology into the European setting and presents results from surveys among 85 CFOs/IROs and 82 financial analysts from Austria, Germany, and Switzerland. Three survey questions are closely related to this study and summarized in the remainder of this section:

- **Which earnings benchmarks are important?** Figure 3.2 summarizes which earnings benchmarks are considered as most important by executives (*Panel A*) and financial analy-

\(^{13}\)Several simulation studies examine the accuracy of earnings management measures. These include, e.g., Dechow et al. (1995), Jeter and Shivakumar (1999), Peasnell et al. (2000b), Alcarria Jaime and De Albornoz Noguer (2004), Kothari et al. (2005), and Stubben (2010) for discretionary accrual measures, Cohen et al. (2011) for real earnings management measures, and Takeuchi (2004) for distributional tests. In Appendix A.3, I provide a similar simulation analysis to identify the most adequate discretionary accrual proxy in the German setting and minimize the risk of erroneous inferences.
lich (Panel B) in the US (gray bars) and Europe (Austria, Germany, and Switzerland; dark gray bars). Executives in both Europe and the US perceive last year’s EPS as the most important benchmark. For European managers, the second most important benchmark is the forecast previously issued by management (no US data available), closely followed by last reported EPS and zero earnings. Their US counterparts consider the analyst consensus forecast and zero earnings as the second and third most important benchmark, respectively. Overall, reporting positive earnings and earnings surprises seems to be more important in the US than in Europe. For European financial analysts, the most prominent benchmark is the management forecast (no US data available), followed by last audited EPS and the analyst consensus estimate. US analysts rank the analyst consensus forecast as most important benchmark, followed by EPS in the same quarter last year and last quarter’s EPS. Interestingly, zero earnings is not considered as an important benchmark for US and European analysts.

- **Why meet earnings benchmarks?** Figure 3.3 summarizes the respondents’ views on the motives for benchmark beating. The majority of executives in the US and Europe state that meeting earnings benchmarks helps to maintain/increase capital market credibility and stock prices. These results are in line with the empirical findings on benchmark-driven market premiums (e.g., Lopez and Rees, 2002; Bartov et al., 2002; Kasznik and McNichols, 2002) and suggest that the earnings game is considerably driven by capital market forces. Alternative explanations such as management reputation, communication with stakeholders and/or contracting seem to be less important than capital market motivations. While a relatively large fraction of respondents name manager reputation as important motive for benchmark beating, employee bonuses and potential violation of debt covenants are not perceived drivers as of the earnings game. The opinions of financial analysts overlap with those of managers in many aspects: Most European and US financial analysts regard capital market credibility, management reputation, and signaling of growth prospects as potential explanations for benchmark beating. Debt covenants and employee bonuses, in contrast, are not viewed as a reason for benchmark importance.

- **What actions are taken to meet earnings benchmarks?** The respondents’ views on preferred actions to avoid missing earnings benchmarks are summarized in Figure 3.4. With regard to the large body of literature on benchmark related earnings management in

---

14Note that the US study refers to EPS of the same quarter last year and the European study to last audited EPS.
accounting research (see Section 3.2) and descriptions of the phenomenon in popular business press (see, e.g., Collingwood, 2001), the results are quite surprising: Except for the active guidance of financial analysts towards beatable earnings targets, European executives and analysts do not agree on the notion that managers deliberately exercise discretion to achieve earnings targets. The US results paint a different picture: A considerable fraction of US managers freely admit to lower discretionary spending and delay costly new investments albeit small sacrifices in firm-value. US analysts, however, do not seem to be aware of these maneuvers and suggest that managers use share repurchases to inflate EPS. Accrual-based earnings management is neither expected by analysts, nor admitted by responding US executives.

The surveys provide results which are in several aspects conflicting with previous empirical evidence. It is noteworthy that, though mitigating some shortcomings of empirical studies (e.g., misspecification and measurement errors), the survey approach introduces another potential source of bias: Managers’ resistance to admit discretionary behavior (Nöldeke, 2007b, p. 71). Hence, survey results should not be taken for granted unless they conform with evidence from, e.g., empirical studies or further qualitative research (e.g., field study analysis).
Panel A: Survey among Financial Executives

Panel B: Survey among Financial Analysts

Fig. 3.2.—Survey Results: Important Earnings Benchmarks. The figures summarize results from surveys conducted in Austria, Germany, and Switzerland (in dark gray) and the US (in gray). The importance score represents the respondents’ average agreement on the importance of the respective benchmarks (0 = strongly disagree, 100 = strongly agree). Panel A summarizes the views of 401 financial executives in the US (Graham et al., 2005) and 85 CFOs/IROs in Austria, Germany, and Switzerland (Nöldeke, 2007a,b). Panel B summarizes the views of 306 financial analysts covering US firms (De Jong et al., 2009) and 82 financial analysts from Austria, Germany, and Switzerland (Nöldeke, 2007a,b). Importance score is based on the authors average importance rating rescaled to 100.

1 no data available for the US sample
Panel A: Survey among Financial Executives

Panel B: Survey among Financial Analysts

FIG. 3.3.—Survey Results: Benchmark Beating Incentives. The figures summarize results from surveys conducted in Austria, Germany, and Switzerland (in dark gray) and the US (in gray). The importance score represents the respondents’ average agreement on the respective incentives for benchmark beating (0 = strongly disagree, 100 = strongly agree). Panel A summarizes the views of 401 financial executives in the US (Graham et al., 2005) and 85 CFOs/IROs in Austria, Germany, and Switzerland (Nöldeke, 2007a,b). Panel B summarizes the views of 306 financial analysts covering US firms (De Jong et al., 2009) and 82 financial analysts from Austria, Germany, and Switzerland (Nöldeke, 2007a,b). Importance score is based on the authors average importance rating rescaled to 100.
Panel A: Survey among Financial Executives

Panel B: Survey among Financial Analysts

FIG. 3.4.—Survey Results: Actions to Achieve Benchmarks. The figures summarize results from surveys conducted in Austria, Germany, and Switzerland (in dark gray) and the US (in gray). The importance score represents the respondents’ average agreement on the respective actions (0 = strongly disagree, 100 = strongly agree). Panel A summarizes the views of 401 financial executives in the US (Graham et al., 2005) and 85 CFOs/IROs in Austria, Germany, and Switzerland (Nöldeke, 2007a,b). Panel B summarizes the views of 306 financial analysts covering US firms (De Jong et al., 2009) and 82 financial analysts from Austria, Germany, and Switzerland (Nöldeke, 2007a,b). Importance score is based on the authors average importance rating rescaled to 100. EG, AEM, and REM stand for earnings guidance, accrual-based, and real earnings management, respectively.

1 no data available for the US sample
Chapter 4

Research Questions, General Methodology, and Contribution

Chapter 3 provides a comprehensive review of previous research on earnings management to achieve benchmarks. In a nutshell, these studies address the following research questions:

– **Is benchmark beating a prevalent phenomenon?** Related studies analyze the distribution of earnings, earnings changes, and earnings surprises to examine the prevalence of earnings management to achieve benchmarks.

– **What are managers’ techniques to meet or beat benchmarks?** Related studies analyze the type and intensity of manipulations used to achieve earnings targets. Possible techniques are accrual-based earnings management, real earnings management, and earnings guidance.

– **Why do benchmarks matter?** Related studies examine why managers engage in earnings or expectations management to meet benchmarks. The majority of these studies considers the capital market consequences of benchmark beating.

This chapter draws on previous evidence of benchmark related earnings management to develop an outline for the current study. Section 4.1 compares institutional characteristics of the US and Germany and shows why patterns of benchmark beating in Germany may differ from those in the US. Research questions and the general methodology of this study are presented in Section 4.2. Section 4.3 highlights the contribution of my research.
4.1 The German Economy as Institutional Setting

The vast majority of the studies presented in the previous chapter are conducted either in the US or the UK. These results may, however, not necessarily be valid for different settings. Accounting research provides compelling evidence that the quality of earnings and the level of earnings management depend on the respective institutional settings and regulatory frameworks (see among others, e.g., Ball et al., 2000, 2003; Leuz et al., 2003; Ball, 2006; Burgstahler et al., 2006). Table 4.1 subsumes some traditional institutional differences between Germany and the US.

<table>
<thead>
<tr>
<th>Table 4.1</th>
<th>Traditional Institutional Differences between Germany and the US</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>US</td>
</tr>
<tr>
<td><strong>Legal Origin</strong></td>
<td>Code Law</td>
</tr>
<tr>
<td><strong>Investor Protection</strong></td>
<td>weak</td>
</tr>
<tr>
<td><strong>Financing</strong></td>
<td>Internal &amp; Bank</td>
</tr>
<tr>
<td><strong>Ownership</strong></td>
<td>concentrated</td>
</tr>
<tr>
<td><strong>Information Asymmetry</strong></td>
<td>low</td>
</tr>
<tr>
<td><strong>Capital Market Size</strong></td>
<td>small</td>
</tr>
<tr>
<td><strong>Tax Alignment</strong></td>
<td>high</td>
</tr>
<tr>
<td><strong>Source of Accounting Rules</strong></td>
<td>Legislation</td>
</tr>
</tbody>
</table>

The table provides a summary of traditional institutional differences between Germany and the US. Similar comparisons can be found in, e.g., Glaum (2000) and Leuz and Wüstemann (2004).

Historically, Germany is described as a typical bank-based or insider economy where capital is provided by internal financing (e.g., pension liabilities, retained earnings) or banks (Leuz and Wüstemann, 2004; Burgstahler et al., 2006). Although steadily increasing, the German capital market is relatively small (Theissen, 2004) and shareholdings are dominated by large non-financial firms (cross holdings) and/or financial institutions (Glaum, 2000). These large stockholders are usually represented in the firm’s supervisory board (“Aufsichtsrat”), the main instrument for corporate governance in Germany (Leuz and Wüstemann, 2004). The stock markets in the US developed much earlier than in continental Europe and are the main source for external financing. In contrast to Germany, ownership is much more dispersed and the role of banks is less important (Glaum et al., 2004). These differences in ownership structure and financing shape diverging environments of corporate control in Germany and the US: Highly concentrated ownership, bank financing, and supervisory board representation in Germany heavily increase the
power of major share- and stakeholders and reduce the market’s demand for investor protection, legal enforcement, and public disclosure. Higher dispersion of ownership and the importance of public equity financing in the US, in contrast, increase the need for regulation that enables smaller shareholders to effectively exercise control rights. This includes, e.g., mandatory public disclosure on a regular basis, investor protection (e.g., voting power of minority shareholders, ease of participation in voting, legal protection against expropriation by management or blockholders), and strong legal enforcement of shareholder rights (La Porta et al., 1998). These structural differences between Germany and the US are also reflected in accounting regulation. As a result, traditional German accounting is rather geared towards efficient (debt) contracting than public information (Leuz and Wüstemann, 2004). In the US, information asymmetry between managers and shareholders is comparably strong and the primary role of accounting is to deliver the information needed to assess current and future firm performance (Leuz and Wüstemann, 2004).

Although the traditional view of two opposing financial systems is still widespread, some remarks on recent institutional changes are noteworthy in this context. In the last 15 years, Germany steered towards becoming a more developed equity market with increased transparency, stronger investor rights, and more rigid enforcement mechanisms. This progress is not only attributable to the capital requirements of German firms, but also to the severe tumults that shook equity markets in the early 2000s and provoked major regulatory changes. These include the strengthening of supervisory boards (e.g., extended responsibility, specific role in risk management and internal control issues), the empowering of shareholders (e.g., easier exertion of voting rights, one-share one-vote rules), legal enforcement (e.g., criminal sanctions for market manipulation and insider trading), financial reporting and disclosure requirements (e.g., mandatory IFRS reporting, extended interim reporting, disclosure of executive compensation, introduction of a corporate governance codex on comply-or-explain basis), and accounting enforcement (e.g., introduction of a two-tier enforcement mechanism).1 In the light of these developments, some important institutional characteristics of Germany seem to align with those of the US. However, the remaining differences are still substantial and it is doubtful whether the two systems will actually converge in the nearer future (Goergen et al., 2008).

Overall, differences in institutional settings are expected to affect the quality of earnings in general and earnings management in particular. Glaum et al. (2004), on the one hand, argue that incentives for benchmark beating in the US are stronger than in Germany because of capital

---

1For detailed reviews of the developments in German corporate governance refer to, e.g., Enriques and Volpin (2007), Goergen et al. (2008), or von Werder and Talaulicar (2011).
market dominance and a higher degree of equity-based contracting with management. On the other hand, they consider the US rule-based accounting system to restrict discretion in financial reporting. Leuz et al. (2003) argue that lower investor protection (as it can be found in Germany) is typically associated with a higher degree of earnings management because managers and/or close shareholders have stronger incentives to conceal private control benefits. Burgstahler et al. (2006) suggest that patterns of earnings management are shaped by capital market forces and the respective institutional setting (e.g., legal enforcement, tax alignment, accounting regulation, investor protection). In the same vein as Leuz et al. (2003), they assume low shareholder protection, a less rigid enforcement environment, high book-tax alignment, low disclosure requirements, and less developed equity markets to increase earnings management activity. In summary, ample differences of German and US financial systems seem to affect the intensity of earnings management. Hence, simply transferring US results into the German setting is not a very meaningful endeavor.

To the best of my knowledge, nine studies address earnings management to achieve benchmarks in Germany, either specifically or in course of an international comparison. Their main results are summarized in Table 4.2.

### Table 4.2

**German Evidence on Benchmark Beating**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample/Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bhattacharya et al. (2003)</td>
<td>3,847 firm-years, 1986–1998; Distributional Approach</td>
<td>Among 34 international countries, Germany ranges on rank 14 in terms of the highest ratio of small profits to small losses. In comparison, the US ranges on rank 23.</td>
</tr>
<tr>
<td>Leuz et al. (2003)</td>
<td>4,440 firm-years, 1990–1998; Distributional Approach</td>
<td>Among 31 international countries, Germany ranges on rank 14 in terms of the highest ratio of small profits to small losses. In comparison, the US ranges on rank 29.</td>
</tr>
<tr>
<td>Glaum et al. (2004)</td>
<td>3,524 firm-years, 1991–2000; Distributional Approach</td>
<td>German and US firms engage in earnings management to achieve the zero earnings, earnings changes, and analyst forecast benchmark. Concerning the first two thresholds, the extent of earnings management does not significantly differ for German and US firms. The analyst forecast benchmark, however, is more important for US firms.</td>
</tr>
<tr>
<td>Holzapfel (2004)</td>
<td>3,112–3,906 firm-years, 1987–2000; Distributional Approach, Event Study</td>
<td>German managers regard zero EPS, EPS changes, and analyst forecasts as important benchmarks. Furthermore managers engage in earnings and expectations management to meet last year’s EPS or the analyst consensus estimate. The market reacts positive to achieving the earnings changes and analyst forecast benchmark.</td>
</tr>
</tbody>
</table>

*Continued on next page.*
4. Research Questions, General Methodology, and Contribution

<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample/Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown and Higgins (2005)</td>
<td>643 firm-years, 1991–2000; Unexpected Forecasts/Forecast Revisions</td>
<td>Among 22 international countries, Germany ranges on rank 11 (based on forecast revisions) and rank 12 (based on unexpected forecasts) in terms of earnings guidance intensity. In comparison, US firms exhibit by far the highest degree of analyst guidance among all countries.</td>
</tr>
<tr>
<td>Coppens and Peek (2005)</td>
<td>3,978 firm-years, 1993–1999; Distributional Approach</td>
<td>The distributions of earnings and earnings changes of public firms exhibit significant discontinuities at zero. The distributions of private firms, in contrast, are comparably smooth. Hence, earnings management seems to be triggered by capital market incentives.</td>
</tr>
<tr>
<td>Burgstahler et al. (2006)</td>
<td>9,430 firm-years, 1997–2003; Distributional Approach</td>
<td>Among 13 European countries, Germany exhibits the third highest ratio of small profits to small losses for public firms and the second highest ratio for private firms. The results suggest that German managers intensively engage in earnings management to avoid small losses.</td>
</tr>
<tr>
<td>Daske et al. (2006)</td>
<td>3,060–9,483 firm-years, 1986–2001; Distributional Approach</td>
<td>German executives engage in earnings management to achieve the zero earnings, earnings changes, and analyst forecast benchmark. In comparison to 13 other European countries, Germany is among the countries with the strongest distributional discontinuities at all three benchmarks. Overall, distributional discontinuities in Europe are far more pronounced than in the US.</td>
</tr>
<tr>
<td>Nöldeke (2007b)</td>
<td>514 firm-years, 400 responses; Distributional Approach, Survey, Forecast Revisions</td>
<td>Distributional tests reveal strong discontinuities at the zero earnings and analyst forecast benchmark. The irregularity at the earnings changes benchmark is considerably lower. Survey results provide conflicting results (see Section 3.3). Managers guide financial analysts down to beatable targets.</td>
</tr>
</tbody>
</table>

The table summarizes the results of studies on benchmark related earnings management in Germany. Note that some of these studies measure benchmark related earnings management as one of several other earnings attributes. The table, however, solely summarizes the results directly related to benchmark beating.

In summary, previous research on benchmark beating suggests that German managers actively engage in earnings or expectations management to achieve benchmarks. Moreover, most of these studies report that earnings management is more prevalent in Germany than in the US.

4.2 Research Questions and General Methodology

The previous section offers two reasons why Germany provides a fruitful setting for my research. First, the German institutional setting differs in many respects from the previously examined anglo-american systems and these differences seem to affect the incentives and constraints of benchmark related earnings management. Second, some of the previous evidence suggests higher
earnings management activity in Germany than in the US.

My first research question considers the capital market reaction to beating important earnings benchmarks. Behavioral theory in finance suggests that investors (and other stakeholders) use prominent earnings benchmarks as focal points in their decision processes (see Section 3.1). Surveys in the financial community support behavioral theory: Asked about the incentives for benchmark beating, the majority of German CFOs and IROs cited “Capital Market Credibility” (83%) and “Maintain/Increase Stock Price” (74%) as main reasons for benchmark importance (see Section 3.3). Furthermore, previous studies conducted in the US (e.g., Lopez and Rees, 2002; Kasznik and McNichols, 2002; Bartov et al., 2002; Brown and Caylor, 2005) and the UK (e.g., Athanasakou et al., 2007) provide empirical evidence of a return premium for benchmark achievement (see Section 3.2.3.1). Since capital market incentives for benchmark beating remain largely untackled in the German setting, my first research question is stated as follows:

RQ (1): Do German investors reward benchmark beating with a premium or penalize missing a benchmark with a discount, so that managers face capital market incentives for earnings management to avoid reporting a loss, earnings declines, or missing the latest analyst consensus forecast?

In the spirit of Lopez and Rees (2002), I measure the market reaction to benchmark beating as market-adjusted cumulative return for a short window of three days surrounding the earnings release. If capital market incentives for earnings management exist in Germany, I expect market-adjusted returns to be positive when the respective benchmark is met or beaten and negative otherwise. In contrast to a prior study by Holzapfel (2004), I use a multivariate analysis that controls for the information content in current earnings and several factors known to affect the relation of returns and earnings.

The second research question addresses the prevalence of earnings management in Germany. Prior evidence suggests that a substantial fraction of firms manage earnings to avoid falling barely short of earnings benchmarks (e.g., Leuz et al., 2003; Glaum et al., 2004; Daske et al., 2006). Furthermore, the intensity of benchmark related earnings management seems to be higher in Germany than in the US or the UK (e.g., Leuz et al., 2003; Daske et al., 2006). Survey results, however, draw a different picture. According to Nöldeke (2007a,b), German managers admit to engage in forecast guidance to lower analyst expectations to a beatable level. In contrast to their US counterparts, however, they deny to engage in any type of earnings management (see Section 3.3). Two reasons potentially explain this discrepancy: 1) The surveyed managers do not tell the
truth and try to mask discretionary behavior; or 2) previous studies measure German earnings management activity with error. To reassess the prevalence of benchmark beating in the German context, my second research question is stated as follows:

**RQ (2):** *Do German firms engage in discretionary activities to avoid reporting a loss, earnings declines, or missing the latest analyst consensus forecast?*

I address the second research question with the distributional approach introduced by Burgstahler and Dichev (1997). In contrast to previous German studies, however, I modify and extend their methodology to account for some major shortcomings. These include bias due to deflation (Dechow et al., 2003; Durtschi and Easton, 2005, 2009), binwidth choice (Holland, 2004), non-normalities in the parent distribution (Christodoulou and McLeay, 2009), rounding errors in the I/B/E/S database (Baber and Kang, 2002; Payne and Thomas, 2003), and effective tax rate asymmetries (Beaver et al., 2007). I use a non-parametric approach to estimate a reference distribution of premanaged earnings and rigid robustness tests to minimize the risk of erroneous inferences.

Previous German studies address the prevalence of earnings management to achieve benchmarks without considering specific types of manipulations. US research provides mixed evidence of specific earnings management techniques (see Section 3.2.2) and several authors attribute this to noisy measures of discretionary activities (e.g., Dechow et al., 2003). However, if earnings management is more intensive in Germany than in the US, the German setting should provide fertile grounds to examine specific patterns of earnings management. The third part of my research has two distinct functions. First, it directly examines management’s techniques to beat benchmarks. Possible alternatives are expectations management as well as accrual-based and real earnings management. Second, it provides a rigid test of the distributional approach: Even after controlling for major caveats of the Burgstahler and Dichev (1997)-methodology, distributional irregularities may have other reasons than earnings management. Dechow et al. (2003), for example, suggest that efficient contracting with employees and management increases efficiency when a firm gets close to important earnings benchmarks. Beaver et al. (2007), as another example, show that distributional irregularities at the zero earnings benchmark are at least in part attributable to different effective tax rates of profit and loss firms. The third research question is split into two parts:
4. Research Questions, General Methodology, and Contribution

**RQ (3.1):** Are discontinuities in the distributions of earnings metrics attributable to discretionary activities?

**RQ (3.2):** What type of manipulations are applied to avoid reporting a loss, earnings declines, or missing the latest analyst consensus forecast?

To test for specific types of earnings management, I regress proxy variables for earnings management on two dummy variables identifying firms that barely miss or just achieve the respective benchmark and a set of control variables. I use nine earnings management proxy variables to test for several different types of accrual-based and real earnings management. These include aggregate abnormal accruals based on the Lagged Model (Dechow et al., 2003), premature revenue recognition (Stubben, 2010), cuts in discretionary expenses (Roychowdhury, 2006; Gunny, 2010), overproduction (Roychowdhury, 2006; Gunny, 2010), and gains from fixed asset sales (Gunny, 2010). For the analyst forecast benchmark, I supplement these proxies with a measure of earnings guidance based on Matsumoto (2002). My test procedure follows the basic idea in Dechow et al. (2003): If managers engage in discretionary actions to avoid missing a benchmark, the average earnings management proxy of “just beat” observations should be significantly larger than the average proxy of 1) “just miss” observations and 2) all remaining observations (i.e., observations that do not fall in the region directly adjacent to the threshold). I use a multivariate regression approach to test for differences between the three groups. To avoid erroneous inferences, the analysis is accompanied by a set of additional analyses and robustness checks.

4.3 Contribution

The outcome of my research is of potential interest for investors, standard setters, government, and regulatory bodies. Investors may gain insights on the reliability of earnings figures when incentives for earnings management are strong. Standard setters might reassess areas of high accounting discretion to restrict leeway for earnings manipulations. For governmental and regulatory bodies, knowledge about financial statement manipulation may help to develop more effective enforcement mechanisms.

My study differs from previous research and provides unprecedented evidence in the following ways:

- It is the first study to provide comprehensive evidence of benchmark related earnings man-
agement under mandatory IFRS accounting. Previous studies focus either on German GAAP reporting or mixed samples with German GAAP and IFRS observations. However, some properties of German GAAP accounting (e.g., high book-tax alignment) may influence the outcomes of these studies.

– It is the first study to address capital market incentives for earnings management in Germany with a multivariate regression approach. In a previous German study, Holzapfel (2004) reports that beating earnings benchmarks is related with positive abnormal returns. However, he does not control for the information content in current earnings and other factors that may affect abnormal returns.

– In comparison to previous studies on earnings management around thresholds, my research design acknowledges the fact that the distribution of (unmanaged) earnings may be not normal (see, e.g., Christodoulou and McLeay, 2009) and uses a non-parametric approach based on kernel density estimation (Bollen and Pool, 2009) to test for irregularities in the distribution of earnings, earnings changes, and earnings surprises. Furthermore, this study directly tackles several other shortcomings of the distributional approach (e.g., binwidth choice, scaling, tax effects, split-adjusted I/B/E/S data) and thus provides more robust results on discontinuities in the distributions of earnings metrics.

– It is the first study to test for specific manipulation techniques in the German setting. A comprehensive set of nine earnings management proxies and one earnings guidance measure allows to shed light on the tools potentially used by managers to achieve earnings targets. Possible manipulation techniques include accrual-based earnings management (e.g., aggregate discretionary accruals, discretionary revenues), real earnings management (i.e., expense cutting, fixed asset sales), and expectations management. Testing for specific types of earnings management further provides a rigid test of the distributional approach and allows to assess the validity this methodology in the German context.
Chapter 5

The Market Reaction to Benchmark Beating

If meeting or beating earnings benchmarks matters, I expect the market to react positively when a benchmark is achieved and negatively when it is missed. In efficient markets, achieving earnings benchmarks affects share price if it provides information about future earnings that is incremental to current year’s earnings information (Kasznik and McNichols, 2002). If achieving earnings thresholds signals firm health and good prospects, firms that do so earn a premium. If the theory of efficient markets does not hold, then behavioral aspects come into play. These include, e.g., investor overreaction around benchmarks due to their use as focal points and hurdles for “yes-no” decisions (Degeorge et al., 2005). No matter whether market rationality or behavioral aspects dominate, if meeting or beating earnings benchmarks affects investors’ decisions, then benchmark beaters will enjoy higher abnormal returns upon the earnings announcement than benchmark missers. Accordingly, I test the following three hypotheses (stated in their alternative form):

**H_A (1a):** Controlling for the information content in current earnings, firms that meet or beat the analyst forecast benchmark earn higher abnormal returns than firms that fail to do so.

**H_A (1b):** Controlling for the information content in current earnings, firms that meet or beat the zero earnings benchmark earn higher abnormal returns than firms that fail to do so.
5. The Market Reaction to Benchmark Beating

$H_A (1c)$: *Controlling for the information content in current earnings, firms that meet or beat the earnings changes benchmark earn higher abnormal returns than firms that fail to do so.*

The remainder of this chapter is structured as follows: Section 5.1 explains the empirical methodology used to test for the capital market effects of benchmark achievement. In Section 5.2, I define the variables required for my analyses and provide details about their derivation. Section 5.3 summarizes the sample selection procedure and descriptive statistics. The empirical results are presented in Section 5.4. Eventually, Section 5.5 closes with a brief summary.

### 5.1 Empirical Methodology

My methodology is designed to test whether firms that meet or beat an earnings benchmark yield abnormal returns in a short window surrounding the earnings announcement date. To gain preliminary insights, I compare the cumulated abnormal returns of benchmark beaters and benchmark missers in the univariate analyses. If achieving benchmarks matters, abnormal returns of the former should be significantly higher in magnitude. In the multivariate analysis, I draw on previous research (see, e.g., Lopez and Rees, 2002; Bartov et al., 2002; Kasznik and McNichols, 2002) and regress cumulated abnormal returns around the earnings release on a dummy variable identifying benchmark achievers, unexpected earnings for the respective period, and an interaction of unexpected earnings and the dummy variable. The basic model (Model 1) is given as

\[
CAR_{it}[-1,1] = \alpha_1 + \alpha_2 MBE_{it}^{Q} + \beta_1 DSURP_{it} + \beta_2 MBE_{it}^{Q} \times DSURP_{it} + \epsilon_{it},
\]

(5.1.1)

where

- $CAR_{it}[-1,1]$: abnormal return of firm $i$ cumulated from one day before until one day after the earnings announcement in year $t$,
- $MBE_{it}^{Q}$: is 1 if firm $i$ has met or beaten benchmark $Q$ (i.e., consensus forecast, zero earnings, or change in earnings) in $t$ and 0 otherwise, and
- $DSURP_{it}$: unexpected earnings for the respective period.
5. The Market Reaction to Benchmark Beating

\[ DSURP_{it} \] is firm \( i \)'s earnings surprise in year \( t \) calculated as the difference between actual EPS and the last EPS consensus forecast before the earnings announcement scaled by end of the year share price.

The interpretation of the regression results is straightforward. If meeting or beating benchmark \( Q \) yields an incremental return irrespective of the magnitude of the earnings surprise, the coefficient on the dummy \( MBE_{it}^Q \) is expected to be significantly greater than zero. Likewise, if achieving a benchmark changes the abnormal return per unit of (deflated) earnings surprise, the interacted earnings response coefficient (ERC) on \( MBE_{it}^Q \times DSURP_{it} \) should be significantly positive.\(^1\)

The basic model is misspecified if variables that affect the relation between the dependent and the independent variables are omitted (omitted variables problem) or variables are measured with error (error in variables problem). Estimated coefficients then either capture effects that are attributable to other (omitted) factors or are attenuated due to measurement errors. Coefficients are then biased and drawn inferences not valid.\(^2\) One possibility to deal with potentially biased coefficient estimates is to include control variables that partial out other factors that affect the relation of interest. With Model 2 and Model 3, I estimate two extended versions of the basic model in Eq. (5.1.1). These models include several control variables suspected to affect the relation of earnings and returns or deemed to correct measurement deficiencies. Control variables are included individually (Model 2) and in interaction with the earnings surprise (Model 3). Hence, Model 2 is given as

\[
CAR_{it}[-1, 1] = \alpha_1 + \alpha_2 MBE_{it}^Q + \beta_1 DSURP_{it} + \beta_2 MBE_{it}^Q \times DSURP_{it} + \sum_{j=1}^{J} \gamma_j Z_{jit} + \epsilon_{it},
\] (5.1.2)

\(^1\)Earnings response coefficients are the coefficients on (unexpected) earnings in a regression of (abnormal) stock returns on (unexpected) earnings. In other words, the ERC measures the market’s reaction to an earnings announcement per unit of earnings surprise. The interacted ERC measures the incremental market reaction to benchmark beating per unit of earnings surprise. For a review of the literature on earnings response coefficients see, e.g., Kothari (2001, pp. 123–143) with further references.

\(^2\)More technically spoken, if relevant variables are omitted from the model or independent variables are measured with error, then disturbances are correlated with some or all of the independent variables and the zero conditional mean assumption of the classic linear regression model is violated. For details and consequences of endogenous variables see, e.g., Greene (2003, pp. 14–15 (assumptions), pp. 83ff. (measurement error), and pp. 148ff. (variable omission)).
and Model 3 as

\[ \text{CAR}_it[-1, 1] = \alpha_1 + \alpha_2 MBE_{it}^{O} + \beta_1 \text{DSURP}_{it} + \beta_2 MBE_{it}^{O} \times \text{DSURP}_{it} \]

\[ + \sum_{j=1}^{J} \gamma_j Z_{jit} + \sum_{k=1}^{K} \delta_k (Z_{kit} \times \text{DSURP}_{it}) + \epsilon_{it}, \]

(5.1.3)

where \(Z_{jit}\) and \(Z_{kit}\) represent control variables for firm \(i\) in year \(t\). These include:\(^3\)

- **Preevent Returns (PRET).** Following Easton and Zmijewski (1989) and Keung et al. (2010), I use preevent returns to control for information leaking to the market after the consensus forecast was reported. Since the information environment may change in the short window between the last consensus forecast and the earnings release, earnings surprise DSURP is a noisy proxy for market expectations. Including preevent returns addresses potential endogeneity bias due to measurement error in market expectations.\(^4\) I expect the coefficient estimate for \(PRET\) to be negative.

- **Growth Expectations (GROW).** Collins and Kothari (1989) show that firms’ growth expectations affect ERCs. If firms signal useful information about the opportunities of a current and future investment set, growth is expected to be positively correlated with the ERC. Including \(GROW\) as a control variable is deemed to control for this effect. I expect the coefficient on \(GROW\) to be positive.

- **Systematic Risk (SRISK).** Collins and Kothari (1989) and Easton and Zmijewski (1989) provide evidence that the relation of abnormal returns and unexpected earnings is affected by firms’ systematic risk. Price response to unexpected earnings is triggered by investors’ revisions of expected future dividends. Since investors discount expected future dividends with the stock’s risk-adjusted rate of return, higher risk is associated with lower stock price reactions to earnings surprises. Ignoring the effect of systematic risk potentially biases ERCs. I add \(SRISK\) as proxy for different levels of securities’ systematic risk and expect it to have a negative sign.

- **Earnings Persistence (PERST).** Kormendi and Lipe (1987), Collins and Kothari (1989),

\(^3\)Firm \((i)\) and time \((t)\) subscripts are suppressed for notational convenience.

\(^4\)Refer to Brown et al. (1987) for a detailed evaluation of using preevent returns to correct for errors in variables.
and Lipe (1990) find that the persistence of unexpected earnings affects ERCs. If earnings provide information about future dividends, then higher earnings persistence should be positively correlated with unexpected returns because investors’ revisions of future dividend expectations and related stock price reactions will be stronger. To ensure that abnormal returns are not attributable to different levels of earnings persistence, I use \( \text{PERST} \) as an additional control variable and expect its coefficient to be positive.

- **Firm Size (SIZE).** Empirical evidence suggests significant differences in the information environments of large and small firms (see, e.g., Atiase, 1980, 1985). Specifically, close following of investors and analysts as well as more intensive information processing for large firms increases the amount of publicly available information in comparison to smaller firms (Freeman, 1987). This faster dissemination of information in the period between the issuance of the last consensus forecast and the earnings announcement induces measurement error in unexpected earnings and biases the ERC (Easton and Zmijewski, 1989). To avoid potential bias, I include \( \text{SIZE} \) as an additional control variable and expect it to have a negative sign.

- **Default Risk (DRISK).** Dhaliwal et al. (1991) find that ERCs are related with a firm’s default risk (i.e., the risk of bankruptcy). Since default risk determines the allocation of firm-value to stockholders and bondholders, an increase in a firm’s default risk should be negatively related with the ERC. I follow Dhaliwal et al. (1991) and use the firm’s financial leverage as a proxy for default risk. I expect the coefficient on \( \text{DRISK} \) to be negative.

- **Analyst Forecast Properties (AGE, FOLL).** I expect the accuracy of analyst forecasts to decrease with forecast age (i.e., the period between the last consensus forecast and the earnings announcement date) and increase with the number of analysts following. That is, the ERC is negatively related with forecast age (\( \text{AGE} \)) and positively related with the number of following analysts (\( \text{FOLL} \)). I add both variables as controls and predict \( \text{AGE} \) to be negative and \( \text{FOLL} \) to be positive.

One issue with large multivariate regression models is multicollinearity, which results in unstable coefficient estimates and/or unreliable \( t \)-statistics due to inflated standard errors (see, e.g., Greene, 2003, pp. 56–59). I detect high correlations between \( \text{DSURP} \) and the interactions
5. The Market Reaction to Benchmark Beating

SIZE $\times$ DSURP (92%) and AGE $\times$ DSURP (85%) in my data.\(^5\) Furthermore, the variance inflation factor analysis (VIF) reveals values clearly above 10 for DSURP and SIZE $\times$ DSURP.\(^6\) To avoid multicollinearity, I remove the control variable interactions SIZE $\times$ DSURP and AGE $\times$ DSURP. A subsequent test of multicollinearity detects no critical VIF values in any of my models.

5.2 Variable Definitions

5.2.1 Dependent Variables

To test whether meeting or beating an earnings benchmark yields an incremental abnormal return requires to identify the expected return $E(R_{id})$ of a security $i$ on day $d$. Given a security’s expected return, the abnormal return $AR_{id}$ is calculated as the difference of the actual return $R_{id}$ and the expected return $(E(R_{id}))$:\(^7\)

$$AR_{id} = R_{id} - E(R_{id}).$$  \hspace{1cm} (5.2.1)

The literature of empirical corporate finance provides several models for the calculation of expected returns. Among these, the Capital Asset Pricing Model (CAPM), the market model as well as the more convenient mean-, market-, size-, or risk-adjusted models are common choices.\(^8\) In this study, I follow Keung et al. (2010) and use the residuals of market model regressions as estimates for unexpected returns. The market model uses a regression approach to estimate the relation of a security’s return ($R_{id}$) and the return of the market ($R_{md}$):

$$R_{id} = \alpha + \beta R_{md} + \varepsilon_{id}.$$  \hspace{1cm} (5.2.2)

\(^5\)In a similar study, Lopez and Rees (2002) also detect multicollinearity between deflated earnings surprises and deflated earnings surprises interacted with firm size.

\(^6\)For critical values of multicollinearity using VIF analysis see, e.g., StataCorp LP (2009, pp. 1573–1576).

\(^7\)For a more detailed introduction to event study methodology in finance and the derivation and cumulation of abnormal returns see, e.g., MacKinlay (1997).

\(^8\)For details on different models developed to estimate (un)expected security returns see, e.g., Brown and Warner (1980), Brown and Warner (1985), Strong (1992), and MacKinlay (1997).
In a first step, Eq. (5.2.2) is estimated for every firm in the sample during a specified period (estimation window) prior to the event of interest. The market return coefficient $\beta$ (market beta) of these time-series regressions describes the relation of the stock return and the return of the market. In a second step, the coefficient estimates $\hat{\alpha}$ and $\hat{\beta}$ are used to calculate expected returns during the event window:

$$E(R_{id}) = \hat{\alpha} + \hat{\beta}R_{md}.$$  \hspace{1cm} (5.2.3)

Eventually, abnormal returns are calculated using Eq. (5.2.1) and cumulated over the event period $D$:

$$CAR_d[0,D] = \sum_{d=0}^{D} (R_{id} - E(R_{id})).$$  \hspace{1cm} (5.2.4)

I follow Keung et al. (2010) and estimate market model coefficients over an estimation window of 255 trading days ending 41 days before the respective firm’s earnings announcement. A clear cut of 41 days between the event and the estimation of the market betas ensures that the estimation is not contaminated by earnings information leaking to the market before earnings are released. The event window spans from one day before until one day after the earnings release. Figure 5.1 illustrates the construction of the estimation and event window on a time line.

![Figure 5.1](image-url)

**FIG. 5.1.—Calculation of Abnormal Returns on a Time Line.** The time line illustrates the estimation and event window used to calculate cumulated abnormal returns in the three days surrounding the earnings announcement.

The application of the market model approach requires the definition of an appropriate market return. Since my sample builds on firms constituting the German composite performance index CDAX, this index is a reasonable choice as reference portfolio for the wide range of firms in my
sample. Daily firm and index returns are calculated as

\[ R_{id} = \frac{P_{id}}{P_{i,d-1}} - 1 \quad \text{and} \quad R_{CDAX,d} = \frac{P_{CDAX,d}}{P_{CDAX,d-1}} - 1, \quad (5.2.5) \]

respectively, where \( P_{id} \) is stock \( i \)'s closing price and \( P_{CDAX,d} \) is the stock index’s closing price on day \( d \). Time-series data for firms (DS: P) and the market index (DS: PI) are gathered from Datastream. All firm prices are adjusted for stock splits and capital changes. Given the firm-specific market model parameter estimates \( \hat{\alpha} \) and \( \hat{\beta} \), cumulated abnormal returns \( CAR_{it}[-1,1] \) are calculated as

\[ CAR_{it}[-1,1] = \sum_{d=-1}^{1} (R_{id} - \hat{\alpha} + \hat{\beta} R_{CDAX,d}). \quad (5.2.6) \]

A critical aspect in the derivation of abnormal returns is to find the correct earnings announcement dates. I collect earnings announcement dates from four different financial databases: I/B/E/S, Worldscope, Bloomberg, and Reuters 3000 Xtra. I further gather announcement dates from the Börsenzeitung’s “Bilanzfahrplan”, a regularly published summary of earnings releases for large German companies.\(^9\) Unfortunately, many database entries turn out to be erroneous when manually checked with daily news. The same issue is reported by Acker and Duck (2009), who find that I/B/E/S earnings announcement dates are affected by data entry errors due to mislabeling in the data collection process. In their large UK dataset, 24% of all earnings announcements do not coincide with hand-collected earnings releases. To increase data accuracy, I merge earnings announcement dates gathered from five different sources into one file and extract dates that match in at least two of these sources. In the relative rare case of two matching pairs, I choose the earlier date.\(^10\) Figure 5.2 illustrates this procedure. In the first row, for example, I/B/E/S, Worldscope, and Börsenzeitung report 2/24/2006 as earnings release date. Hence, it is highly probable that this date is the true announcement date. In the second row are two matching pairs: I/B/E/S and Worldscope report 2/24/2009 as earnings release, while Reuters 3000 Xtra and the Börsenzeitung agree on 2/19/2009. In this case, the earlier date is chosen: 2/19/2009. Although I lose a number of observations with this procedure, manual checks prove that my

---

\(^9\)The Börsenzeitung is a daily and purely financial newspaper reporting about developments on the German stock market.

\(^10\)With two matching pairs, the first date is usually some preliminary earnings release while the second date is the release of the financial report. Since large discrepancies between preliminary and final earnings are rather unusual, I expect the earlier announcement to convey the relevant information.
procedure is highly effective in finding accurate earnings announcements.

<table>
<thead>
<tr>
<th>Database Entries</th>
<th>I/B/E/S</th>
<th>Worldscope</th>
<th>Bloomberg</th>
<th>Reuters 3000</th>
<th>Börsenzeitung</th>
<th>Chosen Date</th>
</tr>
</thead>
</table>

**Fig. 5.2.—Identification of Correct Earnings Announcement Dates.** The example illustrates the procedure to extract the most accurate announcement date from five different sources. Columns I to V show exemplary database entries. Bold entries indicate which announcement date is most accurate (i.e., has the highest probability to coincide with the true announcement date). The release date chosen for the analysis is shown in column VI.

### 5.2.2 Independent Variables

This section summarizes the definition and measurement of the required independent variables:

- **Earnings Surprise** ($DSURP$). I measure deflated earnings surprise $DSURP$ as difference between actual EPS and the last mean consensus forecast before the earnings announcement date scaled by end of the year share price. Fiscal year-end stock price (unadjusted) is retrieved from Datastream (DS: UP). EPS and EPS analyst consensus estimates are gathered from Thomson Reuters’ I/B/E/S Summary History tape.\(^{11}\) Important to note, all summary forecasts are “unadjusted” for stock splits.\(^{12}\) Using unadjusted estimates is substantial when working with earnings surprises calculated from I/B/E/S data. This point was first made by Baber and Kang (2002) and Payne and Thomas (2003) and is described in more detail in Appendix A.2.

- **Benchmark Dummies** ($MBE^\Omega$). The indicator variables identifying firms that meet or beat the analyst consensus forecast ($MBE^{FC}$), the zero earnings benchmark ($MBE^{ZE}$), or the earnings changes benchmark ($MBE^{EC}$) are defined as:

\(^{11}\)Since 1987, I/B/E/S collects sell-side analyst forecasts for non-US firms. Today, I/B/E/S provides the broadest and deepest knowledge base for forecast related analysis (Thomson Reuters, 2010).

\(^{12}\)I thank Dr. Ulrich Weigel for his support in gathering unadjusted EPS estimates and actuals from Thomson Reuters.
\[
MBE^{FC} = \begin{cases} 
1 & \text{if } DSURP \geq 0, \\
0 & \text{if } DSURP < 0, 
\end{cases} \\
MBE^{ZE} = \begin{cases} 
1 & \text{if } EPS \geq 0, \\
0 & \text{if } EPS < 0, 
\end{cases} \\
MBE^{EC} = \begin{cases} 
1 & \text{if } \Delta EPS \geq 0, \\
0 & \text{if } \Delta EPS < 0. 
\end{cases}
\]

\(EPS\) is earnings per share, \(\Delta EPS\) is current year’s earnings per share minus last year’s earnings per share, and \(DSURP\) is earnings surprise as defined above. All data is gathered from I/B/E/S.

- **Growth Expectations (GROW).** To proxy for growth opportunities, I use the ratio of market capitalization to book value of equity (market-to-book ratio or M/B-ratio) at the end of the fiscal year (see, e.g., Collins and Kothari, 1989; Bartov et al., 1999; Lopez and Rees, 2002). I expect the M/B-ratio to be positively related with economic growth (i.e., growth firms exhibit a higher M/B-ratio than small firms). Data required to calculate \(GROW\) is collected from Thomson Reuters’ Worldscope database. Market capitalization is common shares outstanding at the end of the fiscal year (WS: 05031) times closing share price of the year (WS: 05085). Total equity is book value of common equity (WS: 03501) including minority interest (WS: 03426).

- **Systematic Risk (SRISK).** As proxy for systematic risk, I use firm-specific market betas (see, e.g., Collins and Kothari, 1989; Easton and Zmijewski, 1989; Lopez and Rees, 2002; Herrmann et al., 2011). Market betas are derived from market model regressions (see Eq. (5.2.2)) for a 60 day interval ending one day before the earnings announcement. I use CDAX index returns to proxy for market returns. Daily closing prices (DS: P) and index values (DS: PI) required for the calculation of \(SRISK\) are retrieved from Datastream.

- **Earnings Persistence (PERST).** I follow Ali and Zarowin (1992), Cheng et al. (1996), and Lopez and Rees (2002) to proxy for earnings persistence. As shown in Ou and Penman (1989), considerably high/low ratios of earnings to price (earnings-to-price ratio or \(E/P\)-ratio) indicate more transitory earnings. Drawing on this relation, I use the following five steps to calculate \(PERST\) as a proxy for earnings persistence:

1. In each year of the sample period, rank all firms by their \(E/P\)-ratio,
2. Assign rank one to all negative E/P-ratios,
3. Build nine equally sized portfolios ranked by the E/P-ratio of the remaining firms,
4. Assign ranks two to ten to the nine portfolios in ascending order,
5. Set dummy $PERS_{T} = 1$ if rank is between three and eight (i.e., high persistence) and $PERS_{T} = 0$ otherwise (i.e., low persistence).

To calculate the E/P-ratio, I use actual EPS as reported in I/B/E/S and end of the year closing stock price (WS: 05085) from Worldscope.

- **Firm Size ($SIZE$)**. Market capitalization at the fiscal year-end is used to measure firm size (see, e.g., Atiase, 1985; Collins et al., 1987; Collins and Kothari, 1989). Following Easton and Zmijewski (1989), Shevlin and Shores (1993), and Herrmann et al. (2011), I transform market capitalization into its natural logarithm. Market capitalization is derived from the Worldscope database by multiplying common shares outstanding at the end of the fiscal year (WS: 05031) with the closing share price of the year (WS: 05085).

- **Default Risk ($DRISK$)**. I proxy for default risk using firms’ financial leverage (Dhaliwal et al., 1991). Similar to Lopez and Rees (2002), financial leverage is calculated as the ratio of long-term debt (WS: 03251) including the current portion of long-term debt (WS: 18232) to the sum of long-term debt, the current portion of long-term debt, and the book value of total equity (WS: 03501 + 03426). All data is gathered from Worldscope.

- **Analyst Forecast Properties ($AGE, FOLL$)**. $AGE$ is defined as the number of trading days between the issuance of the last mean consensus forecast and the earnings release date. $FOLL$ is the number of analysts included in the calculation of the last mean consensus forecast. Both variables are collected from the I/B/E/S summary history tape.
5.3 Sample Selection and Descriptive Statistics

5.3.1 Sample Selection

The initial sample contains all CDAX constituents between 2005 and 2009 that prepare consolidated financial statements according to IFRS or US GAAP. Regulation 1606/2002 of the European Commission prescribes the compulsory adoption IFRS in the EU for consolidated financial statements of publicly traded companies after the 1st of January 2005. Companies that are traded in non-member states and report according to US GAAP are obliged to switch to IFRS after the 1st of January 2007. Thus, limiting the sample to consolidated reports after 2005 ensures that my findings are not biased by local GAAP specifics (e.g., high book-tax alignment) and strong differences in accounting regulations. Based on this initial sample, the final sample is developed as follows:

- **Common and Preferred Stock.** When firms are listed with both common and preferred stock, I remove preferred stock observations to avoid double counting.14

- **Missing Data.** From the remaining sample, I drop all firm-years with missing data on the intersection of the Worldscope, Datastream, and I/B/E/S databases. Furthermore, I lose observations with missing prior year data required to calculate the earnings changes benchmark and/or missing earnings release dates.

- **Validity of Earnings Releases.** To ensure the validity of earnings announcement dates, I remove all observations with less than 30 or more than 180 days between the fiscal year-end and the earnings release. Early announcements are often preliminary announcements without earnings disclosure or erroneous database entries. Late announcements are frequently incorrect or indicate severe problems within firms (e.g., a recent filing for insolvency).

- **Stale Forecasts.** I exclude all EPS consensus estimates older than 40 trading days since I do not expect them to reflect market expectations at the earnings announcement.

- **Financial Firms.** Following Burgstahler and Dichev (1997) and subsequent studies, I ex-

---

13Firms are included if they are listed on the CDAX index as of June 30th and/or December 31st of the respective year. Historical CDAX constituents are gathered from Bloomberg.

14Note that common and preferred stock may differ in stock valuation but share the same income figures.
clude financial service firms (two-digit SIC codes between 60 and 67) to avoid that results are blurred by regulatory oversight.

- **Influential Observations.** Eventually, I control for influential observations by dropping observations with \( \text{CAR} \) and/or \( \text{DSURP} \) falling into the 1\(^{\text{st}}\) or 99\(^{\text{th}}\) percentile of their yearly distributions.

The sample selection procedure is summarized in Table 5.1. The number of CDAX observations is relatively stable with about 680 observations until 2008 and then drops to about 650 in 2009. This is, most likely, an effect attributable to the subprime crisis that spread out starting in 2007 and led to severe downturns in stock markets all around the world. The number of observations in the final sample follows a different pattern and is constantly increasing from 128 observations in 2005 to 213 observations in 2009. Clearly, most of the initial sample is lost due to data availability: Overall 1,627 or 48.0\% of the initial sample observations are lost because Worldscope and I/B/E/S lack data entries. In the case of forecast data, however, this is not surprising since not all firms in the broad CDAX index are followed by analysts. Comparably low, 251 observations or 7.4\% of all CDAX firms and 84 observations or 2.5\% are lost due to missing prior year and return data, respectively. Besides data availability, a relatively large number of observations has to be dropped because they are operating in the financial service sector (217 firm-years or 6.4\% of the initial sample), listed twice (140 firm-years or 4.1\% of the initial sample), and/or the related forecasts are stale (74 firm-years or 2.2\% of the initial sample). Overall, the final sample covers 927 observations, which makes 27.4\% of all CDAX observations. Although, the final sample is quite small compared to the initial sample, I cover a large percentage of the overall German stock market capitalization. The last three rows of Table 5.1 illustrate these relations. Except for 2005 with 38.7\%, the sample’s coverage increases from 63.6\% in 2006 to 71.9\% in 2009. For the whole period, the final sample covers 61.6\% of German non-financial publicly traded equity.


**Table 5.1**

Sample Selection Procedure

<table>
<thead>
<tr>
<th></th>
<th>Number of Observations</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2005</td>
<td>2006</td>
<td>2007</td>
<td>2008</td>
<td>2009</td>
<td>Pooled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total CDAX Observations</td>
<td>681</td>
<td>684</td>
<td>690</td>
<td>683</td>
<td>652</td>
<td>3,390</td>
<td></td>
<td>100.00</td>
</tr>
<tr>
<td>less: Preferred Stock</td>
<td>30</td>
<td>29</td>
<td>29</td>
<td>27</td>
<td>25</td>
<td>140</td>
<td>4.13</td>
<td></td>
</tr>
<tr>
<td>less: Missing Forecasts and Actuals</td>
<td>356</td>
<td>338</td>
<td>308</td>
<td>317</td>
<td>308</td>
<td>1,627</td>
<td>47.99</td>
<td></td>
</tr>
<tr>
<td>less: Missing Return Data</td>
<td>20</td>
<td>24</td>
<td>24</td>
<td>10</td>
<td>6</td>
<td>84</td>
<td>2.48</td>
<td></td>
</tr>
<tr>
<td>less: Missing Prior Year Data</td>
<td>79</td>
<td>46</td>
<td>49</td>
<td>50</td>
<td>27</td>
<td>251</td>
<td>7.40</td>
<td></td>
</tr>
<tr>
<td>less: Erroneous and/or Late Earnings Releases</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>23</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>less: Stale Forecasts</td>
<td>16</td>
<td>13</td>
<td>18</td>
<td>10</td>
<td>17</td>
<td>74</td>
<td>2.18</td>
<td></td>
</tr>
<tr>
<td>less: Financial Services Firms</td>
<td>39</td>
<td>42</td>
<td>49</td>
<td>45</td>
<td>42</td>
<td>217</td>
<td>6.40</td>
<td></td>
</tr>
<tr>
<td>less: Outliers</td>
<td>8</td>
<td>8</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>47</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td>Final Sample</td>
<td>128</td>
<td>180</td>
<td>198</td>
<td>208</td>
<td>213</td>
<td>927</td>
<td>27.35</td>
<td></td>
</tr>
<tr>
<td>Market Capitalization of Non-financial Firms</td>
<td>828</td>
<td>1,001</td>
<td>1,183</td>
<td>700</td>
<td>777</td>
<td>4,489</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market Capitalization of Sample Firms</td>
<td>320</td>
<td>637</td>
<td>792</td>
<td>458</td>
<td>559</td>
<td>2,766</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample Coverage in %</td>
<td>38.6</td>
<td>63.6</td>
<td>66.9</td>
<td>65.4</td>
<td>71.9</td>
<td>61.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Preferred stock observations are deleted when a firm is listed more than once on the index. Observations are classified as missing if the relevant data fields are not available on I/B/E/S, Worldscope, or Datastream. Financial services firms are defined as having two-digit SIC codes between 60 and 67. Earnings announcements are considered erroneous or late when the date of the earnings release is less than 30 or more than 180 after the fiscal year-end. Forecasts are considered as stale when the latest consensus forecast was calculated more than 40 days before the respective earnings announcement. Outliers are defined as observations below the 1st or above the 99th percentile of the yearly distributions of CAR and DSURP. Year-end market capitalization is retrieved from Datastream (DS: MV). Market capitalization data for all German non-financial firms is gathered from summary statistics of the Deutsche Bundesbank (Deutsche Bundesbank, 2010, p. 45). Market capitalizations are reported in billion EUR.
The sample’s industry composition is summarized in Table 5.2. With 150 observations or 16.2% of the overall sample, the majority of observations is drawn from the “Business Services” sector, representing firms primarily engaged in rendering bussiness-to-business services. With 137 observations or 14.8% of the sample, the second largest sector comprises machinery and equipment including computers. Since the services and the industrial sector constitute the backbone of the German economy, their large stake is not surprising. Industries and services are then followed by electronics (101 firm-years or 10.9% of the sample), chemicals (70 firm-years or 7.6% of the sample), and manufacturing instruments (53 firm-years or 5.7% of the sample).

<table>
<thead>
<tr>
<th>Industry Description</th>
<th>Primary SIC</th>
<th>Observations</th>
<th>in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Services</td>
<td>73</td>
<td>150</td>
<td>16.2</td>
</tr>
<tr>
<td>Industrial and Commercial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machinery/Equipment and computers</td>
<td>35</td>
<td>137</td>
<td>14.8</td>
</tr>
<tr>
<td>Electronic and Other Electrical Equipment</td>
<td>36</td>
<td>101</td>
<td>10.9</td>
</tr>
<tr>
<td>Chemicals and Allied Products</td>
<td>28</td>
<td>70</td>
<td>7.6</td>
</tr>
<tr>
<td>Manufacturing Instruments</td>
<td>38</td>
<td>53</td>
<td>5.7</td>
</tr>
<tr>
<td>Transportation Equipment</td>
<td>37</td>
<td>38</td>
<td>4.1</td>
</tr>
<tr>
<td>Communications</td>
<td>48</td>
<td>33</td>
<td>3.6</td>
</tr>
<tr>
<td>Engineering, Accounting, Research, Management, and Related Services</td>
<td>87</td>
<td>30</td>
<td>3.2</td>
</tr>
<tr>
<td>Apparels and Similar Finished Products</td>
<td>23</td>
<td>23</td>
<td>2.5</td>
</tr>
<tr>
<td>Durable Goods Wholesale</td>
<td>50</td>
<td>20</td>
<td>2.2</td>
</tr>
<tr>
<td>Miscellaneous (below 20%)</td>
<td>272</td>
<td>29.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>927</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

The table summarizes the industry composition of the final sample. Two-digit SIC codes are gathered from the Worldscope database (WS: 07021).

5.3.2 Descriptive Statistics

Variable distributions are summarized in Table 5.3. Cumulated abnormal returns (CAR) are distributed quite symmetrically (skew: 0.02) and are moderately peaked (kurtosis: 4). Mean ab-

---

15In 2010, 47.4% of the German gross domestic product was attributable to the industry sector and service activities (Federal Statistical Office, 2011).
16An overview and detailed descriptions of the Standard Industry Classifications (SIC) can be found on the website of the US Department of Labor (http://www.osha.gov/pls/imis/sic_manual.html).
normal returns are slightly positive (0.001), while median abnormal returns equal zero. The distribution of earnings per share (EPS) is considerably right-skewed with a strong peak (skew: 8.83, kurtosis: 109). Mean EPS (1.59) are significantly larger than zero at the 1% level. Median EPS are also clearly positive (0.68), but lower than average EPS. The distribution of earnings changes (ΔEPS) exhibits a considerable peak (kurtosis: 38) and a slight left-skew (skew: −0.25). Though barely significant (10% level), mean ΔEPS are clearly below zero (−0.15) while the median marginally exceeds zero (0.05). The distribution of scaled earnings surprises (DSURP) is strongly peaked (kurtosis: 72) and skewed to the left (skew: −7.05). Mean DSURP are significantly negative (1% level), supporting prior evidence of forecast optimism in Germany (see, e.g., Capstaff et al., 1998, 2001). Preevent returns (PRET) are on average positive, but close to zero (0.01). A mean market beta (SRISK) of 0.68 indicates that the sample firms are of less than average riskiness. The average firm is valued about twice its book value of equity (GROW), has a logarithmized market capitalization of 5.66 (SIZE), and is followed by roughly 9 analysts. The last consensus estimate before the earnings release date is on average 18 days old.

### Table 5.3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>σ</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAR</td>
<td>0.001</td>
<td>0.050</td>
<td>−0.026</td>
<td>0.000</td>
<td>0.028</td>
</tr>
<tr>
<td>EPS</td>
<td>1.589</td>
<td>4.806</td>
<td>0.120</td>
<td>0.680</td>
<td>1.980</td>
</tr>
<tr>
<td>ΔEPS</td>
<td>−0.146</td>
<td>2.468</td>
<td>−0.400</td>
<td>0.050</td>
<td>0.410</td>
</tr>
<tr>
<td>DSURP</td>
<td>−0.018</td>
<td>0.098</td>
<td>−0.010</td>
<td>0.000</td>
<td>0.006</td>
</tr>
<tr>
<td>PRET</td>
<td>0.005</td>
<td>0.100</td>
<td>−0.046</td>
<td>0.001</td>
<td>0.048</td>
</tr>
<tr>
<td>SRISK</td>
<td>0.678</td>
<td>0.527</td>
<td>0.286</td>
<td>0.662</td>
<td>1.027</td>
</tr>
<tr>
<td>GROW</td>
<td>2.145</td>
<td>1.934</td>
<td>1.120</td>
<td>1.660</td>
<td>2.570</td>
</tr>
<tr>
<td>SIZE</td>
<td>5.659</td>
<td>2.039</td>
<td>3.988</td>
<td>5.262</td>
<td>7.025</td>
</tr>
<tr>
<td>DRISK</td>
<td>0.243</td>
<td>0.210</td>
<td>0.047</td>
<td>0.212</td>
<td>0.394</td>
</tr>
<tr>
<td>AGE</td>
<td>17.7</td>
<td>9.2</td>
<td>9.0</td>
<td>16.0</td>
<td>26.0</td>
</tr>
<tr>
<td>FOLL</td>
<td>8.8</td>
<td>9.0</td>
<td>2.0</td>
<td>5.0</td>
<td>13.0</td>
</tr>
</tbody>
</table>

The table summarizes the distributions of variables for 927 firm-years in the period 2005–2009. CAR is abnormal return cumulated over the three-day event period surrounding the earnings announcement date. Earnings per share (EPS) are actual “street” earnings as reported by I/B/E/S, earnings changes (ΔEPS) are actual EPS minus prior year’s EPS. Deflated earnings surprise DSURP is calculated as difference of actual EPS and the most recent I/B/E/S consensus estimate scaled by end of the year stock price. PRET is preevent abnormal return, SRISK denotes the firm’s market beta, GROW captures firm growth measured as the firm’s M/B-ratio, firm SIZE is the natural logarithm of market capitalization, DRISK is leverage calculated as the ratio of debt to debt plus equity, AGE is forecast age, and FOLL denotes the number of analysts following the firm. Observations with cumulated abnormal returns (CAR) and/or deflated forecast errors (DSURP) below the 1st or above the 99th percentile of their yearly distributions are regarded as outliers and removed. σ stands for the sample standard deviation.

Table 5.4 summarizes benchmark statistics for the earnings forecast threshold in Panel A, the
zero earnings threshold in Panel B, and the change in earnings threshold in Panel C. Observations are classified as MISS if reported EPS are below the respective threshold, MEET if reported EPS are exactly on target, and BEAT if reported EPS exceed the benchmark. The MEET/BEAT group comprises firms that meet or beat the benchmark. The number of firms that miss (442 or 47.7% of all observations) and beat (439 or 47.4%) the analyst forecast benchmark are nearly equal and only 46 (5.0%) observations are exactly on target. On the year level, there is a notable change in the proportion of missers and beaters starting from 2008 on. While more firms beat than miss the forecast until 2007, there are more missers than beaters afterwards. A potential explanation for that pattern is analysts’ failure to anticipate the severe earnings shortfalls during the global economic downturn at that time. In contrast, the proportions around zero earnings are less balanced: For the pooled sample, only 168 (18.1%) observations report a loss while 757 (81.7%) report a profit. This disproportion of profit and loss observations has been documented in several previous studies (see, e.g., Hayn, 1995; Glaum et al., 2004; Daske et al., 2006). Only two firms report exactly zero EPS, which may be an indication for managers’ reluctance to report zero earnings. Interestingly, there is a sharp drop in BEAT observations after 2008; a result most likely attributable to the outspread of the global financial crisis. At the earnings changes threshold, there are more firms that beat the benchmark than those that miss it. With 509 (54.9%) BEAT and 410 (44.2%) MISS observations, however, the asymmetry is significantly less pronounced than around the zero earnings target. Quite similar is the low frequency of MEET observations, suggesting that managers avoid to report exactly the same earnings as in the previous year.

Table 5.5 presents variable correlations, with Pearson correlations in the bottom left and Spearman rank correlations in the top right. Correlations of the dependent variable \( CAR \) and the independent variables are mainly as predicted: \( CAR \) is significantly positively related with \( MBE^{FC} \), the benchmark dummy for analyst forecasts. As expected, firms that meet or beat the analyst forecast seem to earn higher abnormal returns than those that miss it. Likewise, the earnings changes benchmark \( MBE^{EC} \) is positively correlated with \( CAR \). However, compared with analyst forecasts, the relation seems to be less pronounced for earnings changes. Though positive, the zero earnings benchmark dummy \( MBE^{ZE} \) is not significantly related with abnormal returns; a fact indicating that the zero earnings threshold is less important for market participants. The correlation between \( CAR \) and earnings surprise \( DSURP \) is significant and positive, suggesting that consensus forecasts are appropriate proxies for earnings expectations. The correlation between \( CAR \) and preevent returns (\( PRET \)) is significantly negative; a result indicating that \( PRET \) indeed

\[ ^{17} \text{Unless otherwise noted, I refers to Pearson correlations in the remainder of this section.} \]
### TABLE 5.4
*Benchmark Statistics*

**Panel A: Analyst Forecast Benchmark**

<table>
<thead>
<tr>
<th>Year</th>
<th>MISS N</th>
<th>in %</th>
<th>MEET N</th>
<th>in %</th>
<th>BEAT N</th>
<th>in %</th>
<th>MEET/BEAT N</th>
<th>in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>58</td>
<td>45.3</td>
<td>8</td>
<td>6.3</td>
<td>62</td>
<td>48.4</td>
<td>70</td>
<td>54.7</td>
</tr>
<tr>
<td>2006</td>
<td>69</td>
<td>38.3</td>
<td>14</td>
<td>7.8</td>
<td>97</td>
<td>53.9</td>
<td>111</td>
<td>61.7</td>
</tr>
<tr>
<td>2007</td>
<td>86</td>
<td>43.4</td>
<td>5</td>
<td>2.5</td>
<td>107</td>
<td>54.0</td>
<td>112</td>
<td>56.6</td>
</tr>
<tr>
<td>2008</td>
<td>118</td>
<td>56.7</td>
<td>10</td>
<td>4.8</td>
<td>80</td>
<td>38.5</td>
<td>90</td>
<td>43.3</td>
</tr>
<tr>
<td>2009</td>
<td>111</td>
<td>52.1</td>
<td>9</td>
<td>4.2</td>
<td>93</td>
<td>43.7</td>
<td>102</td>
<td>47.9</td>
</tr>
<tr>
<td>Pooled</td>
<td>442</td>
<td>47.7</td>
<td>46</td>
<td>5.0</td>
<td>439</td>
<td>47.4</td>
<td>485</td>
<td>52.3</td>
</tr>
</tbody>
</table>

**Panel B: Zero Earnings Benchmark**

<table>
<thead>
<tr>
<th>Year</th>
<th>MISS N</th>
<th>in %</th>
<th>MEET N</th>
<th>in %</th>
<th>BEAT N</th>
<th>in %</th>
<th>MEET/BEAT N</th>
<th>in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>23</td>
<td>18.0</td>
<td>2</td>
<td>1.6</td>
<td>103</td>
<td>80.5</td>
<td>105</td>
<td>82.0</td>
</tr>
<tr>
<td>2006</td>
<td>24</td>
<td>13.3</td>
<td>0</td>
<td>0.0</td>
<td>156</td>
<td>86.7</td>
<td>156</td>
<td>86.7</td>
</tr>
<tr>
<td>2007</td>
<td>24</td>
<td>12.1</td>
<td>0</td>
<td>0.0</td>
<td>174</td>
<td>87.9</td>
<td>174</td>
<td>87.9</td>
</tr>
<tr>
<td>2008</td>
<td>35</td>
<td>16.8</td>
<td>0</td>
<td>0.0</td>
<td>173</td>
<td>83.2</td>
<td>173</td>
<td>83.2</td>
</tr>
<tr>
<td>2009</td>
<td>62</td>
<td>29.1</td>
<td>0</td>
<td>0.0</td>
<td>151</td>
<td>70.9</td>
<td>151</td>
<td>70.9</td>
</tr>
<tr>
<td>Pooled</td>
<td>168</td>
<td>18.1</td>
<td>2</td>
<td>0.2</td>
<td>757</td>
<td>81.7</td>
<td>759</td>
<td>81.9</td>
</tr>
</tbody>
</table>

**Panel C: Earnings Changes Benchmark**

<table>
<thead>
<tr>
<th>Year</th>
<th>MISS N</th>
<th>in %</th>
<th>MEET N</th>
<th>in %</th>
<th>BEAT N</th>
<th>in %</th>
<th>MEET/BEAT N</th>
<th>in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>36</td>
<td>28.1</td>
<td>1</td>
<td>0.8</td>
<td>91</td>
<td>71.1</td>
<td>92</td>
<td>71.9</td>
</tr>
<tr>
<td>2006</td>
<td>46</td>
<td>25.6</td>
<td>2</td>
<td>1.1</td>
<td>132</td>
<td>73.3</td>
<td>134</td>
<td>74.4</td>
</tr>
<tr>
<td>2007</td>
<td>73</td>
<td>36.9</td>
<td>1</td>
<td>0.5</td>
<td>124</td>
<td>62.6</td>
<td>125</td>
<td>63.1</td>
</tr>
<tr>
<td>2008</td>
<td>104</td>
<td>50.0</td>
<td>0</td>
<td>0.0</td>
<td>104</td>
<td>50.0</td>
<td>104</td>
<td>50.0</td>
</tr>
<tr>
<td>2009</td>
<td>151</td>
<td>70.9</td>
<td>4</td>
<td>1.9</td>
<td>58</td>
<td>27.2</td>
<td>62</td>
<td>29.1</td>
</tr>
<tr>
<td>Pooled</td>
<td>410</td>
<td>44.2</td>
<td>8</td>
<td>0.9</td>
<td>509</td>
<td>54.9</td>
<td>517</td>
<td>55.8</td>
</tr>
</tbody>
</table>

The table summarizes 927 sample observations with respect to benchmark achievement for the period 2005–2009. Earnings per share (EPS) are actual “street” earnings as reported by I/B/E/S, earnings changes (ΔEPS) are actual EPS minus prior year’s EPS. Earnings surprise SURP is calculated as difference of actual EPS and the most recent I/B/E/S consensus estimate. Observations in Panel A are classified as MISS, MEET, BEAT, and MEET/BEAT if \(SURP < 0\), \(SURP = 0\), \(SURP > 0\), and \(SURP \geq 0\), respectively. Observations in Panel B are classified as MISS, MEET, BEAT, and MEET/BEAT if \(EPS < 0\), \(EPS = 0\), \(EPS > 0\), and \(EPS \geq 0\), respectively. Observations in Panel C are classified as MISS, MEET, BEAT, and MEET/BEAT if \(ΔEPS < 0\), \(ΔEPS = 0\), \(ΔEPS > 0\), and \(ΔEPS \geq 0\), respectively.
controls for measurement error in \( CAR \). Firm size (\( SIZE \)) is negatively correlated with \( CAR \) and thus supports the assumption that larger firms provide more pre-announcement information than small firms. Since high default risk lowers the residual claims of shareholders, \( DRISK \) is negatively related with abnormal returns. Except for \( FOLL \), exhibiting a significant correlation with \( CAR \) in the opposite direction as expected, none of the remaining control variables is significant at generally accepted levels.

Although, I do not discuss the correlations of independent variables in detail, some findings merit discussion: Interestingly, all benchmark dummy variables exhibit strong positive correlations. Hence, firms that achieve one threshold are more likely to beat other benchmarks, too. The correlation of earnings surprises (\( DSURP \)) with firm growth (\( GROW \)) and size (\( SIZE \)) is significantly positive. This finding is in line with prior empirical studies in the US (see, e.g., Gu and Wu, 2003; Herrmann et al., 2011). The number of analysts following the firm (\( FOLL \)) is positively related with \( DSURP \). Thus, analyst following seems to trigger forecast pessimism (or vice versa).\(^{18}\) The correlation of systematic risk (\( SRISK \)) and firm size (\( SIZE \)) is highly positive. This finding contrasts to prior evidence from the US (see, e.g., Banz, 1981, and following studies documenting an opposite “size effect”). However, a similar relation has been found on other European markets with thin trading of small stocks.\(^{19}\) Also positive, but less strong is the correlation of \( GROW \) and \( SRISK \). Since growth stocks usually face higher risks than value stocks, this relation seems reasonable and is in line with the US evidence presented in Krishnan (2003).

\(^{18}\)Gu and Wu (2003), in contrast, find a negative relation between forecast errors and analyst following.

\(^{19}\)Martikainen and Perttunen (1991), for instance, document a positive correlation of market beta and firm size and interpret this as a result of downward biased risk estimates for small and infrequently traded firms.
### Table 5.5

*Variable Correlations*

<table>
<thead>
<tr>
<th></th>
<th>CAR</th>
<th>MBE&lt;sup&gt;FC&lt;/sup&gt;</th>
<th>MBE&lt;sup&gt;ZE&lt;/sup&gt;</th>
<th>MBE&lt;sup&gt;EC&lt;/sup&gt;</th>
<th>DSURP</th>
<th>PRET</th>
<th>SRISK</th>
<th>GROW</th>
<th>SIZE</th>
<th>DRISK</th>
<th>PERST</th>
<th>AGE</th>
<th>FOLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAR</td>
<td>0.133&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.051</td>
<td>0.110</td>
<td>0.161</td>
<td>-0.065&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.002</td>
<td>-0.045</td>
<td>-0.056&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-0.080&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.003</td>
<td>0.026</td>
<td>-0.055&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>MBE&lt;sup&gt;FC&lt;/sup&gt;</td>
<td>0.127&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.285&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.311&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.865&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.078&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.004</td>
<td>0.062&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.053</td>
<td>-0.057&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.110&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.040</td>
<td>0.033</td>
<td></td>
</tr>
<tr>
<td>MBE&lt;sup&gt;ZE&lt;/sup&gt;</td>
<td>0.042</td>
<td>0.285&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.275&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.398&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.026</td>
<td>-0.003</td>
<td>0.162&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.248&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.015</td>
<td>0.516&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.003</td>
<td>0.161&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>MBE&lt;sup&gt;EC&lt;/sup&gt;</td>
<td>0.108&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.311&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.275&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.376&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.003</td>
<td>0.064&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.239&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.084&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.080&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.152&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.030</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td>DSURP</td>
<td>0.105&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.339&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.414&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.247&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.064&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.005</td>
<td>0.081&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.067&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.034</td>
<td>0.130&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.006</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>PRET</td>
<td>-0.074&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.098&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.010</td>
<td>-0.003</td>
<td>-0.074&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.005</td>
<td>-0.072&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.037</td>
<td>-0.017</td>
<td>-0.013</td>
<td>0.004</td>
<td>-0.035</td>
<td></td>
</tr>
<tr>
<td>SRISK</td>
<td>-0.016</td>
<td>0.003</td>
<td>-0.014</td>
<td>0.061&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.053</td>
<td>-0.008</td>
<td>0.182&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.450&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.126&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.026</td>
<td>0.038</td>
<td>0.470&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>GROW</td>
<td>-0.047</td>
<td>0.031</td>
<td>0.018</td>
<td>0.176&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.072&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.068&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.118&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.409&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.020</td>
<td>0.239&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.012</td>
<td>0.276&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>-0.065&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.054</td>
<td>0.239&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.083&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.154&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.069&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.401&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.200&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.319&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.207&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.001</td>
<td>0.837&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>DRISK</td>
<td>-0.084&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.066&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.001</td>
<td>-0.083&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.050</td>
<td>-0.031</td>
<td>0.094&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.014</td>
<td>0.308&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000</td>
<td>-0.019</td>
<td>0.300&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>PERST</td>
<td>-0.003</td>
<td>0.110&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.516&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.152&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.188&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.003</td>
<td>-0.047</td>
<td>0.118&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.197&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.008</td>
<td>-0.015</td>
<td>0.160&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>0.026</td>
<td>-0.039</td>
<td>-0.002</td>
<td>0.023</td>
<td>0.052</td>
<td>0.013</td>
<td>0.043</td>
<td>-0.026</td>
<td>0.004</td>
<td>-0.010</td>
<td>-0.017</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>FOLL</td>
<td>-0.070&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.013</td>
<td>0.152&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.008</td>
<td>0.087&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-0.059&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.375&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.100&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.850&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.287&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.148&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.012</td>
<td></td>
</tr>
</tbody>
</table>

The table shows Pearson correlations in the bottom left and Spearman rank correlations in the top right corner. The total sample comprises 927 observations between 2005 and 2009. CAR is abnormal return cumulated over the three-day event period surrounding the earnings announcement date. MBE<sup>FC</sup>, MBE<sup>ZE</sup>, and MBE<sup>EC</sup> are dummy variables indicating whether the firm has achieved (MBE = 1) or missed (MBE = 0) the respective benchmark. DSURP stands for earnings surprises scaled by end of the year stock price, PRET is pre-event abnormal return, SRISK denotes the firm’s market beta, GROW captures firm growth proxied by the M/B-ratio, firm SIZE is the natural logarithm of market capitalization, DRISK is firm leverage measured as the ratio of debt to debt plus equity, PERST is a dummy variable taking on 1 for high and 0 for low earnings persistence, AGE is forecast age, and FOLL denotes the number of analysts following. Significance at the 1%, 5%, and 10% level is indicated by a, b, and c, respectively.
5. Market Reaction to Benchmark Beating

5.4 Empirical Results

In this section, I provide empirical evidence on the capital market response to benchmark achievement. The remainder of this section is divided in three subsections (Section 5.4.1–5.4.3) with univariate and multivariate results for each benchmark under investigation. Section 5.4.4 provides robustness checks to ensure the validity of the results.

5.4.1 The Analyst Forecast Benchmark

If achieving the latest analyst consensus forecast matters, I expect the market response to be significantly positive for benchmark achievers (MBE) and negative for benchmark missers (MISS). Table 5.6 summarizes mean three-day cumulated abnormal returns (CAR) around the earnings release for all MBE and MISS observations (Column II to VII) and related significance levels. With a clear prediction for the sign of abnormal returns above and below the benchmark, I calculate p-values based on one-tailed mean comparison t-tests (i.e., H₀ for MISS: \( \text{CAR} \geq 0 \), H₀ for MBE: \( \text{CAR} \leq 0 \)). The last two columns of the table report differences in \( \text{CAR} \) for MBE and MISS firms with related p-values. Since achieving the forecast is assumed to yield a market premium, \( \text{p} \)-values are calculated using a one-tailed test of differences in means (i.e., H₀: \( \text{CAR}_{\text{MBE}} - \text{CAR}_{\text{MISS}} \leq 0 \)).

### Table 5.6

**Market Rewards for Forecast Achievement—Univariate Results**

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Mean</th>
<th>p-value</th>
<th>N</th>
<th>Mean</th>
<th>p-value</th>
<th>N</th>
<th>Mean</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>58</td>
<td>-0.0162</td>
<td>0.0032***</td>
<td>70</td>
<td>0.0131</td>
<td>0.0058***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>69</td>
<td>0.0011</td>
<td>0.5888</td>
<td>111</td>
<td>0.0057</td>
<td>0.0545*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>86</td>
<td>-0.0092</td>
<td>0.0429**</td>
<td>112</td>
<td>0.0069</td>
<td>0.0558*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>118</td>
<td>-0.0045</td>
<td>0.2267</td>
<td>90</td>
<td>0.0093</td>
<td>0.0796*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>111</td>
<td>-0.0017</td>
<td>0.3476</td>
<td>102</td>
<td>0.0031</td>
<td>0.2458</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled</td>
<td>442</td>
<td>-0.0054</td>
<td>0.0143**</td>
<td>485</td>
<td>0.0072</td>
<td>0.0004***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Columns II to VII depict cumulated abnormal returns for a three-day window spanning from one day before until one day after the earnings announcement date conditional on whether the latest earnings consensus forecast is missed (MISS) or achieved (MBE) with related p-values. Columns VIII and IX show differences in mean returns and related p-values. All p-values are based on one-tailed mean comparison t-tests. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***), respectively.

The last line in Table 5.6 summarizes average \( \text{CAR} \) for the MISS and MBE group and dif-
ferences in means for the pooled sample. As expected, MISS-firms exhibit significant negative abnormal returns (−0.54%), while MBE-firms enjoy positive returns (0.72%). Both values are significantly below/above zero on a level of at least 5%. Comparing abnormal returns of both groups suggests an average market premium for benchmark beating of 1.26%; a difference significant at the 1% level. The yearly analysis generally confirms the inferences drawn from the pooled sample: MBE-firms yield higher cumulated abnormal returns than MISS-firms in every sample year and these differences are significant at at least 10% in three of five sample years. Overall, the univariate analysis suggests that forecast achievement is a desirable earnings target for German firms.

Multivariate regression analysis is a more powerful method to analyze the market rewards for meeting or beating analyst expectations. In comparison to the univariate analysis, regression allows to hold other factors fixed and partial out the effect that is solely attributable to benchmark achievement. In my model, other factors include the level of earnings surprise (DSURP) and a set of control variables. If meeting or beating the analyst forecast benchmark matters, I expect the coefficient on \( MBE^{FC} \) (\( \alpha_2 \)) to be significantly positive. When \( \alpha_2 \) is significantly positive, then achieving the analyst forecast benchmark yields a premium which is unrelated to the magnitude of earnings surprise and other factors that may affect the returns/earnings relationship. Another potential benchmark effect is a different market response per unit of earnings surprise for firms above and below the benchmark. The interacted ERC \( \beta_2 \) measures this incremental premium for benchmark achievers.

The results of pooled OLS regressions in Table 5.7 demonstrate that meeting/beating the analyst consensus yields an incremental return irrespective of the magnitude of earnings surprise. In the basic specification (Model 1), the estimated coefficient \( \alpha_2 \) on the dummy variable \( MBE^{FC} \) is positive (0.0081) and significant at the 1% level. In other words: Beating the analyst forecast benchmark yields an additional cumulated abnormal return of 0.81%, irrespective of whether the benchmark is just met or earnings are far above target. Missing the benchmark, on the contrary, yields a significant negative abnormal return of −0.39%. Hence, meeting or beating the latest analyst consensus estimate is rewarded with a premium, while missing it is penalized. The incremental ERC for benchmark achievement (\( \beta_2 \)) is significantly positive (0.194) and indicates that benchmark beaters earn—on average—an incremental premium of 19.4% for each additional unit of deflated earnings surprise.\(^{20}\) Surprisingly, the coefficient on \( DSURP \) (\( \beta_1 \)) is not

\(^{20}\)The coefficient on price-deflated earnings surprises for benchmark beaters in the study by Lopez and Rees (2002) is significantly higher at 2.1. In comparison, my results suggest a coefficient of 0.22 (\( \beta_1 + \beta_2 \)).
significantly greater than zero. It thus seems as if investors are less sensitive to the magnitude of negative earnings surprises. To illustrate the results, consider two identical firms A and B. At year-end, A and B report a deflated earnings surprise of 0.08 and −0.08, respectively. Based on the coefficient estimates in Table 5.7, the cumulative abnormal return is 2.19% for firm A (−0.39 + 0.81 + 2.74 × 0.08 + 19.40 × 0.08) and −0.61% for firm B (−0.39 + 2.74 × −0.08).

The results remain relatively unchanged when control variables (Model 2) and interacted control variables are included (Model 3). Of the control variables in Model 2, only preevent returns (PRET) and default risk (DRISK) are significant at the 5% level and have the predicted signs. All remaining control variables are not significant at generally accepted levels. Including the control variables slightly increases the main variables of interest: The coefficient on $MBE_{FC}$ increases by 0.0013 to 0.0094, indicating that meeting or beating the benchmark is associated with an incremental premium of 0.94%. The incremental earnings response per unit of $DSURP$ ($\beta_2$), however, drops slightly by 0.0283 to 0.1657. Eventually, the intercept $\alpha_1$ and the ERC $\beta_1$ are not significantly different from zero. In Model 3, one of the non-interacted (PRET) and two of the interacted control variables (DRISK, PERST) are significant at a level of at least 5%. Compared to Model 1, the coefficient on $MBE_{FC}$ increases to 0.0104, suggesting a slightly higher incremental benchmark premium of 1.04%. The interacted ERC ($\beta_2$) decreases only marginally to 0.1889 and remains significantly positive at the 1% level. The constant term $\alpha_1$ and the ERC $\beta_1$ are not significantly different from zero.

Overall, uni- and multivariate tests suggest that achieving the latest analyst consensus forecast is associated with a significant increase in short-term abnormal returns. Important to note, this finding is unrelated to the magnitude of unexpected earnings. Furthermore, benchmark beaters earn significantly higher abnormal returns per unit of earnings surprise. In other words, positive earnings surprises are valued significantly higher than negative earnings surprises. These results provide compelling evidence in support of capital market incentives for beating the latest analyst consensus forecast.
**Table 5.7**

*Market Rewards for Forecast Achievement—Regression Results*

<table>
<thead>
<tr>
<th>Exp. Sign</th>
<th>Model 1 Coefficient</th>
<th>Model 1 p-value</th>
<th>Model 2 Coefficient</th>
<th>Model 2 p-value</th>
<th>Model 3 Coefficient</th>
<th>Model 3 p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_1 )</td>
<td>CONST</td>
<td>-0.0039</td>
<td>0.064*</td>
<td>0.0036</td>
<td>0.645</td>
<td>0.0002</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>MBEFC</td>
<td>+0.0081</td>
<td>0.010***</td>
<td>0.0094</td>
<td>0.004***</td>
<td>0.0104</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>DSURP</td>
<td>+0.0274</td>
<td>0.149</td>
<td>0.0256</td>
<td>0.164</td>
<td>-0.0488</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>MBEFC×DSURP</td>
<td>+0.1940</td>
<td>0.003***</td>
<td>0.1657</td>
<td>0.014**</td>
<td>0.1889</td>
</tr>
<tr>
<td>( \gamma_1 )</td>
<td>PRET</td>
<td>-0.0443</td>
<td>0.013**</td>
<td>0.0041</td>
<td>0.018**</td>
<td></td>
</tr>
<tr>
<td>( \gamma_2 )</td>
<td>SRISK</td>
<td>-0.0015</td>
<td>0.685</td>
<td>0.0002</td>
<td>0.527</td>
<td></td>
</tr>
<tr>
<td>( \gamma_3 )</td>
<td>GROW</td>
<td>+0.0011</td>
<td>0.931</td>
<td>0.0008</td>
<td>0.852</td>
<td></td>
</tr>
<tr>
<td>( \gamma_4 )</td>
<td>SIZE</td>
<td>-0.0006</td>
<td>0.346</td>
<td>-0.0004</td>
<td>0.398</td>
<td></td>
</tr>
<tr>
<td>( \gamma_5 )</td>
<td>DRISK</td>
<td>-0.0152</td>
<td>0.039**</td>
<td>-0.0094</td>
<td>0.133</td>
<td></td>
</tr>
<tr>
<td>( \gamma_6 )</td>
<td>PERST</td>
<td>+0.0003</td>
<td>0.462</td>
<td>0.0017</td>
<td>0.312</td>
<td></td>
</tr>
<tr>
<td>( \gamma_7 )</td>
<td>AGE</td>
<td>-0.0001</td>
<td>0.763</td>
<td>0.0001</td>
<td>0.721</td>
<td></td>
</tr>
<tr>
<td>( \gamma_8 )</td>
<td>FOLL</td>
<td>+0.0002</td>
<td>0.723</td>
<td>-0.0003</td>
<td>0.811</td>
<td></td>
</tr>
<tr>
<td>( \delta_1 )</td>
<td>PRET×SURP</td>
<td>+0.0757</td>
<td>0.667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta_2 )</td>
<td>SRISK×SURP</td>
<td>-0.0354</td>
<td>0.314</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta_3 )</td>
<td>GROW×SURP</td>
<td>0.0082</td>
<td>0.515</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta_4 )</td>
<td>DRISK×SURP</td>
<td>0.2667</td>
<td>0.001***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta_5 )</td>
<td>PERST×SURP</td>
<td>0.1607</td>
<td>0.042**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta_6 )</td>
<td>FOLL×SURP</td>
<td>-0.0092</td>
<td>0.133</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( R^2 \) | 0.0257 | 0.0441 | 0.0656 |
\( N \) | 927 | 927 | 927 |

Regressions of cumulated abnormal returns (CAR) on earnings expectations (DSURP) and an (interacted) dummy variable for the analyst forecast benchmark. Model 1 is the basic specification. Model 2 includes control variables known to affect the returns/earnings relation. Model 3 includes control variables and interactions of control variables and earnings expectations. All models are estimated with pooled OLS using Huber (1967)-White (1980, 1982) sandwich estimators of standard errors to control for heteroskedasticity. P-values are calculated using one-tailed tests if coefficients have predicted signs and two-tailed tests otherwise. Significance at the 10%, 5%, and 1% level is indicated by one (\( * \)), two (\( ** \)), and three (\( *** \)) asterisks, respectively.
5. The Market Reaction to Benchmark Beating

5.4.2 The Zero Earnings Benchmark

Table 5.8 presents univariate results for the market response to positive and negative EPS. If zero earnings is a focal point in investors’ decision processes, reporting a profit should be rewarded with significantly higher returns than reporting a loss. The last line in Table 5.8 compares pooled three-day abnormal returns of firms at or above (MBE) and below (MISS) the zero earnings benchmark: Average CAR is positive for the MBE-group (0.22%) and negative for the MISS-group (−0.32%). However, the difference between MBE- and MISS-observations is, though close to the 10% level, not significant at generally accepted levels. The yearly analysis draws a similar picture: Direct comparison of the MBE- and MISS-group on a yearly level reveals that average abnormal returns of the MBE-group are greater than those of the MISS-group in three of the five sample years. However, only for two years, these differences are statistically significant at least 10%. Overall, the univariate analysis does not suggest that the market regards reporting a profit as an important earnings target.

<table>
<thead>
<tr>
<th>Year</th>
<th>MISS</th>
<th>MBE</th>
<th>MBE–MISS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>p-value</td>
</tr>
<tr>
<td>2005</td>
<td>23</td>
<td>−0.0148</td>
<td>0.0357**</td>
</tr>
<tr>
<td>2006</td>
<td>24</td>
<td>0.0071</td>
<td>0.8062</td>
</tr>
<tr>
<td>2007</td>
<td>24</td>
<td>−0.0046</td>
<td>0.3623</td>
</tr>
<tr>
<td>2008</td>
<td>35</td>
<td>−0.0129</td>
<td>0.1585</td>
</tr>
<tr>
<td>2009</td>
<td>62</td>
<td>0.0031</td>
<td>0.7031</td>
</tr>
<tr>
<td>Pooled</td>
<td>168</td>
<td>−0.0032</td>
<td>0.2197</td>
</tr>
</tbody>
</table>

Columns II to VII depict cumulated abnormal returns for a three-day window spanning from one day before until one day after the earnings announcement date conditional on whether the zero earnings benchmark is missed (MISS) or achieved (MBE) with related p-values. Columns VIII and IX show differences in mean returns and related p-values. All p-values are based on one-tailed mean comparison t-tests. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**) and three (***) asterisks, respectively.

The regression approach controls for other factors that may affect the market’s reaction to loss avoidance. If meeting or beating zero earnings is rewarded with a premium, the coefficient on $MBE^{ZE}$ ($α_2$) should be significantly greater than zero. Furthermore, achieving the zero earnings benchmark may affect the pricing of unexpected earnings. The coefficient on $MBE^{ZE} \times DSURP$ ($β_2$) measures this additional market reaction per unit of earnings surprise. The results in Table 5.9 provide ambiguous evidence of market rewards for loss avoidance. In Model 1 and 2, the coefficients on $MBE^{ZE}$ are positive, but not significant. In Model 3, $α_2$ suggests that avoiding a loss yields an incremental three-day CAR of 0.8%. However, the result is just barely significant.
at a level close to 10\% (p-value: 0.094). The interaction of $DSURP$ and $MBE^{ZE}$ is significantly positive for all three specifications, indicating a higher value relevance of unexpected earnings when the firm reports profits. This finding is generally consistent with Hayn (1995), who documents that losses are less informative with respect to future cash flows than profits because shareholders have an abandonment option that preserves them from indefinite losses.

Overall, I do not find sufficient evidence to document a market premium for meeting or beating the zero earnings benchmark. The results suggest, however, that reporting a profit makes unexpected earnings more value relevant. Meeting or beating the zero earnings benchmark could thus be a reasonable strategy when the company additionally beats the analyst consensus.
### Table 5.9
Market Rewards for Avoiding a Loss—Regression Results

Model 1: \( CAR = \alpha_1 + \alpha_2 MBZE + \beta_1 DSURP + \beta_2 MBZE \times DSURP + \varepsilon \)

Model 2: \( CAR = \alpha_1 + \alpha_2 MBZE + \beta_1 DSURP + \beta_2 MBZE \times DSURP + \sum_{j=1}^{J} \gamma_j Z_j + \varepsilon \)

Model 3: \( CAR = \alpha_1 + \alpha_2 MBZE + \beta_1 DSURP + \beta_2 MBZE \times DSURP + \sum_{j=1}^{J} \gamma_j Z_j + \sum_{k=1}^{K} \delta_k (Z_k \times DSURP) + \varepsilon \)

<table>
<thead>
<tr>
<th>Exp.</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>p-value</td>
<td>Coefficient</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>CONST</td>
<td>-0.0010</td>
<td>0.590</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>MBZE</td>
<td>+0.0010</td>
<td>0.412</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>DSURP</td>
<td>+0.0402</td>
<td>0.071*</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>MBZE \times DSURP</td>
<td>+0.1231</td>
<td>0.034**</td>
</tr>
<tr>
<td>( \gamma_1 )</td>
<td>PRET</td>
<td>-0.0016</td>
<td>0.692</td>
</tr>
<tr>
<td>( \gamma_2 )</td>
<td>RISK</td>
<td>+0.0013</td>
<td>0.937</td>
</tr>
<tr>
<td>( \gamma_3 )</td>
<td>GROW</td>
<td>-0.0009</td>
<td>0.286</td>
</tr>
<tr>
<td>( \gamma_4 )</td>
<td>SIZE</td>
<td>-0.0147</td>
<td>0.044**</td>
</tr>
<tr>
<td>( \gamma_5 )</td>
<td>DRISK</td>
<td>+0.0008</td>
<td>0.583</td>
</tr>
<tr>
<td>( \gamma_6 )</td>
<td>PERST</td>
<td>-0.0001</td>
<td>0.723</td>
</tr>
<tr>
<td>( \gamma_7 )</td>
<td>AGE</td>
<td>-0.0002</td>
<td>0.718</td>
</tr>
<tr>
<td>( \delta_1 )</td>
<td>PRET \times DSURP</td>
<td>-0.0222</td>
<td>0.904</td>
</tr>
<tr>
<td>( \delta_2 )</td>
<td>RISK \times DSURP</td>
<td>-0.0277</td>
<td>0.449</td>
</tr>
<tr>
<td>( \delta_3 )</td>
<td>GROW \times DSURP</td>
<td>0.0096</td>
<td>0.474</td>
</tr>
<tr>
<td>( \delta_4 )</td>
<td>DRISK \times DSURP</td>
<td>0.2685</td>
<td>0.002***</td>
</tr>
<tr>
<td>( \delta_5 )</td>
<td>PERST \times DSURP</td>
<td>0.1084</td>
<td>0.423</td>
</tr>
<tr>
<td>( \delta_6 )</td>
<td>FOLL \times DSURP</td>
<td>-0.0083</td>
<td>0.180</td>
</tr>
</tbody>
</table>

| \( R^2 \) | 0.0213 | 0.0390 | 0.0601 |
| N | 927 | 927 | 927 |

Regressions of cumulated abnormal returns (\( CAR \)) on earnings expectations (\( DSURP \)) and an (interacted) dummy variable for the zero earnings benchmark. **Model 1** is the basic specification. **Model 2** includes control variables known to affect the returns/earnings relation. **Model 3** includes control variables and interactions of control variables and earnings expectations. All models are estimated with pooled OLS using Huber (1967)-White (1980, 1982) sandwich estimators of standard errors to control for heteroskedasticity. \( p \)-values are calculated using one-tailed tests if coefficients have predicted signs and two-tailed tests otherwise. Significance at the 10%, 5%, and 1% level is indicated by one (**), two (**), and three (***) asterisks, respectively.
5. The Market Reaction to Benchmark Beating

5.4.3 The Earnings Changes Benchmark

Table 5.10 compares mean three-day cumulated abnormal returns of firm-years with zero or positive (MBE) and negative change in EPS (MISS). If the market rewards benchmark achievement with a premium and penalizes firms that fall short of prior year’s EPS, I expect mean abnormal returns of the MBE-group to be significantly higher than those of the MISS-group. The results for the pooled sample are shown in the last line of Table 5.10. As expected under the hypothesis of benchmark importance, the MISS-group exhibits negative abnormal returns of $-0.48\%$, while the MBE-group earns positive returns of $0.59\%$ on average. Both means are significantly below/above zero at a significance level of at least 5%. The last two columns suggest that achieving last year’s EPS is (on average) rewarded with an incremental three-day return of $1.07\%$. This result is highly significant below the 1% level. The yearly analysis supports pooled results: All differences in means reported in the last two columns are positive and significant at at least 5% in three of five sample years. Overall, the univariate analysis suggests a positive relation between meeting or beating last year’s EPS and the short-term stock performance at the time of the earnings announcement.

I use the pooled regressions described in Section 5.1 to test for an incremental premium after controlling for the information content in earnings and other control variables. If achieving last year’s EPS matters to investors, I expect the coefficient on $MBE^{EC}$ ($\alpha_2$) to be significantly positive. Additionally, benchmark beating may affect the pricing of unexpected earnings. This effect is measured as the interacted ERC ($\beta_2$). The regression results in Table 5.11 provide evidence in
strong support of benchmark importance. In the basic specification (Model 1), missing the benchmark is associated with a negative abnormal return of $-0.03\%$; a result barely significant at the 10% level. Achieving last year’s earnings, however, yields an incremental cumulated abnormal return of 0.09% significant at the 1% level. Furthermore, meeting the earnings changes benchmark seems to affect the market response to unexpected earnings. The interacted ERC ($\beta_2$) is positive and highly significant at the 1% level. A possible explanation for this finding is provided by Basu (1997), who shows that negative earnings changes are less persistent than positive earnings changes due to accounting conservatism. Hence, achieving the earnings changes benchmark is regarded as a signal for more persistent earnings information and priced accordingly.

The results remain qualitatively unchanged when control variables (Model 2) and interacted control variables (Model 3) are included. In both models, the coefficient on $MBE^{FC}$ ($\alpha_2$) is clearly positive and highly significant. Moreover, the premium for benchmark achievement even increases to 0.95% (Model 2) and 1.05% (Model 3) when control variables are included. The interacted ERC ($\beta_2$) slightly decreases in Model 2 and increases in Model 3, while remaining highly significant at the 1% level.

Overall, the results suggest that avoiding earnings declines yields positive abnormal returns around the earnings announcement date. Moreover, achieving the benchmark seems to intensify the relation of unexpected earnings and abnormal returns.
## Table 5.11

*Market Rewards for Avoiding EPS Declines—Regression Results*

<table>
<thead>
<tr>
<th></th>
<th>Model 1: ( CAR = \alpha_1 + \alpha_2 \text{MBE} + \beta_1 \text{DSURP} + \beta_2 \text{MBE} \times \text{DSURP} + \epsilon )</th>
<th>Model 2: ( CAR = \alpha_1 + \alpha_2 \text{MBE} + \beta_1 \text{DSURP} + \beta_2 \text{MBE} \times \text{DSURP} + \sum_{j=1}^{J} \gamma_j Z_j + \epsilon )</th>
<th>Model 3: ( CAR = \alpha_1 + \alpha_2 \text{MBE} + \beta_1 \text{DSURP} + \beta_2 \text{MBE} \times \text{DSURP} + \sum_{j=1}^{J} \gamma_j Z_j + \sum_{k=1}^{K} \delta_k (Z_k \times \text{DSURP}) + \epsilon )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp.</td>
<td>Coefficient</td>
<td>p-value</td>
<td>Coefficient</td>
</tr>
<tr>
<td>( \alpha_1 ) CONST</td>
<td>-0.0034</td>
<td>0.091*</td>
<td>0.0079</td>
</tr>
<tr>
<td>( \alpha_2 ) MBE</td>
<td>0.0087</td>
<td>0.004***</td>
<td>0.0095</td>
</tr>
<tr>
<td>( \beta_1 ) DSURP</td>
<td>0.0306</td>
<td>0.118</td>
<td>0.0323</td>
</tr>
<tr>
<td>( \beta_2 ) MBE \times DSURP</td>
<td>0.1575</td>
<td>0.006***</td>
<td>0.1470</td>
</tr>
<tr>
<td>( \gamma_1 ) PRET</td>
<td>-0.0393</td>
<td>0.023**</td>
<td>-0.0372</td>
</tr>
<tr>
<td>( \gamma_2 ) SRISK</td>
<td>0.0011</td>
<td>0.636</td>
<td>-0.0002</td>
</tr>
<tr>
<td>( \gamma_3 ) GROW</td>
<td>-0.0014</td>
<td>0.965</td>
<td>-0.0011</td>
</tr>
<tr>
<td>( \gamma_4 ) SIZE</td>
<td>-0.0010</td>
<td>0.265</td>
<td>-0.0008</td>
</tr>
<tr>
<td>( \gamma_5 ) DRISK</td>
<td>-0.0139</td>
<td>0.051*</td>
<td>-0.0082</td>
</tr>
<tr>
<td>( \gamma_6 ) PERST</td>
<td>-0.0009</td>
<td>0.603</td>
<td>0.0006</td>
</tr>
<tr>
<td>( \gamma_7 ) AGE</td>
<td>0.0001</td>
<td>0.667</td>
<td>0.0000</td>
</tr>
<tr>
<td>( \gamma_8 ) FOLL</td>
<td>-0.0001</td>
<td>0.671</td>
<td>-0.0002</td>
</tr>
<tr>
<td>( \delta_1 ) PRET \times DSURP</td>
<td>0.0504</td>
<td>0.768</td>
<td></td>
</tr>
<tr>
<td>( \delta_2 ) SRISK \times DSURP</td>
<td>-0.0337</td>
<td>0.349</td>
<td></td>
</tr>
<tr>
<td>( \delta_3 ) GROW \times DSURP</td>
<td>0.0046</td>
<td>0.711</td>
<td></td>
</tr>
<tr>
<td>( \delta_4 ) DRISK \times DSURP</td>
<td>0.2733</td>
<td>0.001***</td>
<td></td>
</tr>
<tr>
<td>( \delta_5 ) PERST \times DSURP</td>
<td>0.1876</td>
<td>0.052*</td>
<td></td>
</tr>
<tr>
<td>( \delta_6 ) FOLL \times DSURP</td>
<td>-0.0098</td>
<td>0.098*</td>
<td></td>
</tr>
</tbody>
</table>

\( R^2 \) 0.0245 0.0432 0.0660
N 927 927 927

Regressions of cumulated abnormal returns (\( CAR \)) on earnings expectations (\( DSURP \)) and an (interacted) dummy variable for the earnings changes benchmark. *Model 1* is the basic specification. *Model 2* includes control variables known to affect the returns/earnings relation. *Model 3* includes control variables and interactions of control variables and earnings expectations. All models are estimated with pooled OLS using Huber (1967)-White (1980, 1982) sandwich estimators of standard errors to control for heteroskedasticity. \( P \)-values are calculated using one-tailed tests if coefficients have predicted signs and two-tailed tests otherwise. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***)) asterisks, respectively.
5. The Market Reaction to Benchmark Beating

5.4.4 Robustness Tests

This section provides robustness tests for the analyses of stock market reactions to benchmark beating. Section 5.4.4.1 addresses potential non-linearities in the returns/earnings relationship. In Section 5.4.4.2, I test for biased coefficients and invalid test statistics due to omitted fixed-effects and within cluster correlation.

5.4.4.1 Non-linearities in the Returns/Earnings Relationship

Cheng et al. (1992) suspect the classical linear regression of abnormal returns on unexpected earnings to be misspecified due to non-linearities in the returns/earnings relation. Freeman and Tse (1992) and Das and Lev (1994) suggest that the function of abnormal returns and unexpected earnings is rather S-shaped than linear. More specifically, Freeman and Tse (1992) argue that the transitory (permanent) fraction of unexpected earnings is positively (negatively) correlated with the absolute value of unexpected earnings. Because transitory earnings are less value relevant, the marginal price response decreases with the absolute value of unexpected earnings and induces a S-shaped returns/earnings function. Linear and S-shaped returns/earnings relationships are illustrated in Figure 5.3.

![Linear and Non-linear Returns/Earnings Relationships](image)

**Fig. 5.3.**—**Linear and Non-linear Returns/Earnings Relationships.** The left figure displays a hypothetical linear returns/earnings relationship. The right figure illustrates a hypothetical S-shaped returns/earnings relationship as suggested in Freeman and Tse (1992).

Degeorge et al. (2005), Payne and Thomas (2009), and Herrmann et al. (2011) suspect that ignoring the S-shape of the returns/earnings relation under linear OLS biases the coefficients on benchmark achievement ($\alpha_2$) upward. To measure the incremental market reaction beyond the potentially S-shaped relation, Payne and Thomas (2009) and Herrmann et al. (2011) include
non-linear transformations of earnings surprises as additional explanatory variables. Specifically, Payne and Thomas (2009) include the square root of absolute earnings surprise (i.e., $\sqrt{|DSURP|}$) and Herrmann et al. (2011) squared earnings surprise (i.e., $DSURP^2$). Degeorge et al. (2005), in contrast, fit the following non-linear regression equation with non-linear least squares.\(^{21}\)

$$CAR = \alpha_1 + \alpha_2 MBE^Q + \beta_1 \arctan \left( e^{\beta_2} \times DSURP \right) + \ldots + \epsilon \quad (5.4.1)$$

To test whether my results are sensitive to the mentioned non-linearities, I rerun the basic model after including $DSURP^2$ and $\sqrt{|DSURP|}$ as additional independent variables. Furthermore, I follow Degeorge et al. (2005) and fit the non-linear regression in Eq. (5.4.1) by non-linear least squares. The results in Table 5.12 show that the coefficients and their significance remain qualitatively unchanged when controlling for potential non-linearities in the returns/earnings relation.

### Table 5.12

Controlling for Non-linearities

|           | Basic Model | $SURP^2$ | $\sqrt{|DSURP|}$ | Non-linear |
|-----------|-------------|----------|------------------|------------|
|           | Coeff. | p-value | Coeff. | p-value | Coeff. | p-value | Coeff. | p-value |
| $MBE^{FC}$ | 0.008  | 0.010*** | 0.009 | 0.006*** | 0.009 | 0.008*** | 0.010 | 0.005*** |
| $MBE^{ZE}$ | 0.001  | 0.412 | 0.009 | 0.412 | 0.004 | 0.186 | -0.003 | 0.762 |
| $MBE^{EC}$ | 0.009 | 0.004*** | 0.009 | 0.004*** | 0.009 | 0.003*** | 0.009 | 0.005*** |

The table compares the market reaction to benchmark achievement with and without controlling for non-linearities in the returns/earnings relationship. Column Basic Model depicts coefficients and related p-values without controlling for non-linearities. Columns $SURP^2$ and $\sqrt{|DSURP|}$ report coefficients and p-values after including squared deflated earnings surprises and the square root of absolute deflated earnings surprises, respectively. Column Non-linear contains the coefficient estimate $\hat{\alpha}_2$ and related p-values from fitting

$$CAR = \alpha_1 + \alpha_2 MBE^Q + \beta_1 \arctan \left( e^{\beta_2} \times DSURP \right) + \epsilon$$

by non-linear least squares using the Gauss-Newton iterative method. Since all coefficients have predicted signs (+), p-values are based on one-tailed tests. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***) asterisks, respectively.

---

\(^{21}\)Their original notation is adapted to fit the notation in this chapter. Independent variables not required in my analysis are dropped from the original model. Subscripts are omitted for notational convenience.
5.4.4.2 Specification Issues

The regression results reported in this chapter are based on pooled OLS regressions of panel data. Two concerns are noteworthy when pooled OLS is applied in the presence of panel data:

- **Omitted Variables.** Omitted variables (e.g., fixed industry or time effects) may cause endogeneity and yield biased estimates.\(^{22}\) To test whether omitted industry and/or time fixed-effects affect my results, I rerun the regressions after including dummy variables for industry (two-digit SIC) and year.\(^{23}\) The results do not suggest that the coefficients are biased due to endogeneity.

- **Within Cluster Correlation.** Correlated residuals within clusters (e.g., firm or time cluster) cause bias in standard errors and coefficient variability. Although coefficient estimates remain unbiased, hypotheses tests may be unreliable and inferences spurious.\(^{24}\) To account for a possible correlation of residuals within firm clusters (i.e., serial dependence), I rerun the regressions with standard errors clustered by firms.\(^{25}\) The results (not reported) do not indicate that residual correlation within firm clusters biases standard errors. To account for cross-sectional residual correlation, I follow Lopez and Rees (2002) and use the Fama-MacBeth procedure (Fama and MacBeth, 1973).\(^{26}\) The results remain qualitatively unchanged.

Overall, the results do not seem to be spurious due to omitted fixed-effects and/or within cluster correlation.


\(^{23}\) For details on identification and treatment of unobservable fixed-effects in panel regressions refer to, e.g., Wooldridge (2002, pp. 265–279).

\(^{24}\) The problem of correlated residuals in panel data sets in the context of finance and accounting research is described in detail by Petersen (2009) and Gow et al. (2010).

\(^{25}\) Cluster-robust standard errors take into account that errors are correlated (e.g., serially or cross-sectionally) within but not between clusters and adjust for within cluster correlations. They are often referred to as Rogers standard errors. For a description of cluster-robust standard errors see, e.g., Petersen (2009) and Gow et al. (2010).

\(^{26}\) In a recent simulation study, Petersen (2009) confirms the Fama-MacBeth procedure to be effective in the presence of cross-sectional dependence of regression residuals.
5.5 Summary

In this chapter, I analyzed the market response to meeting or beating earnings benchmarks. If benchmarks are indeed used as focal points in investors’ decision processes or facilitate contracting with the stakeholders of a company, I expect the market to react positive on benchmark achievement. Two major insights can be drawn from the evidence in this chapter. First, the German capital market rewards firms that meet or beat the latest analyst EPS forecast or last year’s EPS, irrespective of the information content in current earnings. Achieving the latest analyst consensus forecast is rewarded with an additional $0.81\%$ of abnormal return cumulated over the three days around the earnings announcement. Meeting or beating the last year’s EPS is slightly more favorable with an additional abnormal return of $0.87\%$. These results do, however, not hold for the zero earnings benchmark: Reporting a profit does not yield a significant incremental market return. Second, the market reacts more sensitive to unexpected earnings when one of the three benchmarks is met. While earnings responses are not significantly different from zero for benchmark missers, benchmark achievers exhibit significant positive ERCs. This result suggests that earnings information is considered as less relevant when firms fall short of analyst forecasts, report a loss, or an EPS decline. For benchmark achievers, however, the information conveyed in unexpected earnings is reflected in asset prices.

Overall, the results reveal the importance of analyst forecasts and prior year’s EPS as earnings benchmarks from a capital market perspective. Since the capital market responds positive to these benchmarks, I assume them to provide incentives for earnings management. Whether managers really engage in earnings management to achieve benchmarks is the underlying research question for the empirical analyses in the following two chapters.
Evidence from the previous chapter suggests that the capital market rewards benchmark achievement with a premium. As a result, managers are suspected to engage in earnings management if they would otherwise miss an important benchmark. Following Burgstahler and Dichev (1997) and DeGeorge et al. (1999), a large body of research has addressed the frequency and prevalence of earnings management to meet earnings benchmarks (see the literature review in Section 3.2.1). In the spirit of these studies, I expect that the manipulation of earnings across thresholds is reflected in the shapes of their frequency distributions. Specifically, shifting observations from the “small miss” to the “small beat” region induces a shortfall of observations immediately below and a pile-up of observations just above the benchmark. The related hypotheses (in their alternative form) are stated as follows:

$$H_A (2a):$$ The frequency distribution of German earnings surprises exhibits a discontinuity around zero. Specifically, there are significantly more observations than expected in the region immediately above and less observations than expected just below zero.

$$H_A (2b):$$ The frequency distribution of German earnings levels exhibits a discontinuity around zero. Specifically, there are significantly more observations than expected in the region immediately above and less observations than expected just below zero.
The frequency distribution of German earnings changes exhibits a discontinuity around zero. Specifically, there are significantly more observations than expected in the region immediately above and less observations than expected just below zero.

The remainder of this chapter is structured as follows: In Sections 6.1 and 6.2, I provide details on the methodology underlying the empirical tests in this chapter and variable definitions, respectively. Section 6.3 covers the sample selection process and descriptive statistics. The empirical results are presented in Section 6.4. The chapter closes with a brief summary in Section 6.5.

### 6.1 Empirical Methodology

Starting with the seminal study by Burgstahler and Dichev (1997) a new methodology rapidly spread out and paved ground for a new strand of research on benchmark-driven earnings management. Burgstahler and Dichev (1997) hypothesize that firms manage earnings upward if they would otherwise report a loss or fall short of prior year’s earnings. To test their hypothesis, they examine the pooled distribution of scaled earnings and changes in earnings. If executives manage earnings upward to avoid a loss, the distribution of scaled earnings is expected to exhibit a significant drop just below zero earnings and a pile-up of observations just above. Similarly, if small declines in earnings are managed away, the distribution of earnings changes (i.e., current earnings minus prior year’s earnings) is expected to exhibit a discontinuity at zero. Similar analyses can be performed at other benchmarks: Earnings or expectations management to avoid missing the latest analyst consensus forecast, for example, is expected to induce an irregularity in the distribution of earnings surprises (i.e., actual earnings minus the latest analyst consensus estimate). Throughout this study, I refer to the Burgstahler and Dichev (1997)-methodology as the distributional approach.

The distributional approach usually involves two techniques: Visual inspection (Section 6.1.1) and formal statistical analysis (Section 6.1.2) of distributional discontinuity. Both are

---

1Throughout this study, I use the term “distributional discontinuity” or “distributional irregularity” as interchangeable shorthand terminology for an unusually low frequency of observations immediately below and an unusually high frequency of observations immediately above the benchmark of interest.
6. The Prevalence of Benchmark Beating

6.1.1 Visual Inspection of Earnings Distributions

Visual inspection is based on a frequency plot of the pooled cross-section of (scaled) earnings metrics (e.g., earnings or changes in earnings).\(^2\) Figure 6.1 depicts the frequency distribution of pooled earnings deflated by lagged market value for US firms between 1976 and 1994 as presented in Burgstahler and Dichev (1997). The distribution exhibits a considerable discontinuity around zero earnings: Whereas the distribution is relatively smooth in regions significantly below and above zero, there is a sharp drop of observations immediately below and a considerable pile-up of observations immediately above zero. Burgstahler and Dichev (1997) interpret this distributional pattern as evidence of earnings management to turn small losses into small profits.

![Figure 6.1: Pooled Distribution of Scaled US Earnings](image)

**Fig. 6.1.—Pooled Distribution of Scaled US Earnings.** The figure shows the frequency distribution of annual earnings scaled by lagged market capitalization for 75,999 US firm-years between 1976 and 1994. Interval width of the histogram is 0.005. The figure is taken from Burgstahler and Dichev (1997, p. 109).

Visual inspection using frequency plots is attractive in its simplicity. However, the distribu-

\(^2\)Scaling is used to account for differences in firm size (Burgstahler and Dichev, 1997). Although there is no consensus about the best suitable scaling factor, many studies follow Burgstahler and Dichev (1997) and deflate earnings by lagged market value. Recent studies based on lagged market value as scaling factor include, e.g., Beaver et al. (2007), Jacob and Jorgensen (2007), and Kerstein and Rai (2007). Durtschi and Easton (2005, 2009), however, argue that scaling distorts the underlying distribution and causes a distributional irregularity which is erroneously interpreted as evidence for earnings management.
tional shape strongly depends on the underlying histogram parameters which require cautious calibration to avoid erroneous inferences. To make this point clear, recall that a histogram is defined as (see, e.g., Härdle et al., 2004)

$$\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^{N} \sum_{j} I(X_i \in B_j)I(x \in B_j). \tag{6.1.1}$$

where $\hat{f}(x)$ denotes the estimated density of the underlying population at $x$, $h$ is the width of each histogram bin or interval, $N$ is sample size, and $\sum_{i=1}^{N} \sum_{j} I(X_i \in B_j)I(x \in B_j)$ is the number of observations in the same interval $B_j$ as $x$. Each interval is defined as $B_j = [x_0 + (j - 1)h, x_0 + jh)$ with $j \in \mathbb{Z}$ and $x_0$ denoting the origin of the histogram. From Eq. (6.1.1) it is obvious that histogram origin ($x_0$) and binwidth ($h$) influence the shape of the histogram (Scott, 2010):

- **Histogram Origin.** The histogram origin $x_0$ defines the exact placement of the histogram bins and is often chosen to be the smallest observed data value (Scott, 2010). In the field of earnings management, the researcher is interested in distributional irregularities around a certain threshold. The placement of histogram bins is thus defined by the tested hypothesis. Testing for earnings management around zero earnings, for instance, requires to choose the origin so that zero earnings is the lowest value in one of the histogram bins. Hence, origin choice is not a central issue when applying the distributional approach.

- **Binwidth.** The importance of binwidth choice is illustrated in Figure 6.2. Panel B depicts an exemplary histogram of (scaled) earnings with a binwidth of 0.05. The graph suggests an approximately bell-shaped distribution with a kink adjacent to the threshold at zero. In contrast, a smaller binwidth of 0.02 in Panel A yields a considerably fine structure with several small peaks. Apparently, reducing binwidth decreases histogram smoothness. This becomes even more apparent, when binwidth is increased to 0.08 in Panel C. In comparison to Panel A and B, the distribution becomes so smooth that the kink around zero earnings vanishes. The example illustrates that inferences from visual inspection critically hinge on binwidth choice. Specifically, “ [...] the choice of an interval width must be balanced against the fact that if it is too small, then the spurious fine structure becomes visible [whereas] if it is too large the essential detail is masked” (Holland, 2004, p. 3). The researcher is relatively flexible in setting binwidth. As a result, many authors choose

---

Note that $I(X_i \in B_j) = 1$ if $X_i \in B_j$ and 0 otherwise. Similarly, $I(x \in B_j) = 1$ if $x \in B_j$ and 0 otherwise.
binwidth from eyeball in a rather “ad-hoc” fashion (e.g., Burgstahler and Dichev, 1997; Coppens and Peek, 2005; Ayers et al., 2006; Beaver et al., 2007). On more solid grounds, others test the robustness of their results with several different binwidths (e.g., Coulton et al., 2005) or start with “rules of thumb” that evolved in the discipline of statistics and subsequently adapt these binwidths to their specific setting (e.g., Degeorge et al., 1999; Beatty et al., 2002; Dichev and Skinner, 2002; Glaum et al., 2004; Daske et al., 2006). Rules of thumb are a good starting point for binwidth choice, as their application prevents failure in detecting the essential structure of the data. Several different approaches have been developed to calculate the optimal binwidth from the dataset at hand. Some of these rules are described in detail in Appendix A.4.

![Histograms A, B, and C](image)

**FIG. 6.2.—Effect of Binwidth Choice.** The three exemplary frequency plots of scaled earnings illustrate the effect of binwidth choice on histogram shape. The histograms in Panel A, B, and C have a binwidth of 0.02 (100 bins), 0.05 (40 bins), and 0.08 (25 bins), respectively.

### 6.1.2 Statistical Tests of Distributional Discontinuity

Although visual inspection is an intuitive and comfortable approach to identify distributional irregularities around benchmarks, it lacks formal statistical results. To overcome this limitation,
Burgstahler and Dichev (1997), Degeorge et al. (1999), and Bollen and Pool (2009), among others, developed statistical procedures to test for distributional irregularities. Two of these approaches are described in the remainder of this section.

### 6.1.2.1 The Burgstahler and Dichev Test Statistic

The Burgstahler and Dichev test statistic (BD test statistic) or “standardized difference” is based on the assumption that absent earnings management, the distribution of earnings is relatively smooth. Under this assumption, the number of observations residing in one interval of the distribution is expected to be the average of observations in the two immediately adjacent intervals.\(^4\) A significant difference between the actual and this expected number of observations is considered as evidence of a discontinuity in the distribution. If, for example, executives manage earnings from small losses to small profits, the difference should be significantly positive for the first interval above and negative for the first interval below zero earnings. Let \(X_j\) denote the actual number of observations and \(E(X_j)\) the expected number of observations in interval \(j\), calculated as average of the observations in the immediately adjacent intervals \(j - 1\) and \(j + 1\). With \(X_j\) following a binomial distribution and the number of observations \(N\) large enough for the central limit theorem to take effect, dividing the difference of \(X_j\) and \(E(X_j)\) by its standard deviation \(\sigma(X_j - E(X_j))\) yields a test statistic that is distributed approximately standard normal. This test statistic \(\tau_{BD}^{j}\) for interval \(j\) is given as

\[
\tau_{BD}^{j} = \frac{X_j - E(X_j)}{\sigma(X_j - E(X_j))}.
\]  

(6.1.2)

Defining the probability that an observation resides in bin \(j\) as \(p_j\) and given that \(X_j \sim B(N, p_j)\), \(E(X_j)\) and \(Var(X_j)\) can be expressed as \(Np_j\) and \(Np_j(1 - p_j)\), respectively. If the observations in \(j - 1\) and \(j + 1\) are approximately independent, the standard deviation of the difference between \(X_j\) and \(E(X_j)\) is

\(^4\)In additional tests, Burgstahler and Dichev (1997) use the average of observations in the four bins adjacent to the threshold and the average of observations in the bins next to the directly adjacent intervals. These alternative models of expectations yield qualitatively similar results. Throughout this study, the BD test statistic is based on average observations in the directly adjacent bins.
To test for earnings management, the test statistic $\tau_{j}^{BD}$ is calculated for the intervals immediately adjacent to the benchmark of interest. Since standardized differences are distributed approximately normal, values of $\tau_{j}^{BD}$ above $|2.33|$ ($|1.65|$) imply an irregularity at a significance level below 1% (5%) (one-tailed). As an example consider the distribution of net income scaled by lagged market value in Figure 6.1 taken from Burgstahler and Dichev (1997, p. 109). For the the first interval above zero, the BD test statistic $\tau_{i}^{BD}$ is 5.88. Hence, the interval contains significantly more observations than expected under the null of a smooth distribution. For the first interval below zero $\tau_{i}^{BD}$ is $-8.00$, indicating a significant underrepresentation of observations in that interval. Both results strongly support the hypothesis that executives manage small losses away.

Owing to its intuitive construction and computational ease, the BD test statistic has become a popular measure for distributional irregularities within accounting research. Despite these advantages, several recent papers point out some major shortcomings of the approach:

---

**Independence of Observations.** Burgstahler and Dichev (1997) assume the observations in adjacent bins to be distributed independently and thus the covariance between observations to be zero. Bollen and Pool (2010) use a modified version of the original BD test statistic to test for distributional irregularities in the context of hedge fund returns. In contrast to the BD test statistic, their measure accounts for the covariance of observations in adjacent bins because observations that fall into one bin of the histogram may naturally not lie in any other bin. Since the covariance of the observations in adjacent bins is equal to

$$
Cov(X_{j}, X_{j-1}) = -Np_{j}p_{j-1},
$$

the modified standard deviation of the BD test statistic ($\sigma^{*}_{(X_{j} - E(X_{j}))}$) is given as

$$
\sigma^{*}_{(X_{j} - E(X_{j}))} = \sqrt{\left(\sigma_{(X_{j} - E(X_{j}))}\right)^2 + Np_{j}(p_{j-1} + p_{j+1})}.
$$

---

5These values represent the 5th and 95th (1st and 99th) percentiles of a student-t distribution with $\infty$ degrees of freedom or (equivalently) a standard normal distribution.
Neglecting the term \( Np_j(p_{j-1} + p_{j+1}) \) of \( \sigma^* \) results in an understatement of the standard deviation and an overstatement of the test statistic. Hence, adjusting the denominator of the BD test statistic for the covariance between bins as laid out by Bollen and Pool (2010) yields a more conservative measure of statistical irregularity. Therefore the results reported in the course of this dissertation may be lower than those of other studies applying the “original” BD test statistic.

- **Validity around the Distribution’s Mode.** The BD test statistic is biased towards over-rejecting the null of no earnings management when the threshold of interest is directly adjacent to the peak of the distribution (Holland, 2004). The hypothetical distribution of earnings in Figure 6.3 demonstrates this issue: The distribution appears smooth in all regions of the histogram and visual inspection does not suggest any discontinuities. If the benchmark lies, e.g., between bin \(-5\) and \(-4\), the expected number of observations in the first bin above it is given as \( E(X_{-4}) = \frac{1}{2}(432 + 610) = 521 \). In the figure, this expectation is visualized as the dashed line crossing the adjacent bins. Since the actual number of observations in bin \(-4\) is only one observation below the expected number, the BD test statistic is insignificant \( \tau_{BD}^{-4} = -0.04 \) and does not reject the null of a smooth distribution. If, however, the threshold is adjacent to the peak of the distribution (i.e., between bin \(-1\) and \(+1\)), bin \(+1\) contains 67 observations more than expected under distributional smoothness \( E(X_{+1}) = \frac{1}{2}(725 + 725) = 725 \). In contrast to visual inspection, the BD test statistic \( \tau_{BD}^{+1} = 1.97 \) suggests a discontinuity significant at the 5% level. However, since the distribution peaks at bin \(+1\), the actual number of observations in this interval is per definition above its expected value calculated as the average of the directly adjacent intervals and illustrated with the solid horizontal line (Holland, 2004). As a result, the BD test statistic is per definition positive and prone to erroneously indicate a significant over-representation of observations. Hence, BD test statistics have to be treated with caution when the benchmark of interest coincides with the distribution’s mode.

- **Normality in the Parent Distribution.** Christodoulou and McLeay (2009) remark that inferences about discontinuities in the distribution of accounting earnings are usually based on the assumption of normality in the limit. In other words, one would expect the distribution of earnings to be approximately Gaussian when the sample size gets close to the population’s limit. The assumption of normality, however, does not seem to be justifiable in this context. Thus, finding alternative distributions may improve the power and accu-
racy of distribution-based inferences on earnings management. A non-parametric approach based on kernel density estimates is one alternative method to address non-normalities in the parent distribution. This method is explained in Section 6.1.2.2.

6.1.2.2 Test Statistic based on Kernel Density Estimates

Two approaches may be used to find a more appropriate distribution than under the assumption of asymptotic normality. In studying the distribution of cash flows, Zhang (2009) uses a parametric approach. He first chooses the best fitting density function from a number of candidate density families and subsequently calibrates its parameters to fit it to the underlying sample distribution. In the context of hedge fund misreporting, Bollen and Pool (2009) use a non-parametric approach based on kernel density estimation to model a smooth distribution of (expected) fund returns. As a non-parametric approach, their method does not assume any predefined density function and uses the sample observations to estimate a unique reference density function instead. In this study, I use their approach to test for discontinuities in the distribution of earnings variables. The basic concepts of non-parametric density estimation and its application in the context of earnings management research are summarized in the remainder of this section.

*Bollen and Pool (2009) test whether hedge fund managers overstate returns to avoid reporting small losses on their portfolios. Their refined methodology is based on the idea of the distributional approach developed in earnings management research and can be applied in both contexts.
Kernel density estimation is used to estimate the underlying density function of a sample distribution.\(^7\) In comparison to a histogram estimator, the kernel method uses significantly more data points and is based on a continuous weighting function which provides some striking advantages. These include the allowance for non-normalities in the parent distribution and valid expectations when the distribution peaks near the benchmark of interest. To understand the idea of kernel density estimation, consider the exemplary distribution of scaled earnings in Figure 6.2, Panel B. It has been shown in Section 6.1.1 that the histogram critically hinges on the choice of binwidth and origin. Now, instead of fixing binwidth and origin, a naive estimator can be derived by letting a window \([x - w, x + w]\) continuously slide over the full range of sample observations. The naive estimator is formally described as

\[
\hat{f}(x) = \frac{1}{2Nw} \sum_{i=1}^{N} I(X_i \in [x - w, x + w]),
\]  

(6.1.6)

where \(N\) is the sample size, \(w\) denotes half of the window width, and \(\sum_{i=1}^{N} I(X_i \in [x - w, x + w])\) the number of observations falling into the same window \([x - w, x + w]\) as \(x\). More transparently, one can think of putting a “box” with an area of 1, width 2\(w\) and height 1/2\(w\) on every observation, sum up the boxes over \(x\) and then divide by \(N\). Formally\(^10\), introducing the weighting function

\[
K(u) = \begin{cases} 
\frac{1}{2} & |u| < 1 \\
0 & \text{otherwise}, 
\end{cases}
\]  

(6.1.7)

allows to rewrite the naive estimator in Eq. 6.1.6 as

\[
\hat{f}(x) = \frac{1}{Nw} \sum_{i=1}^{N} K \left( \frac{x - X_i}{w} \right).
\]  

(6.1.8)

Compared to the histogram, the naive estimator yields a seemingly continuous distribution func-
tion. However, the function still exhibits jumps at the points $X_i \pm w$, giving the distribution a rather ragged shape. The reason for this shape lies in the underlying weighting function $K(u)$, since calculating density estimates as “sum of boxes” yields a step function. To avoid this property, the uniform weighting function $K(u)$ in Eq. (6.1.7) can be replaced with a continuous weighting function (or kernel function). When $K(u)$ is continuous, then $\hat{f}(x)$ will also be a continuous density function that shares its differentiability properties with the kernel function (Silverman, 1986, p. 15). Assuming for instance that $K(u)$ is Gaussian (i.e., $K(u) = \frac{1}{\sqrt{2\pi}} e^{(-\frac{1}{2}u^2)}$), all $X_i \in [x + w, x - w]$ closer to $x$ will be assigned with more weight. The solid line in Figure 6.4 depicts a kernel estimate for the underlying density function of the exemplary earnings distribution previously shown in Figure 6.2, Panel B.

![Figure 6.4](image)

**FIG. 6.4.—Kernel Estimate of an Earnings Density Function.** The figure shows a kernel density estimate based on the Gaussian weighting function $K(u) = \frac{1}{\sqrt{2\pi}} e^{(-\frac{1}{2}u^2)}$ for an exemplary distribution of scaled earnings.

The estimated reference distribution allows to calculate the probability $p_j$ that an observation falls in bin $B_j = [x_0 + (j-1)h, x_0 + jh)$, with $j \in \mathbb{Z}$, by means of integration over the bin’s range:

$$p_j = \int_{x_0 + (j-1)h}^{x_0 + jh} \hat{f}(x) \, dx. \quad (6.1.9)$$

Since the number of observations $X_j$ that reside in bin $j$ are distributed binomially with $X_j \sim B(N, p_j)$ and $N$ is large enough for the central limit theorem to take effect, $X_j$ is distributed approximately normal with $X_j \sim N(Np_j, Np_j(1 - p_j))$. Given $p_j$ now allows to calculate the Bollen and Pool test statistic (BP test statistic) as
6. The Prevalence of Benchmark Beating

\[ \tau_j^{BP} = \frac{X_j - E(X_j)}{\sigma(X_j - E(X_j))} = \frac{X_j - Np_j}{\sqrt{Np_j(1 - p_j)}}. \]  

(6.1.10)

The standardized test statistic \( \tau_j^{BP} \) is distributed approximately normal with \( \tau_j^{BP} \sim N(0, 1) \). Hence, values above \( |2.33| \) (\( |1.65| \)) indicate a rejection of the null of no discontinuity below the 1% (5%) significance level and are regarded as evidence for managerial discretion.

The BP approach depends on some critical factors which have to be taken in careful consideration. These include the choice of kernel bandwidth \( w \) and a kernel function \( K \):

- **Kernel Bandwidth.** Estimating the reference distribution requires the choice of a proper window width or bandwidth.\(^{11}\) The bandwidth determines the shape of the estimated density function. Larger bandwidths yield smooth but less detailed functions, while smaller bandwidths draw a more precise picture of the underlying data (see, e.g., the examples in Silverman, 1986, p. 16). Some trade-off has to be found between smoothness and precision to save important features of the underlying distribution without assigning too much weight on the fine and spurious structure of the sample at hand. As for the bin size in histogram estimation (see, e.g., Scott, 1979), researchers in the field of data analysis and statistics have developed simple heuristics to calculate optimal bandwidths from the properties of the underlying data. Following Bollen and Pool (2009), I use Silverman’s (1986) rule of thumb for Gaussian kernel functions and calculate the optimal bandwidth \( w^* \) as

\[ w^* \approx 0.9 \min \left( \hat{\sigma}, \frac{IQR}{1.34} \right) N^{-\frac{1}{5}}, \]  

(6.1.11)

with \( \hat{\sigma} \) denoting the sample standard deviation, \( IQR \) the interquartile range, and \( N \) sample size.\(^{12}\) Important to note, the calculation of the BP test statistic requires to choose the binwidth of the underlying histogram and the bandwidth of the kernel density estimator. For the main analyses in this chapter, I follow Bollen and Pool (2009) and set binwidth \( h \) equal to the optimal bandwidth \( w^* \). In additional analyses, however, I test the robustness of my results for several different bin- and bandwidth sizes (see Section 6.4.4.1).

- **Kernel Function.** In the example above, I use a Gaussian kernel to estimate the density of

---

\(^{11}\)Note that bandwidth is defined as \( \frac{1}{2} \times \) kernel window width.

\(^{12}\)See Silverman (1986, pp. 43–48) and/or Appendix A.4 for a description and the derivation of this rule of thumb.
the reference distribution. Kernel density estimation, however, does not require the kernel $K(u)$ to be Gaussian. Other popular choices are, for example, the Parzen kernel (Parzen, 1962), the Epanechnikov kernel (Epanechnikov, 1969), or the Biweight kernel. Silverman (1986, p. 43) remarks that the differences between these kernel functions are rather negligible in terms of minimizing mean integrated squared error. Hence, the choice of the underlying kernel function seems to be arbitrary in theory. In this study, I follow Bollen and Pool (2009) and choose a Gaussian kernel function for the main analyses. Though differences are assumed to be small in theory, I provide additional robustness tests with alternative kernel functions in Section 6.4.4.2.

6.1.3 Critical Evaluation of the Distributional Approach

Although the distributional approach is intuitive and widespread in the field of earnings management research, several scholars cast doubt on the validity of the results and caution to interpret distributional irregularities as evidence of earnings management. Dechow et al. (2003) challenge the question whether the kink in the earnings distribution is related to earnings management and provide a bunch of several alternative explanations:

- **Real Performance Improvement.** Efficient contracting aligns interests of management, employees and firm owners. Under efficient contracting, management and employees strive to make the firm profitable and expecting small losses provides strong incentives to work harder to cross the “red line”. Thus, the irregularity in the earnings distribution is rather attributable to efficient contracting than to managerial discretion.

- **Sample Selection Bias.** In the US, firms that do not fulfill certain profit and market capitalization thresholds run the risk of becoming delisted. The kink in the distribution around zero may thus simply stem from the fact that stock exchanges prefer profitable firms. In a similar vein, Durtschi and Easton (2005, 2009) show that variable requirements and database coverage censor more loss than profit observations from the sample and induce a discontinuity at the zero earnings benchmark that is unrelated to potential earnings management activity. I address this problem by using two different samples for the main analyses (see Section 6.3 for details on sample selection).

- **Conservatism.** Conservative accounting rules encourage the immediate recognition of losses (e.g., due to impairment charges) and prohibit the recognition of gains unless they
are realized (e.g., income realization in long-term contracts). The asymmetric treatment of gains and losses may shift profitable firms’ income towards zero and unprofitable firms in the high loss region, thus provoking a kink in the distribution of earnings.

- **Variable Scaling.** Deflation raises the strongest controversy within the research community. Beside other interpretations of a kink in the distribution of earnings, distributional irregularities may stem from inappropriate scaling of earnings variables. Instead of increasing homogeneity of the observations, inappropriate deflation potentially induces (or at least) intensifies discontinuities in earnings distributions. Durtschi and Easton (2005, 2009) support this argument and show that deflation by market capitalization and similar measures (e.g., total assets or total sales) significantly distorts the distribution of earnings. Beaver et al. (2007), Jacob and Jorgensen (2007), and Kerstein and Rai (2007), however, argue that deflating earnings by market value does not. The question of scaling remains an empirical one and depends on the underlying sample. Inferences drawn from distributional evidence thus require an in-depth analysis of the relation between earnings and the respective deflators (see Section 6.2) and should be underpinned by robustness test with alternative deflators (see the analyses in Section 6.4.4.3).

- **Asymmetric Tax Effects and Special Items.** Beaver et al. (2007) show that the discontinuity in the distribution of accounting earnings is intensified by differences in effective tax rates and negative special items for profit and loss firms. A higher effective tax rate for profit firms draws observations towards zero. A higher frequency of negative special items for loss firms shifts loss observations into the distribution’s tail. Both effects contribute to a pile-up of observations just above and a shortfall of observations just below zero. I address the potential influence of tax asymmetries in Section 6.4.4.4.

In light of this critical evaluation, evidence of earnings management derived from the distributional approach has to be treated with caution. In the empirical part of this chapter, I use several robustness tests to dispel other interpretations than earnings management. Moreover, I complement the distributional methodology with tests of specific earnings management activity around benchmarks in Chapter 7.
6.2 Variable Definitions

A test for distributional irregularities requires to define adequate earnings and forecast measures. Most of the studies that follow Burgstahler and Dichev (1997) choose deflated variables to test for benchmark related earnings management. Deflating net income by some size proxy seems reasonable since the distribution of net income covers a wide range of heterogeneous firms (Burgstahler and Dichev, 1997). Focusing, for instance, on undeflated net income bears the risk that the analysis primarily covers small and moderately profitable firms that fluctuate in the region of the two intervals surrounding the zero earnings benchmark. Larger firms would, however, be found in the upper and lower regions of the histogram. One way to address this issue and make a heterogeneous set of firms comparable is deflation.

Usually, earnings variables are deflated by lagged market capitalization. Some authors, however, suggest other metrics such as total assets, sales, or number of shares outstanding as alternative deflators. Dechow et al. (2003) show that the irregularity at the zero earnings benchmark is emphasized by scaling with market value and argue that the discontinuity may be rather attributable to deflator choice than managerial discretion. Durtschi and Easton (2005, 2009) take the same line and point out that the market capitalization is smaller for loss firms than for profit firms. Hence, loss firms have a comparably smaller denominator which drives deflated earnings of loss firms farther away and profit observations even closer to the benchmark. To address this asymmetry in market capitalization on the two sides of the threshold, they suggest undeflated EPS as a more suitable earnings measure for distributional tests. Other authors, however, doubt that the discontinuity is simply a statistical effect of scaling (e.g., Beaver et al., 2007; Jacob and Jorgensen, 2007; Kerstein and Rai, 2007). Beaver et al. (2007), in particular, posit that scaling net income by shares outstanding (which is basically the same as undeflated EPS) does not mit-
igate the scaling problem. They show that the number of shares outstanding differs significantly for firms with profits and losses of similar magnitude and argue that deflating by shares outstanding (that is, undeflated EPS) does not eliminate the problem of deflator asymmetry around the threshold.

In this study, I choose (undeflated) per share variables for two reasons:

- **Statistical Attributes of the Scaling Variable.** The most appropriate deflator increases in similar proportion to the magnitude of the respective earnings variable on both sides of the benchmark.\(^{16}\) If, as suggested by Durtschi and Easton (2005, 2009), the deflator is higher for profit than for loss observations of the same magnitude, deflation distorts the underlying distribution of net income and introduces a distributional kink unrelated to earnings management. To assess the appropriateness of potential deflators, I follow Durtschi and Easton (2009) and compare the differences between various potential deflators for profit and loss observations of the same magnitude in Figure 6.1. Specifically, I compare mean and median values of five potential deflators for profit and loss observations in five different ranges of net income (i.e., ±0.0 to 0.2 MEUR, ±0.2 to 0.6 MEUR, ±0.6 to 1.0 MEUR, ±1.0 to 1.5 MEUR, and ± > 1.5 MEUR). Statistical significance of these differences is tested using a two-tailed mean comparison (\(t\)-statistics) and Wilcoxon rank-sum test (\(z\)-statistics). Since it has the lowest \(t\)- and \(z\)-statistics among all other potential deflators, the average number of shares outstanding turns out to be the most adequate deflator in terms of symmetry.\(^{17}\)

- **Importance of the EPS Figure.** It is questionable whether firms manage metrics such as net income deflated by sales or market value. Investors, analysts, and other stakeholders focus on either undeflated net income or EPS. Although EPS is also a deflated earnings figure, its role differs from other deflated earnings numbers since it is mandatorily disclosed in annual and interim financial reports.\(^{18}\) From a US perspective, Durtschi and Easton (2009) note that managers, analysts, and shareholders focus much more on EPS than on shares outstanding during the year should not have a material effect on the distribution of EPS.

\(^{16}\)Beaver et al. (2007, p. 550), for example, plot mean market capitalization for 60 net income intervals (each 100,000 USD wide) on each side of the zero earnings benchmark to visualize its appropriateness as deflator variable.

\(^{17}\)The appropriateness of earnings deflated by number of shares for analyses at the analyst forecast and earnings changes benchmark is confirmed in identical tests (unreported).

\(^{18}\)IAS 33 *Earnings per Share* Par. 2 and 66 require firms whose ordinary shares or potential ordinary shares are traded in a public market to disclose basic and diluted EPS within the Statement of Comprehensive Income.
net income and that both figures are used with at least the same frequency in compensation contracts. They consequently suggest that manipulation efforts are rather directed towards EPS than undeflated or even deflated net income. Under German GAAP, EPS are still a voluntary disclosure. With the increasing importance of capital markets and globalization, however, many publicly listed firms in Germany started to disclose EPS on a voluntary basis even before it became mandatory with the adoption of IFRS. With the shift from a more legally oriented (i.e., aimed on dividend distribution and tax computation) to an information and investor oriented system, EPS is today an important ratio in both financial reports and the press. Moreover, it is one of the figures most frequently used by analysts in the valuation process (Küting and Weber, 2009, p. 307).

Taken together, I consider earnings variables on a per share basis to be the most preferable measures for the distributional approach. As a result, the main analyses and results in Section 6.4 are based on earnings surprises per share, EPS, and changes in EPS. However, to rule out the possibility that my results are driven by using undeflated EPS, I perform additional robustness test with other potential deflators in Section 6.4.4.3.

The remainder of this section summarizes data definitions: Earnings surprises per share are calculated as difference between actual EPS and the last mean consensus forecast before the earnings announcement date. The required data is gathered from the I/B/E/S database. EPS data for the zero earnings and the changes in earnings benchmark are gathered from Worldscope. Worldscope covers several data items for basic EPS. To maximize sample size, I use a combination of two data items, namely Earnings per Share – As Reported (WS: 18193) and Earnings per Share – Fiscal Year-End (WS: 05202). In manual checks I find that the former has the highest level of accuracy but relatively low availability. The latter is, in comparison, slightly less accurate but available for nearly all public firms in the database. To combine the best of both worlds, I use Earnings per Share – As Reported if available and Earnings per Share – Fiscal Year-End otherwise.

---

19See their footnote 13 on p. 563 for further references on the empirical prevalence of net income and EPS in compensation contracts.

20Note, however, that managing EPS usually requires discretion concerning undeflated net income for two reasons. First, repurchasing shares to decrease the denominator and increase EPS is a highly visible and not very effective earnings management device (see, e.g., Guay, 2002; Hribar et al., 2006) Second, if effective at all, repurchases are no “last-minute” earnings management device, since IAS 33 Earnings per Share Par. 10 requires the weighted average number of shares outstanding during the year to calculate EPS. Hence, repurchasing stock close to the fiscal year-end will not have much effect on the EPS figure.
108 6. The Prevalence of Benchmark Beating

<table>
<thead>
<tr>
<th>Net Income Intervals</th>
<th>Profits Mean</th>
<th>Profits Median</th>
<th>Losses Mean</th>
<th>Losses Median</th>
<th>Differences in Means and Location Mean</th>
<th>Differences in Means and Location Median</th>
<th>t-stat</th>
<th>z-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>±0.0 to 0.2</td>
<td>65.3</td>
<td>29.2</td>
<td>33.4</td>
<td>17.4</td>
<td>32.0</td>
<td>11.8</td>
<td>2.48b</td>
<td>2.43b</td>
</tr>
<tr>
<td>±0.2 to 0.6</td>
<td>71.8</td>
<td>21.7</td>
<td>92.8</td>
<td>19.8</td>
<td>–21.0</td>
<td>1.9</td>
<td>0.47</td>
<td>0.24</td>
</tr>
<tr>
<td>±0.6 to 1.0</td>
<td>57.4</td>
<td>40.7</td>
<td>49.9</td>
<td>24.8</td>
<td>7.6</td>
<td>15.8</td>
<td>0.58</td>
<td>2.21b</td>
</tr>
<tr>
<td>±1.0 to 1.5</td>
<td>60.8</td>
<td>39.0</td>
<td>60.8</td>
<td>35.0</td>
<td>0.0</td>
<td>4.0</td>
<td>0.00</td>
<td>0.35</td>
</tr>
<tr>
<td>± &gt; 1.5</td>
<td>3,288</td>
<td>348.6</td>
<td>1,197.3</td>
<td>99.7</td>
<td>2,090.7</td>
<td>248.9</td>
<td>4.35a</td>
<td>12.56a</td>
</tr>
<tr>
<td>Total Assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>±0.0 to 0.2</td>
<td>54.7</td>
<td>24.4</td>
<td>35.1</td>
<td>17.6</td>
<td>19.7</td>
<td>6.8</td>
<td>1.90c</td>
<td>1.99b</td>
</tr>
<tr>
<td>±0.2 to 0.6</td>
<td>54.7</td>
<td>30.5</td>
<td>79.9</td>
<td>24.1</td>
<td>–25.2</td>
<td>6.3</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>±0.6 to 1.0</td>
<td>50.3</td>
<td>36.5</td>
<td>34.6</td>
<td>21.2</td>
<td>15.7</td>
<td>15.2</td>
<td>1.64</td>
<td>2.32b</td>
</tr>
<tr>
<td>±1.0 to 1.5</td>
<td>57.3</td>
<td>42.4</td>
<td>56.5</td>
<td>40.1</td>
<td>0.8</td>
<td>2.3</td>
<td>0.07</td>
<td>0.44</td>
</tr>
<tr>
<td>± &gt; 1.5</td>
<td>4,155.3</td>
<td>284.7</td>
<td>1,062.4</td>
<td>101.5</td>
<td>3,092.9</td>
<td>183.2</td>
<td>4.26a</td>
<td>10.5a</td>
</tr>
<tr>
<td>Equity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>±0.0 to 0.2</td>
<td>22.3</td>
<td>11.9</td>
<td>12.2</td>
<td>4.9</td>
<td>10.1</td>
<td>7.0</td>
<td>2.30b</td>
<td>2.51b</td>
</tr>
<tr>
<td>±0.2 to 0.6</td>
<td>23.3</td>
<td>16.1</td>
<td>29.0</td>
<td>8.7</td>
<td>–5.7</td>
<td>7.4</td>
<td>0.62</td>
<td>2.46b</td>
</tr>
<tr>
<td>±0.6 to 1.0</td>
<td>23.9</td>
<td>21.6</td>
<td>17.0</td>
<td>11.2</td>
<td>6.9</td>
<td>10.3</td>
<td>1.58</td>
<td>2.80a</td>
</tr>
<tr>
<td>±1.0 to 1.5</td>
<td>28.6</td>
<td>19.7</td>
<td>21.0</td>
<td>14.9</td>
<td>7.6</td>
<td>4.8</td>
<td>1.39</td>
<td>1.97b</td>
</tr>
<tr>
<td>± &gt; 1.5</td>
<td>1,116.1</td>
<td>122.3</td>
<td>372.8</td>
<td>38.6</td>
<td>743.3</td>
<td>83.7</td>
<td>4.53a</td>
<td>11.39a</td>
</tr>
<tr>
<td>Market Capitalization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>±0.0 to 0.2</td>
<td>28.9</td>
<td>16.3</td>
<td>24.3</td>
<td>10.9</td>
<td>4.6</td>
<td>5.3</td>
<td>0.78</td>
<td>0.72</td>
</tr>
<tr>
<td>±0.2 to 0.6</td>
<td>33.4</td>
<td>18.4</td>
<td>36.2</td>
<td>11.9</td>
<td>–2.8</td>
<td>6.5</td>
<td>0.21</td>
<td>1.87c</td>
</tr>
<tr>
<td>±0.6 to 1.0</td>
<td>34.0</td>
<td>26.1</td>
<td>20.6</td>
<td>13.9</td>
<td>13.5</td>
<td>12.2</td>
<td>2.44b</td>
<td>3.06b</td>
</tr>
<tr>
<td>±1.0 to 1.5</td>
<td>46.6</td>
<td>31.1</td>
<td>34.1</td>
<td>18.9</td>
<td>12.5</td>
<td>12.2</td>
<td>1.14</td>
<td>2.24b</td>
</tr>
<tr>
<td>± &gt; 1.5</td>
<td>2,063.1</td>
<td>205.1</td>
<td>539.1</td>
<td>46.3</td>
<td>1,524</td>
<td>158.8</td>
<td>5.14a</td>
<td>13.43a</td>
</tr>
<tr>
<td>Number of Shares</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>±0.0 to 0.2</td>
<td>7.2</td>
<td>5.0</td>
<td>6.0</td>
<td>5.5</td>
<td>1.1</td>
<td>–0.5</td>
<td>0.82</td>
<td>0.39</td>
</tr>
<tr>
<td>±0.2 to 0.6</td>
<td>7.9</td>
<td>6.5</td>
<td>6.8</td>
<td>5.2</td>
<td>1.1</td>
<td>1.3</td>
<td>0.87</td>
<td>1.30</td>
</tr>
<tr>
<td>±0.6 to 1.0</td>
<td>8.7</td>
<td>6.6</td>
<td>6.0</td>
<td>5.5</td>
<td>2.6</td>
<td>1.1</td>
<td>1.97b</td>
<td>1.85c</td>
</tr>
<tr>
<td>±1.0 to 1.5</td>
<td>10.0</td>
<td>8.3</td>
<td>8.5</td>
<td>6.4</td>
<td>1.5</td>
<td>1.8</td>
<td>0.81</td>
<td>0.85</td>
</tr>
<tr>
<td>± &gt; 1.5</td>
<td>64.8</td>
<td>14.7</td>
<td>41.5</td>
<td>11.7</td>
<td>23.2</td>
<td>3.0</td>
<td>2.94a</td>
<td>3.77a</td>
</tr>
</tbody>
</table>

The table reports differences in means and medians of various potential deflators for profit and loss observations of a similar magnitude. Means and medians are calculated for the intervals ±0.0 to 0.2 MEUR, ±0.2 to 0.6 MEUR, ±0.6 to 1.0 MEUR, ±1.0 to 1.5 MEUR, and ± > 1.5 MEUR, respectively. The analysis comprises 2,100 observations in the period 2005 to 2009. Net Income is defined as net income attributable to shareholders (WS: 01751). The deflator variables include lagged sales (WS: 01001), lagged total assets (WS: 18184 + 02999), lagged total equity (WS: 03426 + 03501), lagged market value (DS: MV), and weighted average number of shares outstanding during the year (WS: 05192). Deflator observations in the top or bottom 1% of each net income interval are treated as outliers and removed from the sample. Differences in means are tested using a t-test of differences in means (t-stat). Differences in location are tested using a Wilcoxon rank-sum test (z-stat). Significance (two-tailed) at the 1%, 5%, and 10% level is indicated by a, b, and c, respectively. A similar table can be found in Durtschi and Easton (2009, pp. 1263–1264).
6.3 Sample Selection and Descriptive Statistics

6.3.1 Sample Selection

The distributional tests of earnings management in this chapter are based on two different samples. The first, denoted as reference sample, is the same as used for the analyses of the market reaction to benchmark beating in Chapter 5. Working with identical samples allows to interpret the results from distributional tests in the light of previously detected capital market incentives. The second sample (extended sample) has more lenient data requirements and is thus significantly larger. A large sample increases the power of my statistical analyses and avoids potential sample selection bias.

Table 6.2 illustrates the sample selection process for the extended sample. The initial sample covers all firms listed on the German CDAX index between 2005 and 2009 (3,390 firm-years). To avoid double counting, I first delete preferred stock observations when a firm is listed with both preferred and ordinary shares. In a next step, I delete observations with missing data on the I/B/E/S or Worldscope database. Since I/B/E/S solely covers followed firms, the highest number of observations (1,643) is lost in the analyst forecast sample. Then, I drop all potentially erroneous or late earnings announcements (i.e., less than 30 or more than 180 days after the fiscal year-end) and stale forecasts (i.e., forecasts older than 40 days) to ensure data validity. Eventually, I remove all financial firms (i.e., two-digit SIC codes between 60 and 67) and observations below the 1st or above the 99th percentile of their yearly earnings surprise, EPS, or EPS changes distributions.

The final sample contains 1,263 observations (or 37.3% of the initial sample) for the analyst forecast benchmark, 2,100 observations (or 61.9% of the initial sample) for the zero earnings benchmark, and 2,099 observations (or 61.9% of the initial sample) for the changes in earnings benchmark. Although, all three samples are considerably larger than the reference sample, their coverage is still pretty small. Calculating coverage in terms of market capitalization, however, draws a different picture: With 70.1%, 86.3, and 86.5% of total non-financial market capitalization, all three samples cover a considerably large fraction of the German market.

21The sample selection process for the reference sample is described in Section 5.3.1.
22To exclude firms that report German GAAP, I only consider firms that prepare consolidated financial statements. Due to mandatory adoption of IFRS for consolidated financial statements in 2005, my sample is limited to firms reporting either IFRS or US GAAP. See Section 5.3.1 for details.
### Sample Selection Procedure

<table>
<thead>
<tr>
<th>Analyst Forecast Benchmark</th>
<th>Zero Earnings Benchmark</th>
<th>Earnings Changes Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total CDAX Observations</td>
<td>3,390</td>
<td>3,390</td>
</tr>
<tr>
<td>less: Preferred Stock:</td>
<td>140</td>
<td>140</td>
</tr>
<tr>
<td>less: Missing Data:</td>
<td>1,643</td>
<td>705</td>
</tr>
<tr>
<td>less: Erroneous and/or Late Earnings Releases:</td>
<td>23</td>
<td>–</td>
</tr>
<tr>
<td>less: Stale Forecasts:</td>
<td>74</td>
<td>–</td>
</tr>
<tr>
<td>less: Financial Services Firms:</td>
<td>217</td>
<td>395</td>
</tr>
<tr>
<td>less: Outliers:</td>
<td>30</td>
<td>50</td>
</tr>
</tbody>
</table>

Extended Sample (in % of CDAX Observations) 1,263 (37.3) 2,100 (61.9) 2,099 (61.9)

Reference Sample (in % of CDAX Observations) 927 (27.3) 927 (27.3) 927 (27.3)

Market Capitalization of Non-financial firms 4,489 4,489 4,489

Market Capitalization of the Reference Sample (Coverage in %) 2,766 (61.6) 2,766 (61.6) 2,766 (61.6)

Market Capitalization of the Extended Sample (Coverage in %) 3,149 (70.1) 3,873 (86.3) 3,884 (86.5)

Preferred stock observations are deleted when a firm is listed more than once on the index. Observations are classified as missing if the relevant data fields are not available on I/B/E/S, Worldscope, or Datastream. Financial firms are defined as having two-digit SIC codes between 60 and 67. Earnings announcements are considered erroneous or late when the date of the earnings release is less than 30 or more than 180 after the fiscal year-end. Forecasts are considered stale when the latest consensus forecast was calculated more than 40 days before the respective earnings announcement. Outliers are defined as observations below the 1st or above the 99th percentile of their yearly distributions. Year-end market capitalization is retrieved from Datastream (DS: MV). Market capitalization data for all German non-financial firms is gathered from summary statistics of the Deutsche Bundesbank (Deutsche Bundesbank, 2010, p. 45). Market capitalizations are reported in billion EUR.
6. The Prevalence of Benchmark Beating

6.3.2 Descriptive Statistics

Table 6.3 presents summary statistics for the earnings variables. Every panel includes summary statistics for the extended sample and condensed summary statistics for the reference sample. More detailed summary statistics for the reference sample can be found in Section 5.3.2.

Panel A of Figure 6.3 provides summary statistics for the distribution of earnings surprises per share. Confirming prior evidence in Capstaff et al. (1998, 2001), German analyst forecasts are on average optimistic (i.e., the latest consensus forecast is on average above actual reported EPS). Median earnings surprises, in contrast, equal zero in both the extended and the reference sample and do not support analyst optimism. While earnings surprises are right-skewed in the reference sample (skew: 5.89), the extended sample exhibits a slight skew to the left (skew: −1.66). Though skewed in opposite directions, both distributions have a strong peak near zero (extended sample kurtosis: 16.5, reference sample kurtosis: 127.4). Panel B presents the distributions of EPS. Mean and median EPS are positive in the extended and the reference sample. Moreover, mean and median EPS in the extended sample are considerably lower than in the reference sample. This difference seems reasonable when taking into account that the I/B/E/S database covers more profitable firms (Durtschi and Easton, 2005). Both distributions are right-skewed (extended sample skew: 3.25, reference sample skew: 8.83) and peak considerably strong at zero EPS (extended sample kurtosis: 24.9, reference sample kurtosis: 111.5). Panel C summarizes the distributions of changes in EPS. Mean earnings changes are positive in the extended and negative in the reference sample. Though both are above zero, median earnings changes in the extended sample exceed those in the reference sample. The earnings changes distribution of the extended sample has a symmetrical shape (skew: 0.31), while the reference sample is skewed to the right (skew: 8.25). Both distributions exhibit a distinct peak in the immediate vicinity of the threshold at zero (extended sample kurtosis = 20.0, reference sample kurtosis = 37.5).

---

23Mean earnings surprise in the extended sample is significantly different from zero at the 1% level. Mean earnings surprise in the reference sample is not significantly different from zero at generally accepted levels.

24Mean EPS are significantly different from zero at the 1% level in both the extended and the reference sample.

25Mean changes in EPS are not significantly different from zero on any generally accepted levels in both samples.
6. The Prevalence of Benchmark Beating

Table 6.3

Variable Distributions

Panel A: Earnings Surprises per Share

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Mean</th>
<th>σ</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>224</td>
<td>-0.027</td>
<td>0.431</td>
<td>-0.070</td>
<td>0.000</td>
<td>0.090</td>
</tr>
<tr>
<td>2006</td>
<td>247</td>
<td>0.076</td>
<td>0.371</td>
<td>-0.040</td>
<td>0.020</td>
<td>0.140</td>
</tr>
<tr>
<td>2007</td>
<td>271</td>
<td>0.017</td>
<td>0.490</td>
<td>-0.090</td>
<td>0.010</td>
<td>0.110</td>
</tr>
<tr>
<td>2008</td>
<td>271</td>
<td>-0.190</td>
<td>0.682</td>
<td>-0.250</td>
<td>-0.020</td>
<td>0.050</td>
</tr>
<tr>
<td>2009</td>
<td>250</td>
<td>-0.087</td>
<td>0.576</td>
<td>-0.170</td>
<td>-0.010</td>
<td>0.080</td>
</tr>
<tr>
<td>All</td>
<td>1,263</td>
<td>-0.044</td>
<td>0.533</td>
<td>-0.110</td>
<td>0.000</td>
<td>0.090</td>
</tr>
</tbody>
</table>

Panel B: EPS

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Mean</th>
<th>σ</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>412</td>
<td>0.856</td>
<td>2.640</td>
<td>-0.025</td>
<td>0.285</td>
<td>1.300</td>
</tr>
<tr>
<td>2006</td>
<td>423</td>
<td>1.288</td>
<td>3.929</td>
<td>0.030</td>
<td>0.460</td>
<td>1.630</td>
</tr>
<tr>
<td>2007</td>
<td>437</td>
<td>1.390</td>
<td>3.717</td>
<td>0.010</td>
<td>0.460</td>
<td>1.770</td>
</tr>
<tr>
<td>2008</td>
<td>430</td>
<td>0.861</td>
<td>3.783</td>
<td>-0.208</td>
<td>0.365</td>
<td>1.470</td>
</tr>
<tr>
<td>2009</td>
<td>398</td>
<td>0.541</td>
<td>3.174</td>
<td>-0.332</td>
<td>0.180</td>
<td>1.000</td>
</tr>
<tr>
<td>All</td>
<td>2,100</td>
<td>0.995</td>
<td>3.503</td>
<td>-0.091</td>
<td>0.360</td>
<td>1.451</td>
</tr>
</tbody>
</table>

Panel C: Changes in EPS

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Mean</th>
<th>σ</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>412</td>
<td>0.472</td>
<td>2.604</td>
<td>-0.116</td>
<td>0.150</td>
<td>0.605</td>
</tr>
<tr>
<td>2006</td>
<td>423</td>
<td>0.455</td>
<td>2.766</td>
<td>-0.068</td>
<td>0.180</td>
<td>0.627</td>
</tr>
<tr>
<td>2007</td>
<td>436</td>
<td>0.312</td>
<td>2.340</td>
<td>-0.252</td>
<td>0.105</td>
<td>0.575</td>
</tr>
<tr>
<td>2008</td>
<td>430</td>
<td>-0.530</td>
<td>2.838</td>
<td>-0.770</td>
<td>-0.041</td>
<td>0.260</td>
</tr>
<tr>
<td>2009</td>
<td>398</td>
<td>-0.629</td>
<td>2.672</td>
<td>-0.830</td>
<td>-0.095</td>
<td>0.290</td>
</tr>
<tr>
<td>All</td>
<td>2,099</td>
<td>0.021</td>
<td>2.690</td>
<td>-0.370</td>
<td>0.080</td>
<td>0.481</td>
</tr>
</tbody>
</table>

The table presents summary statistics for the key variables underlying the distributional tests of benchmark-driven earnings management. Earnings surprises are calculated as actual EPS minus the last mean consensus forecast issued before the earnings announcement date. Changes in EPS are EPS in t minus EPS in t - 1. Earnings and forecast data for Panel A and the reference sample in Panels B and C are retrieved from I/B/E/S. Earnings data for the extended sample in Panels B and C are gathered from Worldscope.
6.4 Empirical Results

This section presents empirical evidence from distributional tests of benchmark-driven earnings management in Germany. Main results for the analyst forecast, zero earnings, and earnings changes benchmark are summarized in Sections 6.4.1, 6.4.2, and 6.4.3, respectively. Section 6.4.4 provides additional analyses to ensure the robustness of the results.

6.4.1 The Analyst Forecast Benchmark

Figure 6.5 shows the frequency distribution of earnings surprises in Panel A and the corresponding test statistics of distributional discontinuity in Panels B and C. In each panel, black graphs illustrate the results for the extended sample and gray graphs are based on the reference sample. For the histograms in Panel A, interval width is calculated from the sample data using Silverman’s rule of thumb (Rule V in Table A.4.1 of Appendix A.4).\(^\text{26}\) Optimal interval width is 0.032 for the extended sample and 0.036 for the reference sample. To focus on the distributions’ properties around the benchmark, the depicted histograms are limited to 20 intervals above and below zero. The histograms cover earnings surprises in the range of \([-0.64,0.64]\) \([-0.72,0.72]\) with 1,124 (827) observations or 89\% (89\%) of the extended (reference) sample.

Visual inspection of the histograms in Panel A reveals a considerable jump at the threshold of zero earnings surprises. With 115 observations in the first interval below zero and 219 observations in the directly adjacent interval, the extended sample contains 1.90 times more firm-years with zero or small positive than negative earnings surprises. A similar pattern can be found for the reference sample, where the ratio of zero or small positive earnings surprises (165) to negative earnings surprises (88) is 1.88. The jump from interval \(-1\) to \(+1\) generally supports the importance of analyst forecasts as earnings targets. Panels B and C of Figure 6.5 plot BD and BP test statistics of distributional discontinuity for every of the 40 histogram intervals.\(^\text{27}\) Positive (negative) test statistics indicate an overrepresentation (underrepresentation) of observations in the respective bins. To ease the interpretation of the results, Panels B and C each include four horizontal lines representing significance at the 1\% (dashed) and 5\% level (dotted). Test statistics

\(^{26}\)Refer to Appendix A.4 for details on the derivation and calculation of different rules of thumb for optimal bin- and bandwidths.

\(^{27}\)Derivation and calculation of the BD and BP test statistics is discussed in detail in Sections 6.1.2.1 and 6.1.2.2, respectively.
above these lines indicate distributional discontinuities on the respective significance levels. The
test statistics plotted in Figure 6.5 and summarized in Table 6.4 generally confirm visual inspec-
tion: For interval $+1$, the BP test statistic is clearly above the 1% level of significance for both
the extended ($\tau_{+1}^{BD} = 7.85$, $\tau_{+1}^{BP} = 8.27$) and the reference sample ($\tau_{+1}^{BD} = 5.93$, $\tau_{+1}^{BP} = 7.01$). For
the first interval below the benchmark, the BD test statistic is significant at generally accepted
levels in both the extended ($\tau_{-1}^{BD} = -2.79$) and the reference sample ($\tau_{-1}^{BP} = -2.91$). Contrasting
with the earnings management explanation, the more rigid BP test statistic for interval $-1$ does
not indicate a significant underrepresentation of observation in any of the two samples. Given the
lack of significance for interval $-1$, it is surprising that the BP test statistic detects a significant
underrepresentation of observations in interval $+2$ of the extended sample ($\tau_{+2}^{BP} = -3.03$).28 I
provide two potential reasons for the small number of observations in interval $+2$: 1) Managers
create of “cookie jar” reserves when earnings are significantly above target (see, e.g., Degeorge
et al., 1999) or, 2) Managers guide both optimistic and pessimistic analysts, inducing a under-
representation of observations in interval $-1$ and $+2$. Both explain why earnings surprises are
shifted towards zero from both sides of the benchmark and cause pile-up at interval $+1$.

**TABLE 6.4**

| Bin | Extended Sample | | Reference Sample | |
|-----|-----------------|-----------------|
|     | N               | $\tau^{BD}$    | $\tau^{BP}$    | N               | $\tau^{BD}$    | $\tau^{BP}$    |
| $-3$| 53              | -0.61           | -1.43*          | 31              | -3.09***        | -2.92***        |
| $-2$| 88              | 0.35            | -1.04           | 82              | 2.13**          | 0.74            |
| $-1$| 115             | -2.79***        | -1.18           | 88              | -2.91***        | -0.65           |
| $+1$| 219             | 7.85***         | 8.27***         | 165             | 5.93***         | 7.01***         |
| $+2$| 73              | -5.90***        | -3.03***        | 75              | -2.33***        | -0.41           |
| $+3$| 67              | 0.81            | -0.62           | 37              | -2.48**         | -2.76***        |

The table summarizes Burgstahler and Dichev (1997) ($\tau^{BD}$) and Bollen and Pool (2009) ($\tau^{BP}$) test statistics
for irregularities in the distribution of earnings surprises per share. Total number of observations is 927 for the
reference sample and 1,263 for the extended sample. Kernel bandwidth is calculated using Silverman’s
(1986) rule of thumb (Rule V) and equals binwidth (extended sample: 0.032; reference sample: 0.036).
Significance (one-tailed) at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***)
asterisks, respectively.

Figure 6.6 plots the frequency histogram of earnings surprises for a binwidth of 0.032 and
the related kernel density estimate generated with a Gaussian kernel function (solid line). The

28A considerable underrepresentation of earnings surprises in interval $+2$ is also evident in other European studies
(see, e.g., Glaum et al., 2004; Daske et al., 2006). The authors, however, solely focus on the immediate vicinity of
zero earnings surprises.
FIG. 6.5.—Distribution of Earnings Surprises per Share. Panel A displays the frequency histograms of earnings surprises (black: extended sample with 1,263 observations; gray: reference sample with 927 observations) for the period 2005 to 2009. Panels B and C depict the Burgstahler and Dichev (1997) (BD) and Bollen and Pool (2009) (BP) test statistics of distributional discontinuities, respectively. Kernel bandwidths are calculated using Silverman’s (1986) rule of thumb (Rule V). The binwidths (extended sample: 0.032; reference sample: 0.036) of the histograms in Panel A correspond to the bandwidths of the kernel estimators.
dashed lines above and below the density estimate represent the 95% confidence intervals for the estimated density function. The results reinforce the distributional irregularity at interval $+1$, since the histogram bin lies clearly outside the confidence interval of the density estimate. For interval $-1$, the histogram bin is below the reference distribution, but not outside the 95% confidence interval, indicating that the underrepresentation in that interval is not significant at the 5% level. Integration below the estimated density function depicted in Figure 6.6 allows to assess the prevalence of benchmark beating: The probability that an observation falls into interval $+1$, for example, is 0.103. Hence, the expected number of observations for this interval is given as $0.103 \times 1,263 = 130$. With 219 actual observations in interval $+1$, there are 89 observations more than expected under the null of no discontinuities. For interval $-1$, a probability of 0.101 yields an expected number of 128 observations ($0.101 \times 1,263$). Given the actual 115 observations in interval $-1$, there are 13 observations less in that bin than expected. The results suggest that maximal 10% of the observations with small negative earnings surprises engage in income increasing earnings management to achieve the latest analyst consensus. In comparison, a maximum of 28% of the observations in the range of interval $+2$ is shifted towards zero from the other side of the benchmark. A minimum of 48 excess observations in interval $+1$ originate from other regions of the earnings surprise distribution.

Overall, the distribution of earnings surprises exhibits an extreme pile-up of observations just above the benchmark at zero. In contrast to the benchmark beating explanation, however, observations seem to shift to the small earnings surprise interval from both sides of the threshold. Specifically, a maximum of 10% of the observations in interval $-1$ and a maximum of 28% of the observations in interval $+2$ shift to interval $+1$. I expect this pattern to result from guiding optimistic and pessimistic analysts to the small positive earnings surprise region or building up reserves when the benchmark is exceeded by a large amount.

29Confidence intervals are calculated as $\hat{f}(x) \pm z_{1-\alpha/2} \left[ \hat{\sigma}^2_{\hat{f}(x)} \right]^{1/2}$, where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ percentile of the standard normal distribution and $\hat{\sigma}^2_{\hat{f}(x)}$ an estimate of the variance of $\hat{f}(x)$ at point $x$ (Jann, 2007).

30This interpretation grounds on the assumption that observations do not shift from intervals below interval $-1$ to interval $-1$. The assumption is justified because managing earnings upward into the small miss region lacks both theoretical reasoning and economic rationale.
6. The Prevalence of Benchmark Beating

6.4.2 The Zero Earnings Benchmark

Figure 6.7, Panel A, shows the distribution of reported EPS for the extended sample (in black) and the reference sample (in gray). Panels B and C plot the related BD and BP test statistics for distributional discontinuities, respectively. To focus on the benchmark at the center of the distribution, the depicted histograms are limited to 20 intervals on each side of zero EPS. Interval width is calculated using Silverman’s rule of thumb for optimal kernel bandwidths (Rule V in Table A.4.1 of Appendix A.4).\(^\text{31}\) Based on an interval width of 0.224 (0.319), the histograms cover a region of \([-4.40,4.40]\) \([-6.40,6.40]\) with 1,922 (878) observations or 92\% (90\%) of the extended (reference) sample.

The histograms in Panel A exhibit a significant jump in the immediate vicinity of zero EPS. With 136 observations in interval \(-1\) and 330 in interval \(+1\), there are 2.43 times more observa-

\(^{31}\)Refer to Appendix A.4 for details on bin- and bandwidth heuristics.
The Prevalence of Benchmark Beating

In the first interval above than in the first interval below zero EPS. For the reference sample, the ratio of small profits to small losses is even larger with 3.75. Overall, visual inspection reveals a clear discontinuity at zero EPS for both the extended and the reference sample and supports the importance of zero EPS as earnings benchmark. The BD and BP test statistics plotted in Panels B and C of Figure 6.7 and summarized in Table 6.5 support the visual impression of benchmark importance. \( \tau_{BD}^{+1} \) and \( \tau_{BP}^{+1} \) of 7.89 and 7.65, respectively, confirm that there are significantly more observations in interval +1 than expected under the null of a smooth distribution. The same result holds for the reference sample with both test statistics clearly above generally accepted significance levels (\( \tau_{BD}^{-1} = 5.89 \), \( \tau_{BP}^{-1} = 5.65 \)). For the earnings management story to be complete, the documented pile-up in interval +1 should be caused by pushing small losses into the small profit region. The test statistics for interval −1 support this notion: For the extended sample, \( \tau_{BD}^{-1} = -4.73 \) and \( \tau_{BP}^{-1} = -3.06 \) suggest a significant shortfall of observations in the first interval below zero EPS. Even more significant are the results for the reference sample (\( \tau_{BD}^{-1} = -6.07 \), \( \tau_{BP}^{-1} = -3.98 \)). The remaining regions of the histogram are considerably smooth with two exceptions: For the extended sample, \( \tau_{BP}^{-2} \) of −2.50 indicates significantly less observations in interval −2 than expected.32 Two potential reasons explain a shortfall in interval −2: Either managers build up accounting reserves when firms’ premanaged EPS fall considerably short of zero EPS in the current year (see, e.g., Degeorge et al., 1999), or a considerable number of firms manage EPS from the loss region of interval −2 into the small profit region. Furthermore, a \( \tau_{BD}^{+2} \) of −2.04 for the extended sample suggest that managers manipulate EPS downward when premanaged EPS exceed the benchmark by a larger amount.33 However, none of these findings contradict the earnings management hypothesis at the zero earnings benchmark.

Figure 6.8 shows the histogram of the extended sample overlaid by a reference distribution derived from kernel density estimation (solid line). The two dashed lines represent the 95% confidence intervals of the density estimate.34 With histogram bin +1 clearly above the reference distribution and bin −1 considerably below, the plot underpins previous evidence of a shift from the small loss to the small profit region. Integration below the reference distribution in Figure 6.8 allows to assess the prevalence of earnings management to avoid losses. With a probability that an observation falls in interval +1 of 0.106 and 2,100 observations in total, the expected

---

32This effect is, however, only weakly supported by the BD test statistic (\( \tau_{BD}^{-2} = -1.60 \)) nor the results of the reference sample.

33This result is, however, not supported by the BP test statistic (\( \tau_{BP}^{-2} = -0.15 \)) and the reference sample.

34Confidence intervals are calculated as \( \hat{f}(x) \pm z_{1-\alpha/2} \frac{\hat{\sigma}_2}{\hat{f}(x)} \), where \( z_{1-\alpha/2} \) is the 1 − \( \alpha / 2 \) percentile of the standard normal distribution and \( \hat{\sigma}_2 \) an estimate of the variance of \( \hat{f}(x) \) at point \( x \) (Jann, 2007).
A

B

C

Fig. 6.7.—Distribution of EPS. Panel A displays the frequency histograms of EPS (black: extended sample with 2,100 observations; gray: reference sample with 927 observations) for the period 2005 to 2009. Panels B and C depict the Burgstahler and Dichev (1997) (BD) and Bollen and Pool (2009) (BP) test statistics of distributional discontinuities, respectively. Kernel bandwidths are calculated using Silverman’s (1986) rule of thumb (Rule V). The binwidths (extended sample: 0.224; reference sample: 0.319) of the histograms in Panel A correspond to the bandwidths of the kernel estimators.
6. The Prevalence of Benchmark Beating

### Table 6.5

**Discontinuities in the Distribution of EPS**

<table>
<thead>
<tr>
<th>Bin</th>
<th>Extended Sample</th>
<th>Reference Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>(\tau^{BD})</td>
</tr>
<tr>
<td>−3</td>
<td>78</td>
<td>0.95</td>
</tr>
<tr>
<td>−2</td>
<td>88</td>
<td>−1.60*</td>
</tr>
<tr>
<td>−1</td>
<td>136</td>
<td>−4.73***</td>
</tr>
<tr>
<td>+1</td>
<td>330</td>
<td>7.89***</td>
</tr>
<tr>
<td>+2</td>
<td>207</td>
<td>−2.04**</td>
</tr>
<tr>
<td>+3</td>
<td>158</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The table summarizes Burgstahler and Dichev (1997) (\(\tau^{BD}\)) and Bollen and Pool (2009) (\(\tau^{BP}\)) test statistics for irregularities in the distribution of reported EPS. Total number of observations is 927 for the reference sample and 2,100 for the extended sample. Kernel bandwidth is calculated using Silverman’s (1986) rule of thumb (Rule V) and equals binwidth (extended sample: 0.224; reference sample: 0.319). Significance (one-tailed) at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (*** ) asterisks, respectively.

The number of observations is 223. Taking into account the actual number of 330 yields 107 excess observations in interval +1. The probability to reside in interval −1 is significantly lower at 0.083. The expected number of observations in interval −1 (174) suggests 38 observations less than expected under no discontinuity. Hence, a maximum of 22% of the small loss observations in interval −1 has shifted to the small profit interval +1. The remaining excess observations in interval +1 originate from other regions of the EPS distribution.

Overall, visual inspection and statistical tests reveal a considerable shift of observations from the small loss to the small profit region. Specifically, a maximum of 22% of the expected 174 small loss firms in interval −1 is suspected to engage in dicretionary activities to avoid reporting a loss.

### 6.4.3 The Earnings Changes Benchmark

Figure 6.9 provides the distribution of EPS changes in Panel A and related test statistics for discontinuities in Panels B and C. To allow comparison between the extended and the reference sample, the particular results are plotted in black and gray, respectively. The histogram and

---

35This interpretation grounds on the assumption that observations do not shift from intervals below interval −1 to interval −1. The assumption is justified because managing earnings upward into the small miss region lacks both theoretical reasoning and economic rationale.
kernel estimates are based on an interval width of 0.124 for the extended and 0.139 for the reference sample. In both samples, interval width is chosen according to Silverman’s (1986) optimal bandwidth formula for kernel density estimation (Rule V of Table A.4.1 in Appendix A.4). To focus on the threshold of last year’s earnings, the histogram is limited to 20 intervals on each side of zero. This region \([-2.48, 2.48)\) \([-2.78, 2.78]\) covers 1,845 (846) observations, which accounts for approximately 88% (91%) of the extended (reference) sample.

Both histograms in Panel A of Figure 6.9 exhibit a strong peak at the benchmark of zero EPS changes. With a value of 1.73 for the extended and 1.91 for the reference sample, the ratio of observations in interval +1 to observations in interval \(-1\) confirms a considerable pile-up at the threshold. Though less obvious than in the EPS distribution, visual inspection provides evidence of a jump in the distribution of EPS changes and suggests managers’ reluctance to fall short of last year’s earnings. The test statistics shown in Panels B and C of Figure 6.9 and stated in Table

---

36Refer to Appendix A.4 for details on bin- and bandwidth heuristics.
6. The Prevalence of Benchmark Beating

6.6 confirm the first impression from visual inspection. In both samples, BD and BP test statistics for interval +1 are positive and clearly above generally accepted levels of significance (extended sample: $\tau_{+1}^{BD} = 4.77$, $\tau_{+1}^{BP} = 5.11$; reference sample: $\tau_{+1}^{BD} = 3.17$, $\tau_{+1}^{BP} = 3.82$). Moreover, both test statistics for small loss observations in interval –1 are negative and indicate a shortfall of observations. However, only the BD test statistics suggest a significant underrepresentation of observations (extended sample: $\tau_{-1}^{BD} = -1.82$; reference sample: $\tau_{-1}^{BD} = -2.55$). BP test statistics are insignificant in both samples. This lack of significance for the interval immediately below the benchmark contradicts the earnings management explanation. In the wider neighborhood of the threshold, two additional properties of distribution attract attention: For the extended sample, both test statistics indicate less observations in interval –2 than expected under the null of distributional smoothness; a finding that could be either related to income decreasing earnings management when premanaged earnings fall short of last year’s EPS by a large amount or income increasing earnings management from interval –2 to interval +1. Though unexpected, this finding does not generally conflict with the earnings management hypothesis. More puzzling is the significant positive BP test statistic for interval +2 of the reference sample, indicating that executives manage earnings into regions considerably above the threshold and potentially weakening the importance of “just achieving” the benchmark.

### Table 6.6

<table>
<thead>
<tr>
<th>Bin</th>
<th>N</th>
<th>$\tau^{BD}$</th>
<th>$\tau^{BP}$</th>
<th>N</th>
<th>$\tau^{BD}$</th>
<th>$\tau^{BP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>–3</td>
<td>88</td>
<td>0.93</td>
<td>0.18</td>
<td>35</td>
<td>–0.48</td>
<td>–1.03</td>
</tr>
<tr>
<td>–2</td>
<td>102</td>
<td>–1.72**</td>
<td>–2.07**</td>
<td>61</td>
<td>0.97</td>
<td>0.06</td>
</tr>
<tr>
<td>–1</td>
<td>160</td>
<td>–1.82**</td>
<td>–1.07</td>
<td>69</td>
<td>–2.55**</td>
<td>–1.57*</td>
</tr>
<tr>
<td>+1</td>
<td>276</td>
<td>4.77***</td>
<td>5.11***</td>
<td>132</td>
<td>3.17***</td>
<td>3.82***</td>
</tr>
<tr>
<td>+2</td>
<td>210</td>
<td>0.06</td>
<td>1.16</td>
<td>111</td>
<td>1.03</td>
<td>2.23**</td>
</tr>
<tr>
<td>+3</td>
<td>142</td>
<td>–0.20</td>
<td>–0.42</td>
<td>64</td>
<td>–0.85</td>
<td>–0.74</td>
</tr>
</tbody>
</table>

The table summarizes Burgstahler and Dichev (1997) ($\tau^{BD}$) and Bollen and Pool (2009) ($\tau^{BP}$) test statistics for irregularities in the distribution of changes in EPS. Total number of observations is 927 for the reference sample and 2,100 for the extended sample. Kernel bandwidth is calculated using Silverman’s (1986) rule of thumb (Rule V) and equals binwidth (extended sample: 0.124; reference sample: 0.139). Significance (one-tailed) at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***).

Figure 6.10 compares the actual distribution of changes in EPS with a kernel density estimate based on a Gaussian kernel (depicted as the solid line). The two dashed lines indicate the related
Fig. 6.9.—Distribution of EPS Changes. Panel A displays the frequency histograms of changes in EPS (black: extended sample with 2,099 observations; gray: reference sample with 927 observations) for the period 2005 to 2009. Panels B and C depict the Burgstahler and Dichev (1997) (BD) and Bollen and Pool (2009) (BP) test statistics of distributional discontinuities, respectively. Kernel bandwidths are calculated using Silverman’s (1986) rule of thumb (Rule V). The binwidths (extended sample: 0.124; reference sample: 0.139) of the histograms in Panel A correspond to the bandwidths of the kernel estimators.
Confidence intervals at the 95% level.\textsuperscript{37} Confirming previous results, there is a considerable number of excess observations in interval $+1$ and a comparably moderate shortfall of observations in interval $-1$. The over- and underrepresentation of observations in the intervals adjacent to the benchmark can be quantified as the difference between the actual and the expected number of observations based on the kernel density estimate.\textsuperscript{38} With a probability that observations reside in bin $+1$ of 0.098, the first interval above the benchmark contains 70 observations or 34\% more than expected under the null of a smooth distribution. In comparison, the underrepresentation for the first bin below zero is clearly lower: With an actual number of 174 observations, interval $-1$ has only 14 observations or 8\% less than less than expected given a smooth reference distribution.

Although there is a jump in the distribution at the benchmark, visual inspection and test statistics suggests that meeting or beating the last year’s earnings threshold is less important for managers. This result is reinforced by the finding that at maximum 8\% of all “small miss” firms are suspected to manipulate EPS to avoid earnings declines.\textsuperscript{39}

### 6.4.4 Robustness Tests

Although the distributional approach is widespread in accounting research, several authors question its accuracy in measuring earnings management (see Sections 3.2.1.3 and 6.1.3 for details). This section provides a comprehensive set of robustness checks to address some known shortcomings of the approach. In Section 6.4.4.1, I examine the sensitivity of my results with regard to different interval widths. Section 6.4.4.2 tests whether the results hold for different kernel functions. An analysis of alternative deflators is provided in Section 6.4.4.3. Eventually, I examine the influence of differences in effective tax rates for profit and loss firms in Section 6.4.4.4.

\textsuperscript{37}Confidence intervals are calculated as $\hat{f}(x) \pm z_{1-\alpha/2}\sqrt{\hat{\sigma}^2 f(x)}$, where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ percentile of the standard normal distribution and $\hat{\sigma}^2 f(x)$ an estimate of the variance of $\hat{f}(x)$ at point $x$ (Jann, 2007).

\textsuperscript{38}In accordance with the calculations for the other benchmarks, the expected number of observations is derived by multiplying the sample size with the probability that an observation falls into the respective interval under the null of a smooth distribution. This probability equals the area under the estimated reference distribution for the respective interval and can be easily calculated by means of integration.

\textsuperscript{39}This interpretation grounds on the assumption that observations do not shift from intervals below interval $-1$ to interval $-1$. The assumption is justified because managing earnings upward into the small miss region lacks both theoretical reasoning and economic rationale.
6. The Prevalence of Benchmark Beating

---

**FIG. 6.10.**—**EPS Changes Histogram and Reference Distribution.** The figure displays the density histogram of 2,099 EPS changes observations (extended sample) between 2005 and 2009 and a Gaussian kernel estimate used as reference distribution (solid line). The dashed lines represent the 95% confidence intervals of the density estimate. Kernel bandwidth (0.124) is calculated using Silverman’s (1986) rule of thumb (Rule V). The binwidth of the histogram corresponds to the bandwidth of the kernel estimator.

6.4.4.1 Bin- and Bandwidth Choice

The choice of bin- and bandwidth potentially affects the results of the distributional approach.\(^{40}\) Several authors address this issue by testing the robustness of their results with different bin- and bandwidths (see, e.g., Holland and Ramsay, 2003; Glaum et al., 2004). I follow this approach and calculate BD and BP test statistics for the two intervals directly adjacent to the respective threshold and five alternative bin- and bandwidths. These also include the optimal bin- and bandwidths calculated according to Rule II, IV, and V of Table A.4.1 in Appendix A.4.\(^{41}\) The

---

\(^{40}\)Refer to Section 6.1.1 and 6.1.2.2 for a description of the role of bin- and bandwidth choice in density estimation.

\(^{41}\)I exclude Rule I and III because they build on the sample’s standard deviation as a measure of spread. Silverman (1986, p. 46) points out that using the standard deviation potentially oversmoothes the density estimate in the case of bimodality or skewness. Based on the interquartile range as a more robust measure of spread, Rules II, IV, and V are more appropriate in my empirical setting. Important to note, Rule V is based on the minimum of standard deviation \(\hat{\sigma}\) and interquartile range \(\frac{IQR}{1.34}\). However, since \(\hat{\sigma} > \frac{IQR}{1.34}\) for all samples, Rule V is de facto based on \(IQR\).
results are summarized in Table 6.7.\textsuperscript{42}

*Panel A* generally confirms the findings of the main analysis. The BD test statistics remain highly significant except for the drop of $\tau_{-1}^{BD}$ to the 10% level for the largest interval width ($-1.37$). The BP test statistics for interval $+1$ remain significant at the 1% level for all alternative binwidths. Though not significant in the main analysis (Rule V), $\tau_{-1}^{BP}$ gains significance at at least 5% for three alternative interval widths; a finding in further support of the earnings management hypothesis. The results in *Panel B* reveal that BD and BP test statistics are highly significant at the 1% level for all alternative interval widths and thus confirm previous findings. *Panel C* shows that BD and BP test statistics for interval $+1$ remain stable at the 1% significance level for all interval widths. The BP statistic for interval $-1$ gains significance at at least 10% when interval width is increased above Rule V and remains insignificant for smaller interval widths. Similarly, the BD test statistic becomes significant at the 1% level when intervals are wider than recommended by Rule V. Overall, the results in all three panels dispel doubts that my findings are attributable to arbitrary bin- and bandwidth choice.

### 6.4.4.2 Alternative Kernel Functions

Compared to kernel bandwidth, kernel choice is less sensitive in terms of minimizing mean integrated squared error of the kernel density estimate (Silverman, 1986, p. 43).\textsuperscript{43} Related studies in the field of accounting and finance use common weighting functions such as the Gaussian (see, e.g., Bollen and Pool, 2009) or Epanechnikov kernel (Christodoulou and McLeay, 2009). Though kernel choice is theoretically arbitrary, I find some considerable differences in BP test statistics when using different kernel functions. Table 6.8 summarizes BP test statistics for the first interval below ($\tau_{-1}^{BP}$) and above ($\tau_{+1}^{BP}$) the analyst forecast (*Panel A*), zero earnings (*Panel B*), and earnings changes benchmark (*Panel C*).\textsuperscript{44}

The results in Table 6.8 do not conflict with previously reported evidence in Sections 6.4.1

\textsuperscript{42}For the interpretation of the test, recall that all results in the main analyses of this chapter are based on Rule V.

\textsuperscript{43}A comparison of kernel functions in terms of asymptotic efficiency can be found in, e.g., Silverman (1986, p. 43) and Scott (1992, p. 140).

\textsuperscript{44}Note that changing the underlying kernel function also requires to adjust optimal bandwidths (Scott, 1992, pp. 141–143). Hence, I rescale reported BP test statistics using the factors depicted in Scott (1992, Table 6.3, p. 142) to allow for equivalent smoothing across several different kernel functions. Since kernel bandwidth equals binwidth under the BP approach, rescaling optimal bandwidths also affects the binwidth of the underlying histogram. Unreported results, however, suggest categorically similar results when bandwidths are not rescaled.
The table summarizes BD and BP test statistics for the two intervals adjacent to the analyst forecast (Panel A), zero earnings (Panel B), and earnings changes benchmark (Panel C) and five different bin- and bandwidths. Optimal bin- and bandwidths (Rule II, IV, and V) are calculated as described in Appendix A.4. Significance at the 10%, 5%, and 1% level is indicated by one (∗), two (∗∗), and three (∗∗∗) asterisks, respectively.

To 6.4.3. The test statistics in Panel A suggest a stronger discontinuity at interval −1 when others than the Gaussian kernel are used: While $\tau_{-1}^{BP}$ is not significant under a Gaussian kernel, it exceeds the 5% level of significance for all non-normal kernel functions. This result supports the hypothesis of earnings management to meet analyst forecasts. Things are even clearer in Panel B: The test statistics remain significant at the 1% level for all kernel functions and thus strongly support the notion of a discontinuity at zero EPS. In Panel C, the highly significant test statistic $\tau_{+1}^{BP}$ decreases considerably when the Gaussian kernel is replaced by other weighting functions. Except for the Parzen kernel (significant in the 10% level), however, all other test statistics are significant at at least 5%. $\tau_{+1}^{BP}$ remains insignificant for all other kernels except the

### Table 6.7

**Bin- and Bandwidth Analysis**

#### Panel A: The Analyst Forecast Benchmark

<table>
<thead>
<tr>
<th>Size</th>
<th>$N_{-1}$</th>
<th>$N_{+1}$</th>
<th>$\tau_{-1}^{BD}$</th>
<th>$\tau_{+1}^{BD}$</th>
<th>$\tau_{-1}^{BP}$</th>
<th>$\tau_{+1}^{BP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.022</td>
<td>78</td>
<td>174</td>
<td>-3.87***</td>
<td>7.22***</td>
<td>-2.07**</td>
</tr>
<tr>
<td>Rule V</td>
<td>0.032</td>
<td>115</td>
<td>219</td>
<td>-2.79***</td>
<td>7.85***</td>
<td>-1.18</td>
</tr>
<tr>
<td>Rule II</td>
<td>0.037</td>
<td>115</td>
<td>219</td>
<td>-3.58***</td>
<td>6.88***</td>
<td>-2.28**</td>
</tr>
<tr>
<td>Rule IV</td>
<td>0.038</td>
<td>115</td>
<td>219</td>
<td>-3.58***</td>
<td>6.88***</td>
<td>-2.46***</td>
</tr>
<tr>
<td>High</td>
<td>0.048</td>
<td>151</td>
<td>239</td>
<td>-1.37*</td>
<td>5.99***</td>
<td>-1.06</td>
</tr>
</tbody>
</table>

#### Panel B: The Zero Earnings Benchmark

<table>
<thead>
<tr>
<th>Size</th>
<th>$N_{-1}$</th>
<th>$N_{+1}$</th>
<th>$\tau_{-1}^{BD}$</th>
<th>$\tau_{+1}^{BD}$</th>
<th>$\tau_{-1}^{BP}$</th>
<th>$\tau_{+1}^{BP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.184</td>
<td>116</td>
<td>277</td>
<td>-4.42***</td>
<td>6.78***</td>
<td>-2.90***</td>
</tr>
<tr>
<td>Rule V</td>
<td>0.224</td>
<td>136</td>
<td>330</td>
<td>-4.73***</td>
<td>7.89***</td>
<td>-3.06***</td>
</tr>
<tr>
<td>Rule II</td>
<td>0.241</td>
<td>142</td>
<td>356</td>
<td>-5.15***</td>
<td>8.69***</td>
<td>-3.27***</td>
</tr>
<tr>
<td>Rule IV</td>
<td>0.264</td>
<td>152</td>
<td>376</td>
<td>-5.23***</td>
<td>8.91***</td>
<td>-3.37***</td>
</tr>
<tr>
<td>High</td>
<td>0.304</td>
<td>164</td>
<td>421</td>
<td>-6.19***</td>
<td>9.88***</td>
<td>-3.81***</td>
</tr>
</tbody>
</table>

#### Panel C: The Earnings Changes Benchmark

<table>
<thead>
<tr>
<th>Size</th>
<th>$N_{-1}$</th>
<th>$N_{+1}$</th>
<th>$\tau_{-1}^{BD}$</th>
<th>$\tau_{+1}^{BD}$</th>
<th>$\tau_{-1}^{BP}$</th>
<th>$\tau_{+1}^{BP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.104</td>
<td>136</td>
<td>232</td>
<td>-1.96**</td>
<td>3.85***</td>
<td>-1.25</td>
</tr>
<tr>
<td>Rule V</td>
<td>0.124</td>
<td>160</td>
<td>276</td>
<td>-1.82**</td>
<td>4.77***</td>
<td>-1.07</td>
</tr>
<tr>
<td>Rule II</td>
<td>0.133</td>
<td>164</td>
<td>289</td>
<td>-2.57***</td>
<td>4.86***</td>
<td>-1.47*</td>
</tr>
<tr>
<td>Rule IV</td>
<td>0.146</td>
<td>172</td>
<td>313</td>
<td>-3.15***</td>
<td>5.49***</td>
<td>-1.82**</td>
</tr>
<tr>
<td>High</td>
<td>0.166</td>
<td>196</td>
<td>350</td>
<td>-2.61**</td>
<td>6.22***</td>
<td>-1.39*</td>
</tr>
</tbody>
</table>
Epanechnikov kernel (significant at the 5% level). These results support a lower importance of the earnings changes benchmark. Overall, the results remain fairly stable when the Gaussian kernel is replaced by alternative non-normal weighting functions.

### Table 6.8

*Alternative Kernel Functions*

#### Panel A: The Analyst Forecast Benchmark

<table>
<thead>
<tr>
<th>Kernel:</th>
<th>Gaussian</th>
<th>Uniform</th>
<th>Parzen</th>
<th>Epanechnikov</th>
<th>Biweight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{BP}$</td>
<td>-1.18</td>
<td>-2.07**</td>
<td>-2.42***</td>
<td>-2.50***</td>
<td>-2.36***</td>
</tr>
<tr>
<td>$\tau_{BP}^{-1}$</td>
<td>8.27***</td>
<td>7.00***</td>
<td>5.21***</td>
<td>5.81***</td>
<td>5.63***</td>
</tr>
</tbody>
</table>

#### Panel B: The Zero Earnings Benchmark

<table>
<thead>
<tr>
<th>Kernel:</th>
<th>Gaussian</th>
<th>Uniform</th>
<th>Parzen</th>
<th>Epanechnikov</th>
<th>Biweight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{BP}$</td>
<td>-3.06***</td>
<td>-4.07***</td>
<td>-2.51***</td>
<td>-3.17***</td>
<td>-2.65***</td>
</tr>
<tr>
<td>$\tau_{BP}^{-1}$</td>
<td>7.65***</td>
<td>4.45***</td>
<td>2.95***</td>
<td>3.90***</td>
<td>3.65***</td>
</tr>
</tbody>
</table>

#### Panel C: The Earnings Changes Benchmark

<table>
<thead>
<tr>
<th>Kernel:</th>
<th>Gaussian</th>
<th>Uniform</th>
<th>Parzen</th>
<th>Epanechnikov</th>
<th>Biweight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{BP}$</td>
<td>-1.07</td>
<td>-1.10</td>
<td>-1.15</td>
<td>-1.88**</td>
<td>-1.28</td>
</tr>
<tr>
<td>$\tau_{BP}^{-1}$</td>
<td>5.11***</td>
<td>3.42***</td>
<td>1.64*</td>
<td>1.95**</td>
<td>2.29**</td>
</tr>
</tbody>
</table>

The table summarizes BP test statistics based on five different kernel functions for the first interval below ($\tau_{BP}^{-1}$) and above ($\tau_{BP}^{+1}$) the analyst forecast (Panel A), zero earnings (Panel B), and earnings changes benchmark (Panel C). Kernel bandwidth is calculated using Silverman’s (1986) rule of thumb for Gaussian kernels (Rule V). For non-normal kernels, optimal bandwidth is rescaled based on the factors suggested in Scott (1992, Table 6.3, p. 142). Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (****) asterisks, respectively.

### 6.4.4.3 Alternative Deflators

Most studies applying the distributional approach focus on some deflated form of net income or EPS. The analyses in this chapter, however, are based on undeflated EPS for two reasons. First, I assume that managers and stakeholders focus rather on EPS than any other deflated or undeflated earnings figure. As a mandatory disclosure in financial reporting, commonly cited performance measure in the financial press, and key metric in asset valuation, EPS plays an outstanding role in the communication with the stakeholders of the firm (see, e.g., Durtschi and Easton, 2005, 2009). Second, the analysis in Section 6.2 (see Table 6.1) confirms that if net income is scaled, average number of shares outstanding is the most appropriate deflator. Though I regard undeflated EPS as superior in my setting, this section extends the previous tests with an analysis of four alternative
scaling variables. These include lagged sales per share, lagged total assets per share, lagged equity per share, and lagged market capitalization per share. BD and BP test statistics for the first interval above and below the respective benchmark are summarized in Table 6.9. Related frequency histograms are depicted in Figures 6.11 to 6.13.

The second row in Panel A of Table 6.9 shows that both test statistics detect a highly significant (1% level) irregularity at the first interval above the analyst forecast benchmark. This result is consistent with the considerably peaked histograms depicted in Figure 6.11. For the first interval below zero earnings surprises, the BD test statistic remains significant at at least 5% across all deflators. The more rigid BP statistic (not significant for undeflated EPS) gains significance when earnings surprises are deflated by total assets, equity, or market value. Hence, deflating EPS intensifies the underrepresentation of observations in interval $-1$. The frequency histograms in Figure 6.12 illustrate that scaling heavily distorts the distribution of EPS. As a consequence, test statistics slightly increase for interval $-1$ and decrease for interval $+1$. These results are summarized in Panel B of Table 6.9: While all test statistics for interval $-1$ remain significant at the 1% level, BD and BP test statistics for interval $+1$ drop, at worst, to the 5% and 10% level, respectively, when earnings are scaled by total assets, equity, or market value. The results in Panel C are less drastic. For interval $+1$, all test statistics remain highly significant across all alternative deflators. The BD test statistic for interval $-1$ (significant at the 5% level for undeflated EPS) increases to a significance level of 1% when EPS changes are deflated. The BP statistic (not significant for undeflated EPS) gains significance on at least 10% when earnings are deflated by sales, total assets, or equity.

Deflation heavily affects the distribution of EPS and related test statistics, while its effect on the distributions of earnings surprises and EPS changes is comparably low. I attribute these results to systematic differences in deflators for profit and loss observations that distort the underlying distribution (Durtschi and Easton, 2005, 2009). However, irregularities in the distribution of EPS remain significant at generally accepted levels even after deflation. Overall, the results remain fairly stable in the case of earnings deflation.

45Since all deflator variables are on a per share basis, the results are comparable with studies that focus on deflated net income (i.e., net income divided by total sales is, for example, identical to EPS divided by total sales per share). The deflator variables lagged sales (WS: 01001), lagged total assets (WS: 18184 + 02999), lagged total equity (WS: 03426 + 03501), and lagged market value (DS: MV) are gathered from Worldscope and Datastream. The weighted average number of shares outstanding during the year is available on Worldscope (WS: 05192).
### Table 6.9
Deflation Analysis

#### Panel A: The Analyst Forecast Benchmark

<table>
<thead>
<tr>
<th></th>
<th>Undeflated EPS</th>
<th>Sales per Share</th>
<th>Total Assets per Share</th>
<th>Equity per Share</th>
<th>Market Value per Share</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>$\tau^{BD}$</td>
<td>$\tau^{BP}$</td>
<td>N</td>
<td>$\tau^{BD}$</td>
</tr>
<tr>
<td>Bin −1</td>
<td>115</td>
<td>−2.79$^a$</td>
<td>−1.18</td>
<td>119</td>
<td>−1.92$^b$</td>
</tr>
<tr>
<td>Bin +1</td>
<td>219</td>
<td>7.85$^a$</td>
<td>8.27$^a$</td>
<td>219</td>
<td>6.35$^a$</td>
</tr>
</tbody>
</table>

#### Panel B: The Zero Earnings Benchmark

<table>
<thead>
<tr>
<th></th>
<th>Undeflated EPS</th>
<th>Sales per Share</th>
<th>Total Assets per Share</th>
<th>Equity per Share</th>
<th>Market Value per Share</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>$\tau^{BD}$</td>
<td>$\tau^{BP}$</td>
<td>N</td>
<td>$\tau^{BD}$</td>
</tr>
<tr>
<td>Bin −1</td>
<td>136</td>
<td>−4.73$^a$</td>
<td>−3.06$^a$</td>
<td>58</td>
<td>−5.67$^a$</td>
</tr>
<tr>
<td>Bin +1</td>
<td>330</td>
<td>7.89$^a$</td>
<td>7.65$^a$</td>
<td>191</td>
<td>3.89$^a$</td>
</tr>
</tbody>
</table>

#### Panel C: The Earnings Changes Benchmark

<table>
<thead>
<tr>
<th></th>
<th>Undeflated EPS</th>
<th>Sales per Share</th>
<th>Total Assets per Share</th>
<th>Equity per Share</th>
<th>Market Value per Share</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>$\tau^{BD}$</td>
<td>$\tau^{BP}$</td>
<td>N</td>
<td>$\tau^{BD}$</td>
</tr>
<tr>
<td>Bin −1</td>
<td>160</td>
<td>−1.82$^b$</td>
<td>−1.07</td>
<td>153</td>
<td>−2.66$^a$</td>
</tr>
<tr>
<td>Bin +1</td>
<td>276</td>
<td>4.77$^a$</td>
<td>5.11$^a$</td>
<td>274</td>
<td>4.67$^a$</td>
</tr>
</tbody>
</table>

The table summarizes Burgstahler and Dichev (1997) ($\tau^{BD}$) and Bollen and Pool (2009) ($\tau^{BP}$) test statistics for irregularities in the distribution of undeflated and deflated EPS variables. Deflators include lagged sales per share, lagged total assets per share, lagged equity per share, and lagged market capitalization per share. Undeflated and deflated EPS variables are truncated at the upper and lower 1% of their yearly distributions. Kernel bandwidths are calculated using Silverman’s (1986) rule of thumb (Rule V) and equal binwidths. Significance (one-tailed) at the 1%, 5%, and 10% level is indicated by $^a$, $^b$, and $^c$, respectively.
The figure shows frequency histograms of undeflated and deflated earnings surprises per share. Deflators include lagged sales per share, lagged total assets per share, lagged equity per share, and lagged market capitalization per share. Total number of observations is 1,263 (extended sample). Undeflated and deflated earnings surprises are truncated at the upper and lower 1% of their yearly distributions. Binwidths are calculated using Silverman’s (1986) rule of thumb (Rule V).
Fig. 6.12.—**Frequency Histograms of Deflated EPS.** The figure shows frequency histograms of undeflated and deflated EPS. Deflators include lagged sales per share, lagged total assets per share, lagged equity per share, and lagged market capitalization per share. Total number of observations is 2,100 (extended sample). Undeflated and deflated EPS are truncated at the upper and lower 1% of their yearly distributions. Binwidths are calculated using Silverman’s (1986) rule of thumb (Rule V).
FIG. 6.13.—**Frequency Histograms of Deflated EPS Changes.** The figure shows frequency histograms of undeflated and deflated EPS changes. Deflators include lagged sales per share, lagged total assets per share, lagged equity per share, and lagged market capitalization per share. Total number of observations is 2,099 (extended sample). Undeflated and deflated changes in EPS are truncated at the upper and lower 1% of their yearly distributions. Binwidths are calculated using Silverman’s (1986) rule of thumb (Rule V).
6.4.4.4 The Effect of Tax Asymmetries

Beaver et al. (2007) suggest that the discontinuity around zero earnings is at least in part attributable to income tax effects. They show that firms with pretax losses have significantly lower effective tax rates (ETRs) than firms with pretax profits. Consequently, higher ETRs for profit firms draw profit observations towards zero and strengthen the distribution’s peak at the zero earnings benchmark. To rule out the possibility that my results are driven by different ETR levels, this section focuses on how tax rate asymmetries affect the distribution of EPS.46

The main assumption underlying the tax effect is a significant asymmetry in ETRs for profit and loss firms. In the extended sample, mean (median) ETRs of profit and loss firms are 27% (30%) and 6% (0%), respectively.47 In line with the reasoning in Beaver et al. (2007), profit firms have significantly higher ETRs than loss firms.48 Moreover, the most drastic change in ETRs occurs at the zero earnings reference point: As depicted in the last column of Table 6.10, median ETR is 0% for intervals $-3$ and $-2$, increases to 7% in interval $-1$, and subsequently rises to about 30% after crossing the red line of zero earnings.49 To untangle how much of the discontinuity is related to tax rate asymmetries, Figure 6.14 compares the distributions of reported EPS (in black) and pretax EPS (in gray). Since pretax EPS are not available from financial databases, I multiply (after tax) EPS with the ratio of pretax net income to (after tax) net income. If differences in ETRs intensify the irregularity in the distribution of EPS, I expect the distribution of pretax EPS to be considerably smoother than the distribution of EPS. The histograms in Panel A confirm this conjecture: There are significantly less pretax EPS observations immediately above

---

46Important to note, Beaver et al. (2007) do not claim that the discontinuity is solely due to nondiscretionary factors and emphasize that a discontinuity at zero may be the joint effect of both discretionary (i.e., earnings management) and nondiscretionary (i.e., tax asymmetries) effects. Moreover, finding asymmetric ETRs to cause the kink in the distribution of EPS does not rule out the possibility that earnings are managed to meet or beat the benchmark, since tax expenses themselves maybe subject to managerial discretion. However, if the distribution is entirely smooth after controlling for asymmetric income tax effects, there is no reason to assume that management engages in earnings management beyond the manipulation of income taxes.

47To exclude outliers, I delete ETR observations below the 1st and above the 99th percentile of their yearly distributions.

48Based on a two-tailed test of differences in means, ETRs of profit firms are significantly different from those of loss firms at the 1% level.

49German corporate taxes include income taxes, a solidarity surcharge, and a local trade tax. According to KPMG’s Corporate and Indirect Tax Survey (KPMG, 2010), German corporate taxes amounted to about 38% between 2005 and 2007. From 2008 on, average corporate taxes dropped to about 29% as a consequence of a recent reform on corporate taxes. Summed up for the years 2005 to 2009, the average corporate tax rate is about 35%. Taking the frequent utilization of taxable loss carryforwards into account, a slightly lower ETR of about 30% is fairly reasonable.
FIG. 6.14.—Effect of Tax Asymmetries. Panel A displays the frequency histograms of EPS (in black) and pretax EPS (in gray) for 2,100 observations (extended sample) in the period 2005 to 2009. Pretax EPS are calculated as the ratio of pretax net income (WS: 01401) to net income (WS: 01751 + 01501) times reported EPS. Panels B and C depict the Burgstahler and Dichev (1997) (BD) and Bollen and Pool (2009) (BP) test statistics of distributional discontinuities, respectively. Kernel bandwidths are calculated using Silverman’s (1986) rule of thumb (Rule V) based on the distribution of EPS. The binwidths of the histograms in Panel A correspond to the bandwidths of the kernel estimators.
zero than EPS observations, indicating that including income taxes draws positive pretax EPS closer to zero. For negative EPS, in contrast, the distribution is nearly unaffected by the inclusion of income taxes. Although tax asymmetries intensify the jump in the distribution of EPS, the test statistics in Panels B and C of Figure 6.14 suggest that controlling for tax asymmetries does not completely eliminate discontinuities. As shown in Table 6.10, the BP test statistic for interval $-1$ decreases by 1.09 to $-1.97$ when the underlying distribution is based on pretax EPS instead of reported EPS. Similarly, the test statistic for interval $+1$ decreases from $7.65$ to $6.21$. However, both test statistics remain significant at a level of at least 5%.

### Table 6.10

<table>
<thead>
<tr>
<th>Bin</th>
<th>$N$</th>
<th>$\tau_{BP}$</th>
<th>$N$</th>
<th>$\tau_{BP}$</th>
<th>$N$</th>
<th>$\tau_{BP}$</th>
<th>ETR$_{med}$ in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-3$</td>
<td>78</td>
<td>0.09</td>
<td>69</td>
<td>$-0.58$</td>
<td>9</td>
<td>0.67</td>
<td>0</td>
</tr>
<tr>
<td>$-2$</td>
<td>88</td>
<td>$-2.50^{***}$</td>
<td>93</td>
<td>$-1.54^*$</td>
<td>$-5$</td>
<td>$-0.96$</td>
<td>0</td>
</tr>
<tr>
<td>$-1$</td>
<td>136</td>
<td>$-3.06^{***}$</td>
<td>136</td>
<td>$-1.97^{**}$</td>
<td>0</td>
<td>$-1.09$</td>
<td>7</td>
</tr>
<tr>
<td>$+1$</td>
<td>330</td>
<td>$7.65^{***}$</td>
<td>278</td>
<td>$6.21^{***}$</td>
<td>52</td>
<td>1.44</td>
<td>30</td>
</tr>
<tr>
<td>$+2$</td>
<td>207</td>
<td>$-0.15$</td>
<td>181</td>
<td>0.08</td>
<td>26</td>
<td>$-0.23$</td>
<td>30</td>
</tr>
<tr>
<td>$+3$</td>
<td>158</td>
<td>$-0.24$</td>
<td>135</td>
<td>$-0.53$</td>
<td>23</td>
<td>0.29</td>
<td>30</td>
</tr>
</tbody>
</table>

The table summarizes Bollen and Pool (2009) ($\tau_{BP}$) test statistics for irregularities in the distribution of EPS and pretax EPS ($\text{EPS}_{\text{ptax}}$). Pretax EPS are calculated as the ratio of pretax net income (WS: 01401) to net income after taxes (WS: 01751 + 01501) times reported EPS. ETR$_{med}$ denotes median ETR (i.e., tax expense divided by pretax net income) for all observations in the respective interval. The distribution covers 2,100 observations (extended sample) in the period 2005 to 2009. Kernel bandwidths are calculated using Silverman’s (1986) rule of thumb (Rule V) based on the distribution of EPS. Binwidths correspond to the bandwidths of the kernel estimators. Significance (one-tailed) at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***).

Overall, controlling for tax effects does not rule out the possibility that the discontinuity in the distribution of EPS is related to benchmark-driven earnings management activity beyond income tax management.
6.5 Summary

In this chapter, I analyzed the pooled frequency distributions of earnings metrics to shed light on the prevalence of earnings and/or expectations management to achieve targets. The distributions of earnings surprises, EPS, and EPS changes all exhibit significant irregularities around zero. Specifically, I find significantly more observations just above the respective earnings targets than expected under an unmanaged reference distribution. This holds for the large extended sample as well as for the smaller reference sample. The evidence is consistent with managers’ and stakeholders’ preference for above target earnings. For the earnings management story to be complete, I further assume to find significantly less observations just below the target than expected. At the zero earnings benchmark, the two test statistics of distributional discontinuity confirm a highly significant underrepresentation of observations in the “small miss” interval. Based on an estimated reference distribution of unmanaged earnings, a maximum of 22% of the observations in the “small miss” interval supposedly engage in earnings manipulations to avoid reporting a loss. The results at the analyst forecast benchmark and earnings changes benchmark are considerably weaker: While the BD test statistics are highly significant, the BP test statistics heavily lose significance in both samples. Consistent with lower test statistics, only a maximum of 10% of “small miss” firms are suspected to engage in earnings or expectations management to achieve the mean consensus forecast. Similarly, a maximum of 8% of the “small miss” observations potentially engage in earnings management to avoid earnings declines. The results are robust to several different specifications, including alternative interval widths, different scaling variables, kernel choice, and asymmetric income tax effects.

The evidence in this chapter generally confirms the importance of earnings benchmarks in Germany. Interestingly, German managers seem to regard the zero earnings benchmark as the most important earnings target, followed by analyst forecasts, and prior year’s earnings. This result contradicts with prior US evidence of analyst forecast priority (e.g., Brown and Caylor, 2005) and the capital market incentives documented in Chapter 5. The following chapter provides a detailed analysis of “small beat” and “small miss” firms to identify specific types of discretionary activities.
Chapter 7

Management’s Techniques to Achieve Benchmarks

Though different in intensity, the previous section reveals considerable distributional discontinuities at all three potential earnings benchmarks. Commonly, these patterns are attributed to some kind of earnings management to shift earnings across thresholds if they would be barely missed otherwise (e.g., Burgstahler and Dichev, 1997; DeGeorge et al., 1999). Dechow et al. (2003) and subsequent studies (see, e.g., Coulton et al., 2005; Ayers et al., 2006; Lin et al., 2006) use measures of aggregate accruals to challenge whether the kink in the earnings distribution is caused by managerial discretion. In a similar vein, several authors consider management of specific accruals (e.g., tax accounts) to achieve earnings targets (e.g., Plummer and Mest, 2001; Beatty et al., 2002; Dhaliwal et al., 2004; Frank and Rego, 2006; Caylor, 2010). More recently, a fast-growing body of literature examines measures of real earnings management in the vicinity of earnings thresholds (e.g., Roychowdhury, 2006; Gunny, 2010). Eventually, with a focus on the analyst forecast benchmark, Matsumoto (2002), Bartov et al. (2002), and subsequent studies (e.g., Burgstahler and Eames, 2006; Nöldeke, 2007b) investigate whether analyst guidance explains the accumulation of small positive earnings surprises. Following this stream of research, I test the following three hypotheses to investigate whether a distributional kink in the distribution of earnings surprises, EPS, and EPS changes is attributable to specific types managerial discretion (RQ (3)):

\[ H_A (3a): \] German firms engage in accrual-based, real earnings management, and/or analyst guidance to avoid negative earnings surprises.
Hₐ (3b): German firms engage in accrual-based and/or real earnings management to avoid reporting a loss in earnings per share.

Hₐ (3c): German firms engage in accrual-based and/or real earnings management to avoid reporting declines in earnings per share.

Testing these hypotheses serves in two ways: First, it provides a rigid test of earnings management activity at earnings thresholds. Finding a relation between earnings management proxies and benchmark achievement confirms the presumption that distributional discontinuities are evidence of managerial discretion. Failing to detect such a relationship, however, indicates that the observed distributional patterns may stem from other factors than earnings management. Second, it provides explorative evidence of earnings management techniques applied to avoid missing an earnings benchmark.

This chapter is structured as follows: Section 7.1 explains the empirical methodology to test for specific earnings management around benchmarks. In Section 7.2, I define the required variables and explain the derivation of several earnings management measures. The sample selection procedure and descriptive statistics are provided in Section 7.3. Empirical results are presented in detail in Section 7.4 and summarized in Section 7.5.

### 7.1 Empirical Methodology

The empirical methodology follows the basic idea provided by Dechow et al. (2003) for discretionary accruals: If firms boost earnings to achieve a benchmark, then these suspect firms (the small meet or beat group, SMBE) should on average exhibit higher discretionary accruals than the group of “just miss”-firms (the small miss group, SMISS) and all other firms (OTHERS).

I adopt this approach for discretionary accruals and extend it to nine other proxy variables of accrual-based, real earnings management, and analyst guidance. These include:

- $FC'$ as proxy for managers’ influence on financial analysts to provide beatable targets (see Section 7.2.3),
- $DACC$ as proxy for aggregate accruals management (see Section 7.2.1.1),
REV1/2′ as two proxies (REV1′ and REV2′) for discretionary or premature revenue recognition (see Section 7.2.1.2),

DEXP′ as proxy for real earnings management by means of cutting discretionary expenses (see Section 7.2.2.1),

RND′ as proxy for real earnings management by means of cutting R&D spending (see Section 7.2.2.2),

SGA′ as proxy for real earnings management by means of cutting SG&A expenses (see Section 7.2.2.2),

PROD1/2′ as two proxies (PROD1′ and PROD2′) for real earnings management by means of overproduction (see Section 7.2.2.1 and 7.2.2.2), and

DGAIN′ as proxy for real earnings management by means of selling fixed assets (see Section 7.2.2.2).

In univariate tests, I compare the means of these earnings management proxies for the SMBE, the SMISS, and the OTHERS group. Under the earnings management hypothesis, I expect the earnings management proxy of the SMBE group to be significantly larger than the respective proxies of the SMISS and OTHERS group.1 Taking discretionary accruals DACC as an example, I use a simple t-test of differences in means to test the two null hypotheses H0: DACC_{SMBE} ≤ DACC_{SMISS} and H0: DACC_{SMBE} ≤ DACC_{OTHERS}. Rejecting both hypotheses at generally accepted levels of significance indicates that the suspect SMBE group exhibits on average higher discretionary accruals than the SMISS and OTHERS group; a finding in support of accrual-based earnings management at the respective earnings threshold.

Since univariate tests do not control for other factors that may affect the relation of earnings management proxies and benchmark achievement, I expect a multivariate regression approach to provide more reliable results. To test for benchmark related earnings management, I augment the model proposed by Gunny (2010) and estimate the following equation:

---

1For earnings guidance (FC′) and cost cutting proxies (DEXP′, RND′, and SGA′) negative values indicate that these methods are used to increase earnings. To ease interpretation, these proxies are multiplied with −1.
PROXY\textsubscript{it} = \alpha + \beta_1 \text{SMBE}_{t}^Q + \beta_2 \text{SMISS}_{t}^Q \\
+ \beta_3 \text{SIZE}_{t} + \beta_4 \text{MTB}_{t} + \beta_5 \text{ROA}_{t} + \varepsilon_{it} \tag{7.1.1}

where

\begin{itemize}
  \item \text{PROXY}_{it} \text{ represents 10 proxy variables for accrual-based earnings management, real earnings management, and expectations management (see above),}
  \item \text{SMBE}_{t}^Q \text{ is a dummy variable equal to 1 if firm } i \text{ barely achieves benchmark } Q \text{ in year } t \text{ (i.e., is in the SMBE group) and 0 otherwise,}
  \item \text{SMISS}_{t}^Q \text{ is a dummy variable equal to 1 if firm } i \text{ barely misses benchmark } Q \text{ in year } t \text{ (i.e., is in the SMISS group) and 0 otherwise,}
  \item \text{SIZE}_{it} \text{ is the size of firm } i \text{ in year } t,
  \item \text{MTB}_{it} \text{ is the M/B-ratio of firm } i \text{ in year } t, \text{ and}
  \item \text{ROA}_{it} \text{ is firm performance measured as return on total assets of firm } i \text{ in year } t.
\end{itemize}

The interpretation of coefficients is straightforward: If the respective earnings management proxy is significantly larger for the SMBE group than for the OTHERS group, the coefficient \( \beta_1 \) on \( \text{SMBE}_{t}^Q \) is supposed to be significantly positive. If the respective earnings management proxy for the SMBE group is significantly larger than for the SMISS group, the coefficient \( \beta_1 \) on \( \text{SMBE}_{t}^Q \) is expected to be larger than the coefficient \( \beta_2 \) on \( \text{SMISS}_{t}^Q \). I use tests of differences in coefficients to test \( H_0: \beta_1 \leq \beta_2 \).

### 7.2 Variable Definitions

This section summarizes the definition of the dependent and independent variables of the regression model depicted in Eq. (7.1.1). The definition of the independent variable is straightforward: \( \text{SMBE}_{t}^Q \) (\( \text{SMISS}_{t}^Q \)) is 1 if earnings surprises, EPS, or changes in EPS fall in the first interval above (below) the respective benchmark and 0 otherwise. Interval width is thereby identical to the widths used for the distributional tests in Chapter 6. For the analyst forecast benchmark, for instance, \( \text{SMBE}_{t}^{FC} \) is set to 1 if earnings surprise is equal or greater than 0.000 and smaller than...
0.032. Similarly, $SMISS_{FC}$ is 1 if earnings surprises are equal or greater than $-0.032$ and smaller than 0.000. $SIZE_{it}$ is calculated as the natural logarithm of total assets (WS: 18184 + 02999). $MTB_{it}$ is market capitalization (DS: MV) divided by total common equity (WS: 03501). $ROA_{it}$ is the ratio of net income including minority interest (WS: 01751 + 01501) to reported total assets (WS: 18184 + 02999). The data is gathered from Worldscope and Datastream.

The definition of the proxy variables is more complicated and described in detail in the following sections.

### 7.2.1 Proxies for Accrual-based Earnings Management

Measures of accrual-based earnings management focus either on aggregate or specific accruals. Although the analysis of specific accruals may offer a higher degree of measurement precision, its advantages only emerge in settings with strong incentives for the manipulation of the observed accrual (Alcarria Jaime and De Albornoz Noguer, 2004). In contrast, when the researcher assumes motivation for earnings management to be unrelated to specific accruals, aggregate accrual measures prove to be more powerful. I use discretionary accruals as an aggregate measure of accrual-based earnings management and discretionary revenues as proxy for premature revenue recognition (a specific type of accruals management). I focus on revenue recognition because revenues are the largest earnings component for most firms and commonly subject to managerial discretion (Stubben, 2010).

#### 7.2.1.1 Discretionary Accruals

Measuring the extent of accrual manipulation requires an estimate of premanaged accruals. Total accruals are therefore decomposed into two components. The first component comprises the expected premanaged and thus normal or nondiscretionary portion of total accruals. The second component represents the unexpected and thus abnormal or discretionary part of accruals. Figure 7.1 illustrates the different components of earnings. Discretionary accruals are used as a proxy for the magnitude and direction of earnings management. Firms exhibiting high levels of discretionary accruals are suspected to engage in earnings management activity. Positive discretionary accruals are assumed to result from income increasing earnings management, whereas negative discretionary accruals are a consequence of income decreasing manipulations. Given the level of total accruals ($ACC$) and an estimate of nondiscretionary accruals ($NDACC$), discretionary
accruals \((DACC)\) are calculated as:

\[
DACC = ACC - NDACC.
\]  \hspace{1cm} (7.2.1)

The outcome of studies of accrual-based earnings management depends on the researcher’s ability to find valid estimates of nondiscretionary accruals and thus discretionary accruals. In the remainder of this section, I present and explain several different approaches to measure discretionary accruals and subsequently choose the most appropriate model for my research setting.

**Fig. 7.1.**—**Discretionary and Nondiscretionary Accruals.** The figure illustrates the three important components of total earnings considered in earnings management research.

### 7.2.1.1.1 The Jones Model

The early contributions of Healy (1985) and DeAngelo (1986) provide simple estimates for nondiscretionary accruals. The Healy (1985)-approach implicitly assumes that nondiscretionary accruals follow a mean-reverting process with zero growth (Bartov et al., 2001). If no earnings management occurred in prior periods, nondiscretionary accruals can be estimated as the mean of total accruals in the past. In contrast to Healy (1985), DeAngelo (1986) suggests nondiscretionary accruals to follow a random walk (Bartov et al., 2001). Assuming no earnings management occurred in the previous period, the best estimate for nondiscretionary accruals is total accruals of the prior year. A major caveat of these methodologies is that they neglect the economic circumstances of the firm. Kaplan (1985, p. 111), in his comment on Healy (1985), puts it as follows:

This seems like an overly restrictive assumption since the levels of working capital accounts will fluctuate depending upon the economic circumstances of the firm. These working capital accounts serve useful economic purposes beyond providing a domain for the opportunistic behavior of managers. It would be preferable to have
a simple model of how accounts receivable, inventory and accounts payable are expected to vary depending upon changes in sales and production each year.

Jones (1991) addresses this estimation problem by regressing total accruals over several independent variables supposed to affect nondiscretionary accruals. Following Kaplan’s (1985) suggestion concerning working capital accruals, she first adds change in revenues as independent variable. Increasing revenues are usually associated with higher levels of accounts receivable and accounts payable. In addition to the sales effect on working capital accruals, she considers the level of property, plant, and equipment to account for depreciation related nondiscretionary accruals. The Jones Model measures discretionary accruals in analogy to the event study methodology known from finance. The event-period is the period in which earnings management is suspected, while the estimation period includes several precedent periods in which the phenomenon under investigation is not assumed to be present. That is, earnings management is suspected to be zero. The relation between event and estimation period is illustrated in Figure 7.2. During the estimation period, Jones (1991) performs the following time-series regression to estimate how changes in sales (ΔREV) and level of property, plant, and equipment (PPE) affect total accruals:

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{\Delta REV_{it}}{TA_{i,t-1}} + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \epsilon_{it},
\]

where \(ACC_{it}\) is total accruals of firm \(i\) in year \(t\), \(\Delta REV_{it}\) is the change in sales of firm \(i\) from \(t-1\) to \(t\), and \(PPE_{it}\) is firm \(i\)’s property, plant, and equipment at the end of year \(t\). All variables are scaled by lagged total assets \(TA_{i,t-1}\) to account for firm size differences and mitigate heteroskedasticity.

The coefficient estimates \(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1,\) and \(\hat{\beta}_2\) describe the relation between total accruals and the independent variables. The coefficient on scaled \(PPE_{it}\) (\(\hat{\beta}_2\)), for instance, is supposedly negative, because a higher level of property, plant, and equipment implies increased depreciation charges. In the event period (the period with potential earnings manipulation), the coefficient

---

2I augment the original Jones Model with an intercept \((\alpha_1)\). Adding an intercept controls for heteroskedasticity and alleviates problems related to an omitted size variable (Kothari et al., 2005).

3The disturbances \(\epsilon_t\) are called heteroskedastic when their variances are not constant across observations. Although coefficients remain unbiased in the presence of heteroskedasticity, standard errors are biased and hypotheses tests potentially invalid. For details on causes, consequences, and remedies of heteroskedasticity see, e.g., Greene (2003, pp. 216–249) and Baltagi (2008, pp. 98–109). Since lagged assets are assumed to be positively associated with the variance of the disturbance term, weighting both sides of the equation with \(TA_{i,t-1}\) is supposed to mitigate heteroskedasticity (Jones, 1991).
7. Management’s Techniques to Achieve Benchmarks

<table>
<thead>
<tr>
<th>Estimation Period</th>
<th>Event Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>(No EM Activity Suspected)</td>
<td>(EM Suspected)</td>
</tr>
<tr>
<td>$t - n$</td>
<td>$t - 1$</td>
</tr>
<tr>
<td>$t$</td>
<td>$t + 1$</td>
</tr>
</tbody>
</table>

**Fig. 7.2.** — **Event Study Methodology in Earnings Management Research.** The figure illustrates the idea of estimating discretionary accruals from a time-series of accounting variables. The estimation window, usually spanning over several years, is the period for which the accrual model’s coefficients are estimated. The event window is the period when earnings management activity is expected to take place. Abnormal accruals in the event period are calculated as difference of actual total accruals and expected total accruals based on the coefficient estimates during the estimation period.

Estimates $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\beta}_1$, and $\hat{\beta}_2$ are plugged in equation Eq. (7.2.3) to calculate nondiscretionary accruals ($NDACC_{it}$):$^4$

\[
NDACC_{it} = \hat{\alpha}_1 + \hat{\alpha}_2 \frac{1}{TA_{i,t-1}} + \hat{\beta}_1 \frac{\Delta REV_{it}}{TA_{i,t-1}} + \hat{\beta}_2 \frac{PPE_{it}}{TA_{i,t-1}}.
\]  

Subtracting nondiscretionary ($NDACC_{it}$) from total accruals $ACC_{it}$ then gives discretionary accruals ($DACC_{it}$).

### 7.2.1.1.2 The Cross-Sectional Jones Model

The Jones Model is a firm-specific measure of discretionary accruals based on a time-series estimation. To deliver unbiased estimates of discretionary accruals, time-series models assume the relation between the dependent and independent variables to be stationary (Jones, 1991). Taking structural changes (e.g., business or accrual policy decisions) into account, long time-series may produce erroneous estimates of discretionary accruals (Ronen and Yaari, 2008, pp. 411–414). Discretionary accruals are assumed to be zero during the estimation period. If firms, however, manage earnings during the estimation period, then coefficients are biased towards not detecting or incorrectly detecting earnings management (Ronen and Yaari, 2008, pp. 407–411). DeFond and Jiambalvo (1994) circumvent these problems and develop a cross-sectional version of the standard Jones Model. Instead of estimating coefficients in time-series regressions, they use data from a sample of firms matched on year and industry. The cross-sectional Jones Model relaxes some of the assumptions of the time-series model (such as the no earnings management assump-

---

$^4$In short, scaled discretionary accruals equal the residuals $\hat{\varepsilon}_{it}$ from fitting regression equation Eq. (7.2.2).
tion for prior periods and stationarity). Moreover, it does not require a long and continuous
time-series of observations and is thus less prone to survivorship bias. Albeit these advantages,
it introduces new conjectures about the sample’s cross-sectional properties (Bartov et al., 2001):
First, the model assumes no earnings management activity among matched firms. If matched
firms engage in earnings management, discretionary accrual estimates are potentially biased.
Second, the model requires that matched firms are comparable. Matching on broad industry cat-
egories with a wide range of heterogeneous firms and business models, for example, impedes the
isolation of discretionary from nondiscretionary accrual components. Though most researchers
today agree that the cross-sectional approach dominates the time-series design (Ronen and Yaari,
2008, p. 417), the choice between the two models remains an empirical question (Bartov et al.,
2001). Throughout this study, I solely use cross-sectional accrual models. I do so not only due to
the general preference for cross-sectional specifications in the academic community. More im-
portantly, it is the lack of reliable and long-term time-series data which hampers the application
of the classical Jones Model in the German setting.

7.2.1.1.3 Adapted Versions of the Jones Model

Modified Jones Model (Dechow et al., 1995; Subramanyam, 1996)

Jones Model estimates are biased towards zero when earnings management is based on dis-
cretionary revenue recognition, because the model implicitly assumes no discretion on revenues
in both the estimation and the event period (Dechow et al., 1995). With the modified Jones
Model, Dechow et al. (1995) assume all credit sales in the event period to be discretionary. To
attenuate bias in the case of revenue manipulations, they first estimate the coefficients of the
Jones Model and then use the difference of change in sales and change in trade receivables
(i.e., $\Delta REV - \Delta REC$) to predict nondiscretionary accruals. Subramanyam (1996) adopt Dechow
et al.’s (1995) idea for the cross-sectional specification and estimate the following equation to
predict nondiscretionary accruals:

$$\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \frac{1}{TA_{i,t-1}} + \beta_1 \left[ \frac{\Delta REV_{it} - \Delta REC_{it}}{TA_{i,t-1}} \right] + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \epsilon_{it}. \quad (7.2.4)$$

\[5\] I augment the Modified Jones Model with an intercept ($\alpha_1$). Adding an intercept controls for heteroskedasticity
and alleviates problems related to an omitted size variable (Kothari et al., 2005).
Adapted, Lagged, and Forward-looking Model (Dechow et al., 2003)

The modified Jones Model treats all credit sales as discretionary. Although credit sales are obviously more susceptible to managerial discretion than cash sales, classifying all credit sales as discretionary seems to be an overrestrictive assumption. To overcome this caveat, Dechow et al. (2003) estimate the expected change in accounts receivable for a given change in sales and include this factor into the regression equation of the classical modified Jones Model. This model, referred to as the Adapted Model, is given as:

\[
\frac{\text{ACC}_{it}}{\text{TA}_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{\text{TA}_{i,t-1}} + \beta_1 \left[ \frac{(1 + \hat{\xi}) \Delta \text{REV}_{it} - \Delta \text{REC}_{it}}{\text{TA}_{i,t-1}} \right] + \beta_2 \frac{\text{PPE}_{it}}{\text{TA}_{i,t-1}} + \epsilon_{it}, \tag{7.2.5}
\]

where factor \( \hat{\xi} \) captures the increase in accounts receivable dependent on sales and is estimated as slope coefficient of regressing Eq. (7.2.6) for every industry-year combination:

\[
\frac{\Delta \text{REC}_{it}}{\text{TA}_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{\text{TA}_{i,t-1}} + \hat{\xi} \frac{\Delta \text{REV}_{it}}{\text{TA}_{i,t-1}} + \epsilon_{it}. \tag{7.2.6}
\]

The Lagged Model is an extension of the Adapted Model and takes advantage of accruals’ time-series properties. Although accruals are less persistent than cash flows and reverse through time, Dechow et al. (2003) conjecture that lagged accruals at least partly explain current accruals. The model can be described with Eq. (7.2.7) as

\[
\frac{\text{ACC}_{it}}{\text{TA}_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{\text{TA}_{i,t-1}} + \beta_1 \left[ \frac{(1 + \hat{\xi}) \Delta \text{REV}_{it} - \Delta \text{REC}_{it}}{\text{TA}_{i,t-1}} \right] + \beta_2 \frac{\text{PPE}_{it}}{\text{TA}_{i,t-1}} + \beta_3 \frac{\text{ACC}_{i,t-1}}{\text{TA}_{i,t-1}} + \epsilon_{it}, \tag{7.2.7}
\]

6I augment all three original models from Dechow et al. (2003) with an intercept scaled by lagged total assets \( \frac{1}{\text{TA}_{i,t-1}} \). Allowing the intercept to vary with lagged assets controls for heteroskedasticity (Dechow et al., 2003).
with lagged total accruals \( ACC_{i,t-1} \) as additional regressor.

Augmenting the Lagged Model by future revenue growth as another explanatory variable leads to the Forward-looking Model. The idea is that anticipated sales growth yields a nondiscretionary increase of current accruals; an effect not captured by the aforementioned models. Formally, this model is described as:

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \left( \frac{1}{TA_{i,t-1}} \right) + \beta_1 \left[ \frac{(1 + \hat{\xi}) \Delta REV_{it} - \Delta REC_{it}}{TA_{i,t-1}} \right] \\
+ \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \beta_3 \frac{\Delta REV_{it}}{REV_{i,t-1}} + \epsilon_{it}, \tag{7.2.8}
\]

**Performance-matched and Performance-adjusted Models (Kothari et al., 2005)**

A main caveat of the traditional accrual models described above is their misspecification when applied to firms with extreme performance (see, e.g., Dechow et al., 1995; Kothari et al., 2005). Extreme performance often exhibits patterns of mean reversion or momentum (Kothari et al., 2005). Performance shocks of value stocks, for example, are often mean reverting, while growth stocks’ performance exhibits momentum (Kothari et al., 2005). In both cases, past performance provides information about future performance and expected accruals.\(^7\) Ignoring the effect of past or current performance in estimating discretionary accruals potentially causes severe coefficient bias and unreliable estimates of discretionary accruals. Kothari et al. (2005) address the problem of performance related model misspecification by either including a performance proxy into the model (performance-adjusted model) or by comparison with a control group matched on industry, year, and performance (performance-matched model). The performance-adjusted cross-sectional model is given as:

\(^7\)Dechow et al. (1998) model total accruals \( ACC_t \) as a function of the firm’s expected long-term operating cycle expressed as a fraction of a year (\( \delta \)) and changes in sales (\( \Delta REV_t \)): \( ACC_t = \delta \Delta REV_t \). Hence, expected changes in accruals depend on expected changes in sales and thus expected sales performance.
\[
\frac{ACCI_t}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \left[ \frac{\Delta REV_{it} - \Delta REC_{it}}{TA_{i,t-1}} \right] + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \beta_3 ROA_{it} or_{i,t-1} + \epsilon_{it},
\]

(7.2.9)

where \( ROA_{it} or_{i,t-1} \) is actual or lagged return on total assets as proxy for firm performance. The performance-matched model is based on the cross-sectional modified Jones Model (see Eq. (7.2.4)). Performance-matched discretionary accruals are calculated as discretionary accruals of the observed firm minus discretionary accruals of a firm matched on industry, year, and the closest current or prior year’s return on assets. By doing so, firms that exhibit earnings management through discretionary accruals are those that manage more than would be expected given their level of performance.

### 7.2.1.4 Accuracy of Accrual Models and Model Choice

The validity of earnings management studies depends on the researcher’s ability to identify the most accurate model to measure discretionary accruals. Since studies on earnings management are always joint tests of the applied measurement method and earnings management (Kothari et al., 2005), it is hard to determine whether results are attributable to “true” earnings management or measurement error. This raised the interest in research on discretionary accrual models, their empirical specifications, and the accuracy of the results (see, e.g., Dechow et al., 1995; Jeter and Shivakumar, 1999; Peasnell et al., 2000b; Alcarria Jaime and De Albornoz Noguer, 2004; Kothari et al., 2005). The accuracy of an accrual model is evaluated in two dimensions, namely power and specification (Dechow et al., 1995):

- **Model Power.** A powerful model correctly detects earnings management. Statistically spoken, the model (correctly) rejects the null of no earnings management (at a certain level of confidence) in situations where earnings management has taken place and thus minimizes Type II error.\(^8\)

---

\(^8\)Type II errors arise when the null hypothesis (i.e., no earnings management) is not rejected although it is false.
– **Model Specification.** A well specified model does not incorrectly detect earnings management. In statistical terms, the well specified model does not (incorrectly) reject the null of no earnings management and thus minimizes Type I error.9

Since “true” earnings management is not observable, accuracy tests are based on simulations. To choose the optimal accrual model for my analyses, I use a similar simulation approach as applied in Kothari et al. (2005). To test for model power, I artificially induce accrual-based earnings management to a subset of 50 randomly chosen firms from the overall sample by increasing or decreasing total accruals with a specified amount. Adjustments of total accruals thereby vary between −8% and +8% of total assets. I then calculate discretionary accruals based on the accrual model under investigation and test whether the model correctly detects the induced earnings management activity. More specifically, I apply a *t*-test to assess whether average abnormal accruals for the subset firms are significantly greater (H₀: $DACC_{mean} \leq 0$) or smaller than zero (H₀: $DACC_{mean} \geq 0$). The number of correct detections in 250 simulation runs is used to calculate the relative detection frequency. A high detection rate, ideally close to 100%, suggests a powerful accrual model. To test for model specification, the simulation procedure is carried out the same way, except for the fact that I choose a subset of 50 firms without adjusting total accruals. Since this subset is drawn randomly, a well specified model should not (incorrectly) detect earnings management. I apply *t*-tests to test for incorrect detection of income increasing (H₀: $DACC_{mean} \leq 0$) or decreasing (H₀: $DACC_{mean} \geq 0$) earnings management. The number of incorrect detections in 250 simulation runs then indicates the relative detection frequency. In contrast to power testing, models are better specified the closer their detection rates are to 0%.

A detailed description of the simulation study and related results is provided in Appendix A.3. For the eleven accrual models under investigation, the Lagged Model has the best properties in terms of specification: Income increasing (decreasing) earnings management is incorrectly detected in only 5.2% (2.0%) of all cases (see the results in Table A.3.1 on p. 202). Considering model power in detecting income increasing earnings management, the Adapted Model outperforms all remaining models for every adjustment level and thus seems to be the best choice (see the results in Table A.3.2 on p. 203). For the detection of income decreasing earnings management, however, the Lagged Model is most powerful. For this study, I prefer a conservative measure of earnings management to avoid erroneous inferences about managerial discretion. With the lowest (incorrect) detection frequency among all tested models, the Lagged Model pro-

---

9Type I errors arise when the null hypothesis (i.e., no earnings management) is rejected although it is true.
vides the most conservative measure of earnings management. Furthermore, the Lagged Model is above average in terms of detection power for all positive adjustment levels. Accordingly, I consider the Lagged Model as the optimal accrual model for my setting. Figure 7.3 illustrates the accuracy of the chosen model.

![Graph](image)

**FIG. 7.3.** — **Accuracy of the Lagged Model.** The graph illustrates power and specification of the Lagged Model (Dechow et al., 2003). Adjustment in % of total assets is the artificial adjustment to total accruals and related accounts to simulate earnings management activity. The relative rejection frequency is the number of H0-rejections (i.e., $DACC_{\text{mean}} \leq 0$) divided by the number of simulation runs. The graph is based on the simulation study summarized in Appendix A.3.

It is important to note that my simulation results are based on random sampling. If, however, achieving benchmarks is correlated with firm characteristics known to affect discretionary accruals (i.e., non-random samples), then some discretionary accrual models may suffer in terms of specification. For times series models, Dechow et al. (1995) show that the Jones Model and the modified Jones Model are severely misspecified when earnings management incentives are correlated with firm performance. Similar misspecifications are reported by Kothari et al. (2005), who test several cross-sectional models on stratified-random samples. Kothari et al. (2005) find performance-adjusted and performance-matched accrual models to be most reliable in terms of specification when earnings management incentives and firm performance are correlated. To ensure that my results are not spurious due to omitted firm performance, I check the robustness of my results by repeating the analyses with performance-adjusted and performance-matched discretionary accruals in Section 7.4.4.3.
7.2.1.2 Discretionary Revenues

As an alternative to measures of aggregate discretionary accruals, Stubben (2010) develops a model aimed on identifying discretion in revenue recognition (i.e., specific accruals management).10 His approach is based on the idea that reported revenues \( \text{REV}_{it} \) consist of an unmanaged \( \text{REV}^U_{it} \) and managed part \( \text{REV}^M_{it} \), whereas a fraction \( c \) of unmanaged revenues are uncollected at year-end, and, naturally, managed revenues remain completely uncollected. Assuming that all receivables outstanding are collected within one year, trade accounts receivable at year-end can be expressed as

\[
\text{REC}_{it} = c \times \text{REV}^U_{it} + \text{REV}^M_{it}.
\] (7.2.10)

Substituting \( \text{REV}^U_{it} \) by \( \text{REV}_{it} - \text{REV}^M_{it} \), rearranging terms, and first differencing then gives

\[
\Delta \text{REC}_{it} = c \times \Delta \text{REV}_{it} + (1 - c) \times \text{REV}^M_{it}.
\] (7.2.11)

Transferred into a regression equation, Eq. (7.2.11) can be expressed as

\[
\Delta \text{REC}_{it} = \alpha_1 + \beta_1 \Delta \text{REV}_{it} + \varepsilon_{it}.
\] (7.2.12)

The residual of Eq. (7.2.12) is the managed component of reported revenues multiplied with the fraction of revenues collected at year-end (i.e., \( 1 - c \)). Since \( c \) is, per definition, smaller than or equal to one, the residual may be regarded as a conservative proxy for premature revenue recognition. Eq. (7.2.12) is the most basic version of Stubben’s (2010) discretionary revenue model.

In my analyses, I run two modified versions of the basic discretionary revenue model in Eq. (7.2.12) and use their residuals as proxies for discretionary revenue recognition \( \text{REV}_1' \) and \( \text{REV}_2' \). The first version acknowledges that the basic model treats revenues from early in the year the same as revenues from later in the year. However, early revenues are usually collected

---

10This section summarizes the methodology developed by Stubben (2010) to measure discretionary revenue recognition. For more details on model development and application refer to Stubben (2010), particularly pp. 700-702.
at year-end, while a larger fraction of late revenues will remain uncollected. Consequently, Stubben (2010) allows factor $c$ to vary for early and late revenues during the year and specifies Eq. (7.2.12) with regard to revenue timing as\textsuperscript{11}

$$\Delta \text{REC}_{i,t} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{\Delta \text{REV}_{Q1-3}^{i,t}}{TA_{i,t-1}} + \beta_2 \frac{\Delta \text{REV}_{Q4}^{i,t}}{TA_{i,t-1}} + \epsilon_{it},$$

(7.2.13)

where $\Delta \text{REV}_{Q1-3}^{i,t}$ denotes reported revenues (WS: 01001) of the first three quarters in $t$ minus the first three quarters’ revenues in $t-1$ and $\Delta \text{REV}_{Q4}^{i,t}$ the difference of fourth quarter revenues in $t$ and fourth quarter revenues in $t-1$. $TA_{i,t-1}$ is lagged total assets (WS: 18184 + 02999).

The second model proposed by Stubben (2010) acknowledges that changes in receivables depend on a firm’s financial strength, its operational performance (relative to industry peers), and its stage in the business cycle. To control for these factors, he uses firm size as proxy for the firm’s financial strength, firm size and age as indicators of a firm’s business cycle, and revenue growth and gross profit margin as surrogates for operational performance. Interacting these controls with changes in revenues leads to the following adapted version of the basic model:

$$\Delta \text{REC}_{i,t} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{\text{REV}_{it}}{TA_{i,t-1}} + \sum_{k=1}^{K} \gamma_k Z_{kit} \times \frac{\Delta \text{REV}_{it}}{TA_{i,t-1}} + \epsilon_{it},$$

(7.2.14)

where all variables are defined as before and $Z_{kit}$ describes a set of the following control variables:\textsuperscript{12}

- $\text{SIZE}_{it}$ is the natural logarithm of total assets (WS: 18184 + 02999) of firm $i$ in period $t$,
- $\text{FAGE}_{it}$ represents firm age measured as $t$ minus the year firm $i$ was founded (WS: 18272),
- $\text{FAGE}_{SQ}^{it}$ represents squared firm age to control for potential non-linearities in the

\textsuperscript{11}All variables (including the intercept) are scaled by lagged total assets ($TA_{i,t-1}$) to account for size differences and heteroskedasticity. Stubben (2010) deflates variables by average total assets. To be consistent with deflation in the other models of this study, I choose lagged total assets instead.

\textsuperscript{12}Industry-median-adjusted variables are derived by calculating the median of all observations in the same industry-year combination and subsequently taking the difference of the respective actual value and the industry median.
relation of age and a firm’s credit policy,

\[ \text{GRR}_{P_{it}} \]

is the industry-median-adjusted growth rate in revenues (WS: 18272) if the growth rate is positive and 0 otherwise,

\[ \text{GRR}_{N_{it}} \]

is the industry-median-adjusted growth rate in revenues (WS: 08631) if the growth rate is negative and 0 otherwise,

\[ \text{GRM}_{it} \]

is industry-median-adjusted gross margin (WS: 08306), and

\[ \text{GRM}_{SQ_{it}} \]

denotes squared industry-median-adjusted gross margin to control for potential non-linearities in the relation of gross margin and a firm’s credit policy.

The proxy variables \( \text{REVI}^\prime \) and \( \text{REV2}^\prime \) are derived as the residuals from estimating Eq. (7.2.13) and Eq. (7.2.14), respectively, for all firm-years in the same industry-year combination (two-digit SIC codes) with at least eight observations. High values of \( \text{REVI}^\prime \) and \( \text{REV2}^\prime \) indicate that firms engage in revenue manipulations.

### 7.2.2 Proxies for Real Earnings Management

Measuring real earnings management depends on the researcher’s ability to detect discretionary business activities. Therefore, measures of real earnings management require expectation models that effectively identify normal business patterns and allow an extraction of abnormal activities. As an example, consider the case of overproduction. Due to fixed production costs, increasing production volume decreases cost of goods sold (COGS) and inflates earnings. To identify overproduction, one requires an estimate of the normal production level (i.e., the production level absent discretion). Reliable evidence of earnings management by means of overproduction, sales manipulation, cost cutting, or sale of fixed assets thus depends on the accuracy of the underlying expectation model. Compared to the broad field of accrual models, only a few models have been developed in the area of real earnings management. These include, e.g., Berger (1993), Perry and Grinaker (1994), and the more recent models in Gunny (2010) and Roychowdhury (2004, 2006).\(^{13}\)

\(^{13}\)Recent applications of the models proposed by Roychowdhury (2004, 2006) and Gunny (2010) can be found in, e.g., Lin et al. (2006), Athanasakou et al. (2007), Zang (2007), Cohen et al. (2008), Chen et al. (2010), Cohen et al. (2010a) and McInnis and Collins (2011).
7.2.2.1 The Roychowdhury Model

Roychowdhury (2004, 2006) analyzes whether managers manipulate real activities to avoid reporting an annual loss. Real earnings management activities thereby include discretionary expenditures, overproduction, and sales manipulation. Inferring conclusions on these activities from empirical data requires assumptions about their financial statement effects:

- **Expense Cutting.** Reducing discretionary expenses inflates current period earnings and leads to discretionary expenses below the expected level. If these expenses are not generally paid on credit, cutting expenses will result in higher cash flow from operating activities (CFO) than expected under conditions absent earnings management.

- **Overproduction.** Overproduction results in production costs above their expected level. Although overproduction decreases COGS, production costs as sum of COGS and change in inventory are supposed to rise. Furthermore, increased inventories tie up cash and thus CFO falls below its expected level.

- **Sales Manipulations.** Boosting sales by discounts or more lenient credit terms requires a higher production volume and thereby increases production costs. Additionally, discounts and more lenient credit terms reduce CFO relative to sales.

Testing for real earnings management now requires valid models to estimate the normal levels of CFO, production costs, and discretionary expenses. Building on the theoretical work of Dechow et al. (1998), Roychowdhury (2004, 2006) develops the following cross-sectional expectation models:

**Expected Level of CFO**

\[
\frac{CFO_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{REV_{it}}{TA_{i,t-1}} + \beta_2 \frac{\Delta REV_{it}}{TA_{i,t-1}} + \epsilon_{it},
\]  

(7.2.15)

where \(CFO_{it}\) denotes cash flow from operating activities of firm \(i\) in \(t\), \(REV_{it}\) stands for sales, and \(\Delta REV_{it}\) is the change in sales from \(t-1\) to \(t\). All variables are deflated by lagged total assets \(TA_{i,t-1}\) to account for firm heterogeneity and avoid heteroskedasticity.
Expected Level of Production Costs

\[
\frac{\text{PROD}1_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{\text{REV}_{it}}{TA_{i,t-1}} + \beta_2 \frac{\Delta \text{REV}_{it}}{TA_{i,t-1}} + \beta_3 \frac{\Delta \text{REV}_{i,t-1}}{TA_{i,t-1}} + \epsilon_{it}, \tag{7.2.16}
\]

where \(\text{PROD}1_{it}\) is the sum of COGS (WS: 01051) and change in inventories (WS: 02101) and \(\Delta \text{REV}_{i,t-1}\) is the change in sales from \(t-2\) to \(t-1\). All other variables are defined as before.

Expected Level of Discretionary Expenses

\[
\frac{\text{DEXP}_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{\text{REV}_{i,t-1}}{TA_{i,t-1}} + \epsilon_{it}, \tag{7.2.17}
\]

where \(\text{DEXP}_{it}\) is the sum of SG&A (WS: 01101) and R&D costs (WS: 01201). All other variables are defined as before.\(^{14}\)

A significant drawback of Roychowdhury’s approach, is the ambiguous interpretation of abnormal CFO. As Gunny (2010) notes, increasing earnings by reductions of discretionary expenses (usually paid in cash) incurs abnormally high CFO at year-end.\(^{15}\) Conversely, increasing earnings by means of overproduction results in lower CFO than expected. To circumvent the conflictive interpretation of abnormal CFO, I solely focus on abnormal production costs and discretionary expenses. To do so, I estimate Eq. (7.2.16) and Eq. (7.2.17) in cross-sectional regressions for each industry-year combination (two-digit SIC code). I require a minimum of eight observations per industry-year to ensure reliable coefficient estimates. The residuals from these regressions are used as proxies for abnormal production (\(\text{PROD}1'\)) and abnormal discretionary expenses (\(\text{DEXP}'\)). To ease the interpretation of the results, the cost cutting proxy \(\text{DEXP}'\) is multiplied with \(-1\) (i.e., positive values indicate income increasing earnings management).

\(^{14}\) Roychowdhury (2004, 2006) includes advertising costs (gathered from Compustat) as a separate component of \(\text{DEXP}_{it}\). I do not include advertising costs separately since they are already included in the Worldscope data item for SG&A (WS: 01101).

\(^{15}\) See her footnote 2 on p. 856 for this argument.
7. Management’s Techniques to Achieve Benchmarks

7.2.2.2 The Gunny Model

The models developed by Gunny (2010) measure four different real earnings management activities: discretionary decreases of R&D costs, discretionary decreases of SG&A expenses, the timing of fixed assets/investment sales, and overproduction. Her approach differs from Roychowdhury (2004, 2006) in three aspects: First, she does not measure abnormal CFO due to its ambiguous role as real earnings management proxy. Second, she examines the timing of fixed assets and/or investments as additional real earnings management device. Eventually, she augments Roychowdhury’s expectation models by additional independent variables to control for firm characteristics. Her models are specified as follows:

\[ \text{Expected Level of R&D Expenses} \]

The expected level of R&D expenses is assumed to be a function of prior year’s R&D costs, firm size, internal funds available for R&D spending, and Tobin’s Q as proxy for the marginal benefit to marginal cost of an additional unit of a new investment:

\[ \frac{RND_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{RND_{i,t-1}}{TA_{i,t-1}} \]
\[ + \beta_2 \frac{INT_{it}}{TA_{i,t-1}} + \beta_3 SIZE_{it} + \beta_4 Q_{it} + \epsilon_{it}, \]  

(7.2.18)

where \( RND_{it} \) and \( RND_{i,t-1} \) are current and lagged R&D expenses (WS: 01201), respectively, \( INT_{it} \) is internal funds available calculated as net income (WS: 01751) plus R&D expenses (WS: 01201) and depreciation (WS: 01151), \( SIZE_{it} \) is calculated as the natural logarithm of market value (DS: MV), and \( Q_{it} \) is defined as the ratio of market capitalization (DS: MV) plus total assets (WS: 18184 + 02999) minus total equity (WS: 03426 + 03501) to total assets. \( RND_{it} \), \( RND_{i,t-1} \), and \( INT_{it} \) are deflated by lagged total assets to address heteroskedasticity.

\[ \text{Expected Level of SG&A Expenses} \]

Expected SG&A spending (\( SGA_{it}, \) WS: 01101) is assumed to depend on the change in revenues (\( \Delta REV_{it}, \) WS: 01001), size (\( SIZE_{it} \)), Tobin’s Q (\( Q_{it} \)), and an additional interaction term (\( DD_{it} \)) controlling for “sticky” cost behavior (i.e., the relative decrease of SG&A expenses in times of shrinking demand is smaller than the relative increase in these expenses when demand is
rising):\(^{16}\)

\[
\frac{SGA_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{\Delta REV_{it}}{TA_{i,t-1}} + \beta_2 \frac{\Delta REV_{it}}{TA_{i,t-1}} \times DD_{it}
\]

\[
+ \beta_3 \frac{INT_{it}}{TA_{i,t-1}} + \beta_4 SIZE_{it} + \beta_5 Q_{it} + \epsilon_{it},
\]

(7.2.19)

where \(INT_{it}, SIZE_{it}\) and \(Q_{it}\) are derived as in Eq. (7.2.18) and \(DD_{it}\) is a dummy variable for sales decreases set to 1 if \(\Delta REV < 0\) and 0 otherwise.

**Expected Level of Production Costs**

Gunny’s expectation model for production costs \((PROD2_{it})\) is identical to Roychowdhury’s version except for the additional regressors size \((SIZE_{it})\), and Tobin’s Q \((Q_{it})\). Hence, production costs are assumed to depend on the sales level in \(t\) \((REV_{it})\), change in sales \((\Delta REV_{it})\), and lagged change in sales \((\Delta REVi, t-1)\):

\[
\frac{PROD2_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{REV_{it}}{TA_{i,t-1}} + \beta_2 \frac{\Delta REV_{it}}{TA_{i,t-1}}
\]

\[
+ \beta_3 \frac{\Delta REVi, t-1}{TA_{i,t-1}} + \beta_4 SIZE_{it} + \beta_5 Q_{it} + \epsilon_{it},
\]

(7.2.20)

where \(PROD2_{it}\) is the sum of COGS (WS: 01051) and change in inventories (WS: 02101) and all other variables are defined as before.

**Expected Gains and Losses from Selling Fixed Assets**

Gunny (2010) assumes that gains and losses from fixed assets sales are a function of the disposal proceeds \((DISP_{it})\), internal funds available \((INT_{it})\), size \((SIZE_{it})\) and Tobin’s Q \((Q_{it})\):\(^{17}\)

---

\(^{16}\)The term “sticky” costs is adopted from Anderson et al. (2003), who present empirical evidence of an asymmetric response of SG&A costs in times of increasing and decreasing demand. Specifically, they regress SG&A expense on changes in revenues and changes in revenues interacted with a dummy variable set to 1 if sales are decreasing and 0 otherwise. A significantly negative coefficient on the interaction term supports their conjecture of “sticky” cost behavior.

\(^{17}\)The original model also includes gains and losses from investment sales. Due to data limitations, my model only includes gains and losses from selling fixed assets.
where \( D\text{GAIN}_{it} \) is gains and losses from fixed asset sales (WS: 01306). Proceeds from these sales (\( D\text{ISP}_{it}, \) WS: 04351) are multiplied with \(-1\) when the respective asset was sold with a loss (i.e., \( D\text{GAIN}_{it} < 0 \)).\(^{18}\) All other variables are defined as before.

The residuals \( RND', SGA', PROD2', \) and \( D\text{GAIN}' \) are derived by estimating the respective expectation models (Eq. (7.2.18) to Eq. (7.2.21)) for every industry-year combination (two-digit SIC codes) with at least eight observations. Since negative values of the cost cutting proxies indicate income increasing earnings management, I multiply \( RND' \) and \( SGA' \) with \(-1\) to ease the interpretation of the results (i.e., positive values indicate income increasing earnings management).

### 7.2.3 Proxy for Expectations Management

When the target of interest is the latest analyst consensus forecast, an appealing alternative to earnings management is the active guidance of financial analysts. To measure expectations management, I follow Matsumoto’s (2002) approach and compare the period’s actual consensus forecast with an expected forecast. Firms are assumed to engage in expectations management when the actual analyst forecast is below its expected value. Since expected forecasts are unknown, I use the fitted values of a simple regression model to estimate the expected change in EPS in a first step.\(^{19}\) The regression model is given by Eq. (7.2.22):

\[
\frac{D\text{GAIN}_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{D\text{ISP}_{it}}{TA_{i,t-1}} + \beta_2 \frac{INT_{it}}{TA_{i,t-1}} + \beta_3 S\text{IZE}_{it} + \beta_4 Q_{it} + \epsilon_{it},
\]

(7.2.21)

\(^{18}\)Multiplication with \(-1\) is needed to ensure that the relationship between \( D\text{GAIN}_{it} \) and \( D\text{ISP}_{it} \) is monotonic. See Gunny (2010, p. 865) for details.

\(^{19}\)The original model is estimated on a quarterly basis and thus more accurate than my annual specification. Burgstahler and Eames (2006) calculate Matsumoto’s proxy on annual basis by estimating the expected change in earnings for the fourth quarter and subsequently add EPS realizations of the first three quarters. As a matter of fact, my model is less accurate and a more noisy measure of forecast guidance. However, due to a limited number of quarterly EPS figures available for the German market, I am not able to follow this approach.
\[
\frac{\Delta EPS_{it}}{P_{i,t-1}} = \alpha_1 + \beta_1 \frac{\Delta EPS_{i,t-1}}{P_{i,t-2}} + \beta_2 RET_{it} + \varepsilon_{it}, \tag{7.2.22}
\]

where

\(\Delta EPS_{it/i,t-1}\) denote changes in EPS from \(t-1\) to \(t\) and \(t-2\) to \(t-1\), respectively,

\(P_{i,t-1/i,t-2}\) is the closing share price of year \(t-1\) and \(t-2\), respectively, and

\(RET_{it}\) is monthly market-adjusted return cumulated from the first month after year \(t-1\)’s earnings announcement until the month preceding year \(t\)’s earnings release.

All EPS variables are gathered from the I/B/E/S database to ensure comparability with the related forecasts. Share prices are retrieved from Datastream (DS: UP). Market-adjusted excess returns are the monthly returns of the respective firm (WS: 08807) minus the monthly return of the CDAX index. Monthly CDAX returns are calculated based on the Datastream index history (DS: PI). Eq. (7.2.22) is estimated for each industry-year combination with at least eight observations. The fitted values from these industry-year regressions are then multiplied with \(P_{i,t-1}\) to derive the expected change in EPS:

\[
E(\Delta EPS_{it}) = \hat{\alpha}_1 + \hat{\beta}_1 \frac{\Delta EPS_{i,t-1}}{P_{i,t-2}} + \hat{\beta}_2 RET_{it} \times P_{i,t-1}. \tag{7.2.23}
\]

Adding the expected change in EPS to prior year’s EPS then gives the expected final consensus forecast for period \(t\) absent any managerial guidance:

\[
E(FC_{it}) = EPS_{i,t-1} + E(\Delta EPS_{it}). \tag{7.2.24}
\]

Eventually, the unexpected consensus forecast \((FC'_{it})\) is calculated as the difference of the actual forecast \(FC_{it}\) and \(E(FC_{it})\) from Eq. (7.2.24). Since \(FC'_{it}\) is the deviation of the actual consensus forecast from the expected consensus forecast, negative values of \(FC'_{it}\) suggest that firms guide analysts down to beatable earnings targets. To ease the interpretation of the results, I multiply \(FC'\) with \(-1\) so that positive values indicate downward earnings guidance.
7.3 Sample Selection and Descriptive Statistics

7.3.1 Sample Selection

The samples for the tests of specific earnings management techniques are based on the extended samples used in Chapter 6. The sample selection procedure for the extended samples are summarized in Section 6.3.1. The initial sample covers 1,263 observations for the analyst forecast, 2,100 observations for the zero earnings, and 2,099 observations for the earnings changes benchmark. In a first step, I drop all firm-years with missing data to calculate the respective earnings management measures. Subsequently, I exclude all observations that belong to industry-year combinations with less than eight available firm-years to ensure valid estimates of the proxy variables. Eventually, I account for influential observations by excluding the (yearly) upper and lower 1% of the variables required to calculate the respective earnings management and guidance measures. The sample selection procedure is summarized in Table 7.1.

I lose a large number of firm-years due to insufficient observations per industry-year combination and missing data. Data availability varies considerably for the respective earnings management proxies. The highest number of firm-years is lost for the $DGAIN'$ and $RND'$ samples, since both lack the required database entries for approximately 50% of the initial sample. As a result, the final sample sizes vary considerably for different earnings management proxies: For the analyst forecast benchmark (Panel A), the final sample contains between 361 ($DGAIN'$ sample) and 703 ($REV'$ sample) firm-years. In terms of German non-financial market capitalization, the sample coverage varies between 12.5% ($REV'$ sample) and 43.0% ($FC'$ sample). The zero earnings benchmark sample (Panel B) is significantly larger with between 680 ($DGAIN'$ sample) and 1,368 ($PROD'$ sample) firm-years, representing between 29.7% ($REV'$ sample) and 65.5% ($PROD'$ sample) of total non-financial market value. For the earnings changes benchmark (Panel C), the final sample has between 692 ($DGAIN'$ sample) and 1,417 firm-years ($PROD'$ sample) and covers between 28.5% ($REV'$ sample) and 67.1% ($DACC$ sample) of overall non-financial market value.

7.3.2 Descriptive Statistics

Table 7.2 provides summary statistics for the earnings management proxies and control variables used in the analysis of earnings management techniques. For the analyst forecast benchmark
sample (Panel A), mean earnings management proxies are close to zero and vary between $-0.001$ ($DACC$) and $0.005$ ($PROD2'$). As the only undeflated variable, the earnings guidance proxy ($FC'$) is considerably larger with a mean of $0.032$. None of these proxy means is significantly different from zero at generally accepted levels. The summary statistics for the zero earnings benchmark sample in Panel B draw a similar picture. All earnings management proxy means are close to zero and vary between $-0.002$ ($PROD2'$) and $0.006$ ($SGA'$). None of these means is significantly different from zero. Eventually, the proxy means in the earnings changes benchmark sample (Panel C) range between $-0.005$ ($PROD2'$) and $0.004$ ($SGA'$) and are also not significantly different from zero. Median values are generally close to mean values except for $FC'$ and $DEXP'$.

Table 7.3 presents correlations between the variables and provides some preliminary insights on earnings management techniques. In every panel of Table 7.3, the bottom left values are Pearson and the top right are Spearman rank correlations. Unless these correlations differ significantly, I henceforth refer to Pearson correlations. The correlations in the first column of Panel A indicate a positive relation between just achieving the analyst forecast benchmark and discretionary revenues. All other correlations between proxy variables and the $SMBE$ dummy lack significance; a finding inconsistent with strong managerial intervention to beat the analyst consensus. The results in Panel B of Table 7.3 indicate that just achieving the zero earnings benchmark is positively related with discretionary accruals ($DACC$), SG&A expense cutting ($SGA'$), and overproduction ($PROD1'$). Pearson correlations further suggest a positive correlation of $SMBE$ with discretionary expense cutting ($DEXP'$) and discretionary revenues ($REV2'$). These results, however, do not hold for Spearman rank correlations. Overall, the correlations suggest earnings management activity around the zero earnings benchmark. Panel C of Table 7.3 presents correlations for the earnings changes benchmark sample. Positive correlations between achieving last year’s EPS and $SGA'$ suggest that firms engage in cost cutting to avoid missing the benchmark. Overproduction ($PROD2'$) is negatively related to benchmark beating. This indicates a potential measurement error of the overproduction proxy. The remaining correlations between the proxy variables and the $SMBE$ dummy lack significance for at least one correlation measure. In sum, the results do not provide strong support for the earnings management hypothesis. The remaining correlations in Panel A to C reveal two additional findings that deserve consideration: First, there are high and nearly exclusively positive correlations between the proxy variables within the accrual-based and the real earnings management group. This suggests that firms do not limit discretion to one single type of accrual-based or real earnings management. Second, firm performance ($ROA$) is highly correlated with nearly all of the proxy variables. This finding underpins the importance of controlling for $ROA$ in the regression analysis.
The table summarizes the sample selection procedure for the earnings management techniques tests. \( FC' \) is the proxy variable for earnings guidance (Matsumoto, 2002), \( DACC \) for aggregate accruals management (Jones, 1991; Dechow et al., 2003), \( REV1' \) and \( REV2' \) for premature revenue recognition (Stubben, 2010), \( DEXP' \), \( RND' \), and \( SGA' \) for discretionary cost cutting, \( PROD1' \) and \( PROD2' \) for overproduction, and \( DGAIN' \) for abnormal gains from selling fixed assets (Roychowdhury, 2004, 2006; Gunny, 2010). Market capitalization data for all German non-financial firms is gathered from summary statistics of the Deutsche Bundesbank (Deutsche Bundesbank, 2010, p. 45).

<table>
<thead>
<tr>
<th>Table 7.1</th>
<th>Sample Selection Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FC'</td>
</tr>
<tr>
<td>Panel A: Analyst Forecast Benchmark</td>
<td></td>
</tr>
<tr>
<td>Extended Sample</td>
<td>1,263</td>
</tr>
<tr>
<td>less: Missing Data</td>
<td>163</td>
</tr>
<tr>
<td>less: &lt; 8 Obs./Outliers</td>
<td>410</td>
</tr>
<tr>
<td>Final Sample</td>
<td>690</td>
</tr>
<tr>
<td>% of Non-fin. Firms’ Market Value</td>
<td>43.0</td>
</tr>
<tr>
<td>Panel B: Zero Earnings Benchmark</td>
<td></td>
</tr>
<tr>
<td>Extended Sample</td>
<td>–</td>
</tr>
<tr>
<td>less: Missing Data</td>
<td>–</td>
</tr>
<tr>
<td>less: &lt; 8 Obs./Outliers</td>
<td>–</td>
</tr>
<tr>
<td>Final Sample</td>
<td>–</td>
</tr>
<tr>
<td>% of Non-fin. Firms’ Market Value</td>
<td>–</td>
</tr>
<tr>
<td>Panel C: Earnings Changes Benchmark</td>
<td></td>
</tr>
<tr>
<td>Extended Sample</td>
<td>–</td>
</tr>
<tr>
<td>less: Missing Data</td>
<td>–</td>
</tr>
<tr>
<td>less: &lt; 8 Obs./Outliers</td>
<td>–</td>
</tr>
<tr>
<td>Final Sample</td>
<td>–</td>
</tr>
<tr>
<td>% of Non-fin. Firms’ Market Value</td>
<td>–</td>
</tr>
</tbody>
</table>
### Table 7.2

**Variable Distributions**

<table>
<thead>
<tr>
<th></th>
<th>FC'</th>
<th>DACC</th>
<th>REV1'</th>
<th>REV2'</th>
<th>DEXP</th>
<th>RND'</th>
<th>SGA'</th>
<th>PROD1'</th>
<th>PROD2'</th>
<th>DGAIN'</th>
<th>SIZE</th>
<th>MTB</th>
<th>ROA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Analyst Forecast Benchmark</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>690</td>
<td>682</td>
<td>703</td>
<td>404</td>
<td>686</td>
<td>475</td>
<td>644</td>
<td>686</td>
<td>643</td>
<td>361</td>
<td>1,262</td>
<td>1,262</td>
<td>1,262</td>
</tr>
<tr>
<td>Mean</td>
<td>0.032</td>
<td>-0.001</td>
<td>0.002</td>
<td>0.004</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.004</td>
<td>0.005</td>
<td>0.000</td>
<td>5.903</td>
<td>2.115</td>
<td>0.019</td>
</tr>
<tr>
<td>σ</td>
<td>1.874</td>
<td>0.077</td>
<td>0.055</td>
<td>0.071</td>
<td>0.188</td>
<td>0.021</td>
<td>0.173</td>
<td>0.200</td>
<td>0.219</td>
<td>0.011</td>
<td>2.067</td>
<td>2.117</td>
<td>0.139</td>
</tr>
<tr>
<td>25%</td>
<td>-0.374</td>
<td>-0.041</td>
<td>-0.025</td>
<td>-0.017</td>
<td>-0.090</td>
<td>-0.008</td>
<td>-0.061</td>
<td>-0.084</td>
<td>-0.081</td>
<td>-0.003</td>
<td>4.293</td>
<td>1.035</td>
<td>0.010</td>
</tr>
<tr>
<td>50%</td>
<td>-0.039</td>
<td>-0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.025</td>
<td>0.000</td>
<td>0.006</td>
<td>0.008</td>
<td>-0.001</td>
<td>-0.001</td>
<td>5.467</td>
<td>1.628</td>
<td>0.040</td>
</tr>
<tr>
<td>75%</td>
<td>0.327</td>
<td>0.038</td>
<td>0.029</td>
<td>0.013</td>
<td>0.122</td>
<td>0.009</td>
<td>0.083</td>
<td>0.106</td>
<td>0.095</td>
<td>0.001</td>
<td>7.157</td>
<td>2.574</td>
<td>0.072</td>
</tr>
</tbody>
</table>

|       |     |      |       |       |      |      |      |        |        |        |      |     |     |
| **Panel B: Zero Earnings Benchmark** |     |      |       |       |      |      |      |        |        |        |      |     |     |
| N    | –   | 1,320 | 1,264 | 738   | 1,345 | 686  | 1,222 | 1,368  | 1,276  | 680    | 2,100 | 2,098 | 2,100|
| Mean | –   | 0.004 | 0.001 | 0.002 | 0.004 | 0.000 | 0.006 | 0.005  | -0.002 | 0.000  | 5.237 | 2.333 | -0.024|
| σ    | –   | 0.084 | 0.063 | 0.056 | 0.192 | 0.021 | 0.181 | 0.209  | 0.208  | 0.014  | 2.177 | 11.247 | 0.738|
| 25%  | –   | -0.035 | -0.027 | -0.020 | -0.086 | -0.008 | -0.061 | -0.083 | -0.079 | -0.002 | 3.795 | 0.933 | -0.016|
| 50%  | –   | 0.004 | 0.002 | 0.000 | 0.029 | 0.001 | 0.007 | 0.008  | 0.002  | 0.000  | 4.927 | 1.493 | 0.031 |
| 75%  | –   | 0.043 | 0.029 | 0.020 | 0.122 | 0.009 | 0.082 | 0.110  | 0.090  | 0.002  | 6.408 | 2.498 | 0.065 |

|       |     |      |       |       |      |      |      |        |        |        |      |     |     |
| **Panel C: Earnings Changes Benchmark** |     |      |       |       |      |      |      |        |        |        |      |     |     |
| N    | –   | 1,379 | 1,272 | 752   | 1,376 | 695  | 1,270 | 1,417  | 1,305  | 692    | 2,099 | 2,097 | 2,099|
| Mean | –   | 0.001 | 0.001 | 0.003 | 0.001 | 0.000 | 0.004 | -0.001 | -0.005 | 0.000  | 5.244 | 2.322 | -0.027|
| σ    | –   | 0.082 | 0.058 | 0.068 | 0.189 | 0.020 | 0.157 | 0.202  | 0.194  | 0.015  | 2.174 | 11.236 | 0.742|
| 25%  | –   | -0.038 | -0.026 | -0.021 | -0.086 | -0.008 | -0.055 | -0.091 | -0.083 | -0.002 | 3.793 | 0.933 | -0.015|
| 50%  | –   | 0.003 | 0.002 | -0.001 | 0.026 | 0.001 | 0.008 | 0.008  | 0.001  | 0.000  | 4.927 | 1.491 | 0.031 |
| 75%  | –   | 0.041 | 0.029 | 0.017 | 0.119 | 0.008 | 0.081 | 0.097  | 0.083  | 0.002  | 6.413 | 2.498 | 0.066 |

Panels A to C depict summary statistics of the variables required to test for earnings management techniques. FC' is the proxy variable for earnings guidance (Matsumoto, 2002), DACC for aggregate accruals management (Jones, 1991; Dechow et al., 2003), REV' and REV2' for premature revenue recognition (Stubben, 2010), DEXP', RND', and SGA' for discretionary cost cutting, PROD1' and PROD2' for overproduction, and DGAIN' for abnormal gains from selling fixed assets (Roychowdhury, 2004, 2006; Guny, 2010). SIZE is firm size measured as natural logarithm of total assets, MTB is the market value of equity divided by its book value, and ROA is return on (total) assets.
### Table 7.3

**Panel A: Analyst Forecast Benchmark**

<table>
<thead>
<tr>
<th>Variable</th>
<th>SMBE</th>
<th>SMISS</th>
<th>FC</th>
<th>DACC</th>
<th>REV1</th>
<th>REV2</th>
<th>DEXP</th>
<th>RND</th>
<th>SGA</th>
<th>PROD1</th>
<th>PROD2</th>
<th>DGAIN</th>
<th>SIZE</th>
<th>MTB</th>
<th>ROA</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMBE</td>
<td>0.15</td>
<td>0.01</td>
<td>0.04</td>
<td>0.08</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.05</td>
<td>0.03</td>
<td>-0.06</td>
<td>-0.12</td>
<td>0.02</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>SMISS</td>
<td>-0.15</td>
<td>0.02</td>
<td>0.05</td>
<td>0.00</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.04</td>
<td>-0.05</td>
<td>0.05</td>
<td>-0.09</td>
<td>0.04</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FC</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.03</td>
<td>0.04</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.07</td>
<td>-0.05</td>
<td>0.02</td>
<td>-0.05</td>
<td>0.05</td>
<td>0.08</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>DACC</td>
<td>-0.03</td>
<td>0.02</td>
<td>0.04</td>
<td>0.22</td>
<td>0.19</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
<td>0.06</td>
<td>0.10</td>
<td>0.08</td>
<td>0.02</td>
<td>-0.01</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>REV1</td>
<td>0.08</td>
<td>0.05</td>
<td>0.03</td>
<td>0.26</td>
<td>0.46</td>
<td>0.03</td>
<td>-0.05</td>
<td>0.04</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.08</td>
<td>0.03</td>
<td>0.06</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>REV2</td>
<td>0.10</td>
<td>-0.03</td>
<td>0.02</td>
<td>0.16</td>
<td>0.33</td>
<td>-0.10</td>
<td>-0.02</td>
<td>-0.14</td>
<td>-0.15</td>
<td>-0.15</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.10</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>DEXP</td>
<td>0.03</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.07</td>
<td>0.04</td>
<td>0.01</td>
<td>0.17</td>
<td>0.65</td>
<td>0.67</td>
<td>0.60</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.07</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>RND</td>
<td>0.04</td>
<td>-0.04</td>
<td>0.06</td>
<td>0.04</td>
<td>-0.09</td>
<td>0.13</td>
<td>0.22</td>
<td>0.13</td>
<td>0.05</td>
<td>0.01</td>
<td>-0.01</td>
<td>-0.04</td>
<td>0.01</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>SGA</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.08</td>
<td>0.04</td>
<td>0.03</td>
<td>-0.26</td>
<td>0.64</td>
<td>0.19</td>
<td>0.50</td>
<td>0.55</td>
<td>0.08</td>
<td>0.03</td>
<td>0.01</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>PROD1</td>
<td>0.04</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.09</td>
<td>-0.01</td>
<td>-0.07</td>
<td>0.69</td>
<td>0.06</td>
<td>0.47</td>
<td>0.83</td>
<td>0.02</td>
<td>-0.05</td>
<td>-0.12</td>
<td>-0.28</td>
<td></td>
</tr>
<tr>
<td>PROD2</td>
<td>0.04</td>
<td>-0.04</td>
<td>0.03</td>
<td>0.20</td>
<td>0.00</td>
<td>-0.14</td>
<td>0.50</td>
<td>0.09</td>
<td>0.48</td>
<td>0.77</td>
<td>0.05</td>
<td>-0.07</td>
<td>0.06</td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td>DGAIN</td>
<td>0.10</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.01</td>
<td>0.02</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.04</td>
<td>-0.11</td>
<td>0.11</td>
<td>0.05</td>
<td>0.03</td>
<td>-0.04</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>SIZE</td>
<td>-0.13</td>
<td>-0.09</td>
<td>0.06</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.07</td>
<td>0.01</td>
<td>0.01</td>
<td>-0.01</td>
<td>-0.03</td>
<td>0.00</td>
<td>0.03</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>MTB</td>
<td>0.01</td>
<td>0.03</td>
<td>0.02</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.12</td>
<td>-0.17</td>
<td>0.00</td>
<td>-0.06</td>
<td>-0.20</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.05</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>ROA</td>
<td>0.07</td>
<td>0.02</td>
<td>-0.01</td>
<td>0.20</td>
<td>0.16</td>
<td>-0.03</td>
<td>0.04</td>
<td>0.20</td>
<td>0.07</td>
<td>-0.24</td>
<td>-0.12</td>
<td>0.03</td>
<td>0.14</td>
<td>-0.01</td>
<td></td>
</tr>
</tbody>
</table>

*Continued on next page.*
### Panel B: Zero Earnings Benchmark

<table>
<thead>
<tr>
<th></th>
<th>SMBE</th>
<th>SMISS</th>
<th>DACC</th>
<th>REV1'</th>
<th>REV2'</th>
<th>DEXP'</th>
<th>RND'</th>
<th>SGA'</th>
<th>PROD1'</th>
<th>PROD2'</th>
<th>DGAIN'</th>
<th>SIZE</th>
<th>MTB</th>
<th>ROA</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMBE</td>
<td>-0.11a</td>
<td>0.09a</td>
<td>0.03</td>
<td>0.05</td>
<td>0.04</td>
<td>0.02</td>
<td>0.05c</td>
<td>0.06b</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.24a</td>
<td>-0.07a</td>
<td>-0.04c</td>
<td></td>
</tr>
<tr>
<td>SMISS</td>
<td>-0.11a</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.08b</td>
<td>0.01</td>
<td>-0.06</td>
<td>-0.01</td>
<td>0.05c</td>
<td>0.06b</td>
<td>0.09b</td>
<td>-0.19a</td>
<td>-0.02</td>
<td>-0.26c</td>
<td></td>
</tr>
<tr>
<td>DACC</td>
<td>0.11a</td>
<td>-0.01</td>
<td>0.23a</td>
<td>0.27a</td>
<td>-0.01</td>
<td>0.11a</td>
<td>0.01</td>
<td>0.03</td>
<td>0.07c</td>
<td>0.05c</td>
<td>0.00</td>
<td>0.22c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REV1'</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.25a</td>
<td>0.57a</td>
<td>-0.02</td>
<td>-0.04</td>
<td>0.05</td>
<td>-0.10a</td>
<td>-0.08b</td>
<td>0.00</td>
<td>0.05c</td>
<td>0.09a</td>
<td>0.15c</td>
<td></td>
</tr>
<tr>
<td>REV2'</td>
<td>0.09b</td>
<td>-0.04</td>
<td>0.20a</td>
<td>0.49a</td>
<td>-0.07c</td>
<td>0.14a</td>
<td>-0.08b</td>
<td>-0.12a</td>
<td>-0.01</td>
<td>-0.07b</td>
<td>0.01</td>
<td>0.02</td>
<td>0.07c</td>
<td></td>
</tr>
<tr>
<td>DEXP'</td>
<td>0.05c</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.10b</td>
<td>0.18a</td>
<td>0.69a</td>
<td>0.59a</td>
<td>0.02</td>
<td>-0.01</td>
<td>-0.07b</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RND'</td>
<td>0.02</td>
<td>-0.04</td>
<td>0.11a</td>
<td>-0.08c</td>
<td>0.12a</td>
<td>0.18a</td>
<td>0.09b</td>
<td>0.08c</td>
<td>0.07c</td>
<td>-0.11b</td>
<td>-0.02</td>
<td>0.01</td>
<td>0.12c</td>
<td></td>
</tr>
<tr>
<td>SGA'</td>
<td>0.07b</td>
<td>-0.02</td>
<td>-0.02</td>
<td>0.05c</td>
<td>-0.12b</td>
<td>0.69a</td>
<td>0.09b</td>
<td>0.51a</td>
<td>0.54a</td>
<td>0.02</td>
<td>-0.06b</td>
<td>0.01</td>
<td>0.07b</td>
<td></td>
</tr>
<tr>
<td>PROD1'</td>
<td>0.05c</td>
<td>0.04</td>
<td>0.03</td>
<td>-0.05c</td>
<td>-0.14a</td>
<td>0.69a</td>
<td>0.08b</td>
<td>0.52a</td>
<td>0.85a</td>
<td>0.01</td>
<td>-0.08a</td>
<td>-0.12a</td>
<td>-0.29c</td>
<td></td>
</tr>
<tr>
<td>PROD2'</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.03</td>
<td>-0.03</td>
<td>-0.16a</td>
<td>0.60a</td>
<td>0.10b</td>
<td>0.59a</td>
<td>0.86a</td>
<td>-0.01</td>
<td>-0.05c</td>
<td>0.02</td>
<td>-0.25c</td>
<td></td>
</tr>
<tr>
<td>DGAIN'</td>
<td>0.04</td>
<td>0.09b</td>
<td>0.13a</td>
<td>0.01</td>
<td>-0.07</td>
<td>0.01</td>
<td>-0.12b</td>
<td>-0.03</td>
<td>0.00</td>
<td>0.02</td>
<td>-0.04</td>
<td>-0.01</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>-0.21a</td>
<td>-0.18a</td>
<td>-0.04</td>
<td>0.04</td>
<td>-0.03</td>
<td>0.02</td>
<td>-0.01</td>
<td>-0.03</td>
<td>-0.01</td>
<td>-0.02</td>
<td>0.06a</td>
<td>0.20c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTB</td>
<td>0.03</td>
<td>0.01</td>
<td>0.00</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.06b</td>
<td>0.00</td>
<td>0.02</td>
<td>-0.04c</td>
<td>0.26c</td>
<td></td>
</tr>
<tr>
<td>ROA</td>
<td>0.04c</td>
<td>-0.09a</td>
<td>0.22a</td>
<td>0.13a</td>
<td>0.05</td>
<td>0.04</td>
<td>0.08b</td>
<td>0.05c</td>
<td>-0.22a</td>
<td>-0.17a</td>
<td>0.06c</td>
<td>0.21a</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

*Continued on next page.*
### Panel C: Earnings Changes Benchmark

<table>
<thead>
<tr>
<th></th>
<th>SMBE</th>
<th>SMISS</th>
<th>DACC</th>
<th>REVI'</th>
<th>REV2'</th>
<th>DEXP'</th>
<th>RND'</th>
<th>SGA'</th>
<th>PROD1'</th>
<th>PROD2'</th>
<th>DGAIN'</th>
<th>SIZE</th>
<th>MTB</th>
<th>ROA</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMBE</td>
<td>-0.11^a</td>
<td>0.00</td>
<td>0.05^c</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.01</td>
<td>0.05^c</td>
<td>-0.02</td>
<td>-0.06^b</td>
<td>0.00</td>
<td>-0.12^a</td>
<td>0.05^b</td>
<td>0.14^b</td>
<td></td>
</tr>
<tr>
<td>SMISS</td>
<td>-0.11^a</td>
<td>0.03</td>
<td>0.04</td>
<td>0.03</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.01</td>
<td>-0.13^a</td>
<td>0.03</td>
<td>0.17^c</td>
<td></td>
</tr>
<tr>
<td>DACC</td>
<td>0.03</td>
<td>0.03</td>
<td>0.22^a</td>
<td>0.24^a</td>
<td>-0.01</td>
<td>0.12^a</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td>0.07^c</td>
<td>0.02</td>
<td>-0.01</td>
<td>0.24^c</td>
<td></td>
</tr>
<tr>
<td>REVI'</td>
<td>0.03</td>
<td>0.05^c</td>
<td>0.26^a</td>
<td>0.55^a</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.04</td>
<td>-0.09^a</td>
<td>-0.08^b</td>
<td>-0.01</td>
<td>0.06^b</td>
<td>0.11^a</td>
<td>0.29^c</td>
<td></td>
</tr>
<tr>
<td>REV2'</td>
<td>0.11^a</td>
<td>-0.01</td>
<td>0.14^a</td>
<td>0.40^a</td>
<td>-0.05</td>
<td>0.17^a</td>
<td>-0.06</td>
<td>-0.10^a</td>
<td>-0.12^a</td>
<td>-0.05</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.15^b</td>
<td></td>
</tr>
<tr>
<td>DEXP'</td>
<td>0.02</td>
<td>-0.03</td>
<td>-0.02</td>
<td>0.00</td>
<td>-0.10^b</td>
<td>0.20^a</td>
<td>0.70^a</td>
<td>0.66^a</td>
<td>0.60^a</td>
<td>0.04</td>
<td>-0.01</td>
<td>-0.05^b</td>
<td>-0.14^b</td>
<td></td>
</tr>
<tr>
<td>RND'</td>
<td>-0.02</td>
<td>-0.04</td>
<td>0.11^a</td>
<td>-0.06</td>
<td>-0.05</td>
<td>0.23^a</td>
<td>0.13^a</td>
<td>0.09^b</td>
<td>0.06</td>
<td>-0.10^b</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.15^b</td>
<td></td>
</tr>
<tr>
<td>SGA'</td>
<td>0.06^b</td>
<td>-0.03</td>
<td>-0.05^c</td>
<td>0.01</td>
<td>-0.13^a</td>
<td>0.70^a</td>
<td>0.15^a</td>
<td>0.50^a</td>
<td>0.53^a</td>
<td>0.03</td>
<td>-0.05^c</td>
<td>0.01</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>PROD1'</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.06^b</td>
<td>-0.15^a</td>
<td>0.67^a</td>
<td>0.08^b</td>
<td>0.53^a</td>
<td>0.88^a</td>
<td>0.05</td>
<td>-0.06^b</td>
<td>-0.11^a</td>
<td>-0.37^c</td>
<td></td>
</tr>
<tr>
<td>PROD2'</td>
<td>-0.07^b</td>
<td>0.00</td>
<td>-0.04</td>
<td>-0.06^c</td>
<td>-0.19^a</td>
<td>0.58^a</td>
<td>0.08^b</td>
<td>0.55^a</td>
<td>0.88^a</td>
<td>0.01</td>
<td>-0.02</td>
<td>0.00</td>
<td>-0.34^c</td>
<td></td>
</tr>
<tr>
<td>DGAIN'</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.15^a</td>
<td>0.04</td>
<td>-0.03</td>
<td>0.01</td>
<td>-0.10^b</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.03</td>
<td>0.00</td>
<td>-0.01</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>-0.11^a</td>
<td>-0.12^a</td>
<td>0.03</td>
<td>0.06^b</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.04</td>
<td>0.00</td>
<td>0.06^a</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>MTB</td>
<td>-0.02</td>
<td>0.01</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.01</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.07^b</td>
<td>-0.01</td>
<td>0.03</td>
<td>-0.04^c</td>
<td>0.39^c</td>
<td></td>
</tr>
<tr>
<td>ROA</td>
<td>0.03</td>
<td>0.02</td>
<td>0.25^a</td>
<td>0.13^a</td>
<td>0.07^c</td>
<td>0.03</td>
<td>0.12^a</td>
<td>0.07^b</td>
<td>-0.23^a</td>
<td>-0.20^a</td>
<td>0.05</td>
<td>0.21^a</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

Panels A to C depict Pearson (bottom left) and Spearman rank correlations (top right) for the variables required to test for earnings management techniques. $SMBE^Q$ (SMISSQ) is a dummy variable set to 1 if benchmark $Q$ is barely achieved (missed) and 0 otherwise. $FC'$ is the proxy variable for earnings guidance (Matsumoto, 2002), $DACC$ for aggregate accruals management (Jones, 1991; Dechow et al., 2003), $REVI'$ and $REV2'$ for premature revenue recognition (Stubben, 2010), $DEXP'$, $RND'$, and $SGA'$ for discretionary cost cutting, $PROD1'$ and $PROD2'$ for overproduction, and $DGAIN'$ for abnormal gains from selling fixed assets (Roychowdhury, 2004, 2006; Gunny, 2010). $SIZE$ is firm size measured as natural logarithm of total assets, $MTB$ is the market value of equity divided by its book value, and $ROA$ is return on (total) assets. Significance at the 1%, 5%, and 10% level is indicated by $^a$, $^b$, and $^c$, respectively.
7.4 Empirical Results

This section presents results from univariate and multivariate tests for specific patterns of earnings and expectations management around earnings benchmarks. Main results for the analyst forecast, zero earnings, and earnings changes benchmark are summarized in Sections 7.4.1, 7.4.2, and 7.4.3, respectively. Additional analyses and robustness tests are provided in Section 7.4.4.

7.4.1 The Analyst Forecast Benchmark

To test whether the pile-up of observations in distribution of earnings surprises is caused by earnings or expectations management, I split the overall sample in three groups. The first group, SMBE, covers all firms that just achieve the benchmark. The second group, SMISS, consists of all observations that fall barely short of the latest analyst consensus forecast. Eventually, the third group, OTHERS, comprises all remaining firms in the sample. Under the maintained hypothesis that firms engage in manipulative activities to cross the benchmark, I predict the earnings and expectations management proxy variables of SMBE group to be on average higher than those of the SMISS and OTHERS group. Formally, I expect

\[ \text{PROXY}_{\text{SMBE}} > \text{PROXY}_{\text{OTHERS}} \quad \text{and} \quad \text{PROXY}_{\text{SMBE}} > \text{PROXY}_{\text{SMISS}}. \]

To ensure consistency with the previous analyses, I assign firms as SMBE (SMISS) if they reside in interval \(+1 (−1)\) of the histogram underlying the tests in Section 6.4.1. Hence, the SMBE and SMISS group cover earnings surprise observations within \([0.000, 0.032)\) and \([-0.032, 0.000)\), respectively. The OTHERS group contains all observations outside interval \(+1\) and \(-1\) (i.e., \(-0.032 > \text{earnings surprise} \geq 0.032\)).

I use uni- and multivariate tests to compare the proxies of the three groups. Univariate results are summarized in Table 7.4. Panel A and B compare mean proxy variables of the SMBE group with the OTHERS and SMISS group, respectively. In both panels, the last two columns show differences in means and related \(p\)-values. If firms engage in specific types of earnings management or forecast guidance to achieve the latest consensus estimate, the respective difference in means is expected to be greater than zero and statistically significant. In the multivariate analysis, I test for differences between the three groups using a pooled regression approach (see Section 7.1 for details). In contrast to the univariate analysis, regression allows to control for firm characteristics supposed to affect earnings management proxies (i.e., size, M/B-ratio, and firm performance). If
the SMBE group exhibits higher average proxy values than the OTHERS group, the coefficient on $SMBE^{FC}$ ($\beta_1$) is expected be significantly positive. Similarly, if the SMBE group exhibits higher proxy values than the SMISS group, the coefficient on $SMBE^{FC}$ ($\beta_1$) is expected to be significantly larger than the coefficient on $SMISS^{FC}$ ($\beta_2$). To test for differences between $\beta_1$ and $\beta_2$, I use one-tailed tests of differences in coefficient estimates (i.e., $H_0: \beta_1 \leq \beta_2$).

The last column in Panel A of Table 7.4 reveals that differences in means are positive for seven of ten proxy variables. However, most of these lack significance at generally accepted levels and only the proxies for discretionary revenues ($REV1'$ and $REV2'$) exhibit differences which are significant at at least 5%. Comparisons of the SMBE and the SMISS group in Panel B of Table 7.4 show that though eight of ten differences are positive, none is significantly larger than zero at generally accepted levels. These results do not suggest earnings or expectations management to avoid missing the analyst consensus forecast. Multivariate regression analysis summarized in Table 7.5 confirms univariate results: The coefficient estimate of $\beta_1$ ($SMBE^{FC}$) is above zero for seven of the ten proxy variables. Except for the $REV1'$-specification (significant at the 10% level), however, none of the estimates is statistically significant at generally accepted levels. Furthermore, comparing the estimates of $\beta_1$ ($SMBE^{FC}$) and $\beta_2$ ($SMISS^{FC}$) does not suggest significant differences in any of the ten regression specifications (see the lines labeled “$H_0: \beta_1 \leq \beta_2$”). Taken together, none of the proxy variables indicates earnings or expectations management activity around the analyst forecast benchmark.

Overall, the results do not suggest that the discontinuity in the distribution of earnings surprises is caused by any form of earnings management or forecast guidance. This evidence contrasts with the results of the distributional approach in Chapter 6 and calls into question whether discontinuities in the distribution of earnings surprises are indeed attributable to managerial intervention. Furthermore, it conflicts with German survey results that confirm the use management forecasts to guide analysts down to beatable targets (Nöldeke, 2007a). One potential explanation for this conflict is severe measurement error in unexpected forecasts $FC'$. To address this problem, Section 7.4.4.1 provides additional analyses with an alternative measure of forecast guidance.
### TABLE 7.4

Techniques to Achieve the Consensus Forecast—Univariate Analysis

**Panel A: Small Meet or Beat (SMBE) vs. All Other Firms (OTHERS)**

<table>
<thead>
<tr>
<th>Proxy</th>
<th>SMBE</th>
<th>OTHERS</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>N</td>
</tr>
<tr>
<td>FC'</td>
<td>134</td>
<td>0.1401</td>
<td>482</td>
</tr>
<tr>
<td>DACC</td>
<td>129</td>
<td>-0.0051</td>
<td>485</td>
</tr>
<tr>
<td>REV1'</td>
<td>132</td>
<td>0.0119</td>
<td>496</td>
</tr>
<tr>
<td>REV2'</td>
<td>83</td>
<td>0.0176</td>
<td>277</td>
</tr>
<tr>
<td>DEXP'</td>
<td>132</td>
<td>0.0129</td>
<td>483</td>
</tr>
<tr>
<td>RND'</td>
<td>87</td>
<td>0.0022</td>
<td>337</td>
</tr>
<tr>
<td>SGA'</td>
<td>128</td>
<td>0.0012</td>
<td>448</td>
</tr>
<tr>
<td>PROD1'</td>
<td>132</td>
<td>0.0192</td>
<td>485</td>
</tr>
<tr>
<td>PROD2'</td>
<td>126</td>
<td>0.0245</td>
<td>452</td>
</tr>
<tr>
<td>DGAIN'</td>
<td>66</td>
<td>-0.0007</td>
<td>268</td>
</tr>
</tbody>
</table>

**Panel B: Small Meet or Beat (SMBE) vs. Small Miss Firms (SMISS)**

<table>
<thead>
<tr>
<th>Proxy</th>
<th>SMBE</th>
<th>SMISS</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>FC'</td>
<td>134</td>
<td>74</td>
<td>-0.0007</td>
</tr>
<tr>
<td>DACC</td>
<td>129</td>
<td>68</td>
<td>0.0037</td>
</tr>
<tr>
<td>REV1'</td>
<td>132</td>
<td>75</td>
<td>0.0096</td>
</tr>
<tr>
<td>REV2'</td>
<td>83</td>
<td>44</td>
<td>-0.0019</td>
</tr>
<tr>
<td>DEXP'</td>
<td>132</td>
<td>71</td>
<td>-0.0144</td>
</tr>
<tr>
<td>RND'</td>
<td>87</td>
<td>51</td>
<td>-0.0019</td>
</tr>
<tr>
<td>SGA'</td>
<td>128</td>
<td>68</td>
<td>-0.0082</td>
</tr>
<tr>
<td>PROD1'</td>
<td>132</td>
<td>69</td>
<td>-0.0153</td>
</tr>
<tr>
<td>PROD2'</td>
<td>126</td>
<td>65</td>
<td>-0.0209</td>
</tr>
<tr>
<td>DGAIN'</td>
<td>66</td>
<td>27</td>
<td>0.0016</td>
</tr>
</tbody>
</table>

Column SMBE in Panel A and B depicts the means of ten different earnings and expectations management proxy variables for firms with zero or slightly positive earnings surprises (i.e., $0.000 \leq$ earnings surprise $< 0.032$). Columns SMISS (Panel B) and OTHERS (Panel A) depict the proxy means for firms with small negative earnings surprises (i.e., $-0.032 \leq$ earnings surprise $< 0.000$) and all remaining firms (i.e., $-0.032 >$ earnings surprise $\geq 0.032$), respectively. FC' is a proxy variable for earnings guidance (Matsumoto, 2002), DACC for aggregate accruals management (Jones, 1991; Dechow et al., 2003), REV1' and REV2' for premature revenue recognition (Stuben, 2010), DEXP', RND', and SGA' for discretionary cost cutting, PROD1' and PROD2' for overproduction, and DGAIN' for abnormal gains from selling fixed assets (Roychowdhury, 2004, 2006; Gunny, 2010). FC', DEXP', RND', and SGA' are multiplied with $-1$ so that all proxy variables have the same predicted signs. $P$-values are calculated using one-tailed mean comparison t-tests. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***).
### Table 7.5
**Techniques to Achieve the Consensus Forecast—Regression Analysis**

Model: $PROXY = \alpha + \beta_1 SMBE^{FC} + \beta_2 SMISS^{FC} + \beta_3 SIZE + \beta_4 MTB + \beta_5 ROA + \epsilon$

<table>
<thead>
<tr>
<th></th>
<th>FC'</th>
<th>DACC</th>
<th>REV1'</th>
<th>REV2'</th>
<th>DEXP'</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>CONST</td>
<td>-0.3624</td>
<td>0.169</td>
<td>0.0037</td>
<td>0.686</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>SMBE^{FC}</td>
<td>0.1804</td>
<td>0.133</td>
<td>-0.0093</td>
<td>0.845</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>SMISS^{FC}</td>
<td>0.0392</td>
<td>0.777</td>
<td>-0.0001</td>
<td>0.991</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>SIZE</td>
<td>0.0571</td>
<td>0.156</td>
<td>-0.0006</td>
<td>0.614</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>MTB</td>
<td>0.0192</td>
<td>0.612</td>
<td>-0.0015</td>
<td>0.299***</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>ROA</td>
<td>-0.2756</td>
<td>0.483</td>
<td>0.1417</td>
<td>0.000***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>FC'</th>
<th>DACC</th>
<th>REV1'</th>
<th>REV2'</th>
<th>DEXP'</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$: $\beta_1 \leq \beta_2$</td>
<td>0.164</td>
<td>0.801</td>
<td>0.410</td>
<td>0.126</td>
<td>0.200</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.005</td>
<td>0.043</td>
<td>0.037</td>
<td>0.025</td>
<td>0.037</td>
</tr>
<tr>
<td>$N$</td>
<td>690</td>
<td>682</td>
<td>703</td>
<td>404</td>
<td>686</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>RND'</th>
<th>SGA'</th>
<th>PROD1'</th>
<th>PROD2'</th>
<th>DGAIN'</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>CONST</td>
<td>0.0009</td>
<td>0.768</td>
<td>0.0117</td>
<td>0.677</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>SMBE^{FC}</td>
<td>0.0006</td>
<td>0.429</td>
<td>-0.0040</td>
<td>0.560</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>SMISS^{FC}</td>
<td>-0.0030</td>
<td>0.339</td>
<td>-0.0086</td>
<td>0.637</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>SIZE</td>
<td>-0.0001</td>
<td>0.771</td>
<td>-0.0001</td>
<td>0.988</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>MTB</td>
<td>-0.0003</td>
<td>0.373</td>
<td>-0.0054</td>
<td>0.234</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>ROA</td>
<td>0.0443</td>
<td>0.002***</td>
<td>0.0987</td>
<td>0.012**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>RND'</th>
<th>SGA'</th>
<th>PROD1'</th>
<th>PROD2'</th>
<th>DGAIN'</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$: $\beta_1 \leq \beta_2$</td>
<td>0.209</td>
<td>0.443</td>
<td>0.116</td>
<td>0.111</td>
<td>0.834</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.043</td>
<td>0.009</td>
<td>0.087</td>
<td>0.019</td>
<td>0.003</td>
</tr>
<tr>
<td>$N$</td>
<td>475</td>
<td>644</td>
<td>686</td>
<td>643</td>
<td>361</td>
</tr>
</tbody>
</table>

The table summarizes the results of regressing ten proxy variables (PROXY) for earnings guidance, accrual-based, and real earnings management on dummy variables identifying firms that barely achieve/miss ($SMBE^{FC}$/SMISS^{FC}), the analyst forecast benchmark and control variables. $P$-values for $SMBE^{FC}$ and $H_0$: $\beta_1 \leq \beta_2$ are calculated using one-tailed tests; all others are based on two-tailed tests. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***') asterisks, respectively.
7.4.2 The Zero Earnings Benchmark

If firms engage in earnings manipulations to move from the small loss to the small profit region, I expect firms that just meet or beat the zero earnings benchmark (SMBE group) to exhibit higher earnings management proxies than all other sample firms (OTHERS group) and firms that fall barely short of the benchmark (SMISS group). As in the previous section, I use the interval widths of the histogram in Section 6.4.2 to assign observations to groups as follows: Observations are treated as SMBE if they reside in interval $+1$ (i.e., $0.000 \leq \text{EPS} < 0.224$), as SMISS if they reside in interval $-1$ (i.e., $-0.224 \leq \text{EPS} < 0.000$), and as OTHERS if they lie outside interval $+1$ and $-1$ (i.e., $-0.224 > \text{EPS} \geq 0.224$). I use uni- and multivariate analyses to compare the earnings management proxies of the three groups. Uni- and multivariate results are summarized Table 7.6 and Table 7.7, respectively.

Panel A of Table 7.6 compares the manipulation proxies of the SMBE and the OTHERS group. All proxy means of the SMBE group exceed those of the OTHERS group except for Gunny’s (2010) measure of overproduction ($\text{PROD}_2'$. Six of these differences are statistically significant: Average discretionary accruals ($\text{DACC}$), discretionary revenues ($\text{REV}_2'$), and the proxy for discretionary cuts in SG&A expenses ($\text{SGA}'$) are significantly greater for the SMBE group at the 1% level. The proxies for overproduction ($\text{PROD}_1'$) and discretionary expense cutting ($\text{DEXP}'$) are significantly greater at the 5% level. Eventually, discretionary gains from selling fixed assets ($\text{DGAIN}'$) are significantly greater at the 10% level. Comparing the SMBE group and the SMISS group in Panel B of Table 7.6 weakens preliminary inferences from Panel A. In conflict with the earnings management hypothesis, the proxies for overproduction ($\text{PROD}_1'$, $\text{PROD}_2'$) and fixed asset sales ($\text{DGAIN}'$) are smaller when compared with the SMISS group. For all remaining proxies, positive differences in the last column indicate higher proxy values for the SMBE group. Except for discretionary accruals ($\text{DACC}$), however, none of these differences is significant at generally accepted levels. Hence, above average proxies for discretion in revenues ($\text{REV}_2'$), cost cutting ($\text{DEXP}'$, $\text{SGA}'$), and fixed asset sales ($\text{DGAIN}'$) in Panel A seem to be a common feature of the SMBE and the SMISS group; a finding conflicting with the earnings management explanation. Multivariate analysis generally supports these results: After controlling for firm size, M/B-ratio, and firm performance, the results in Table 7.7 show that the difference between the SMBE and OTHERS group is positive and statistically significant for four earnings management measures (see the $\beta_1$ coefficients in Table 7.7). These include overproduction ($\text{PROD}_1'$) at the 1% level, discretionary accruals ($\text{DACC}$) and discretionary expenses ($\text{DEXP}'$) at the 5% level, and gains from asset sales ($\text{DGAIN}'$) at the 10% level. Comparing the
SMBE with the SMISS group, however, weakens the results. Apart from discretionary accruals
\((DACC)\), none of the SMBE proxy means is significantly greater than those of the SMISS \emph{and}
OTHERS group (see \(\beta_1\) and \(\text{``}H_0: \beta_1 \leq \beta_2\text{''}\)). Taken together, the results suggest that German
managers engage in accrual-based management to avoid reporting a loss. The same result does,
however, not hold for any of the remaining earnings management proxies.

In a recent survey study, Nöldeke (2007b) reports that German managers deny to engage
in earnings management to achieve earnings benchmarks. The results in this section provide
contrary evidence and suggest that the strong discontinuity in the distribution of EPS (see Chapter
6) is at least in part attributable to accrual-based earnings management. However, the relatively
weak evidence also raises doubt that earnings management is the primary and only factor that
causes a kink in the distribution of EPS.
### TABLE 7.6

Techniques to Avoid Losses—Univariate Results

**Panel A: Small Meet or Beat (SMBE) vs. All Other Firms (OTHERS)**

<table>
<thead>
<tr>
<th>Proxy</th>
<th>SMBE N</th>
<th>SMBE Mean</th>
<th>OTHERS N</th>
<th>OTHERS Mean</th>
<th>Difference Mean</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DACC</td>
<td>218</td>
<td>0.0259</td>
<td>83</td>
<td>0.0011</td>
<td>0.0248</td>
<td>0.041**</td>
</tr>
<tr>
<td>REV1'</td>
<td>213</td>
<td>0.0036</td>
<td>86</td>
<td>−0.0070</td>
<td>0.0106</td>
<td>0.143</td>
</tr>
<tr>
<td>REV2'</td>
<td>116</td>
<td>0.0141</td>
<td>37</td>
<td>−0.0084</td>
<td>0.0225</td>
<td>0.104</td>
</tr>
<tr>
<td>DEXP'</td>
<td>215</td>
<td>0.0247</td>
<td>98</td>
<td>0.0081</td>
<td>0.0166</td>
<td>0.263</td>
</tr>
<tr>
<td>RND'</td>
<td>111</td>
<td>0.0014</td>
<td>37</td>
<td>−0.0029</td>
<td>0.0042</td>
<td>0.188</td>
</tr>
<tr>
<td>SGA'</td>
<td>200</td>
<td>0.0349</td>
<td>86</td>
<td>−0.0092</td>
<td>0.0441</td>
<td>0.120</td>
</tr>
<tr>
<td>PROD1'</td>
<td>220</td>
<td>0.0280</td>
<td>90</td>
<td>0.0371</td>
<td>−0.0091</td>
<td>0.611</td>
</tr>
<tr>
<td>PROD2'</td>
<td>206</td>
<td>−0.0112</td>
<td>83</td>
<td>0.0290</td>
<td>−0.0402</td>
<td>0.841</td>
</tr>
<tr>
<td>DGAIN'</td>
<td>74</td>
<td>0.0020</td>
<td>33</td>
<td>0.0059</td>
<td>−0.0039</td>
<td>0.929</td>
</tr>
</tbody>
</table>

**Panel B: Small Meet or Beat (SMBE) vs. Small Miss Firms (SMISS)**

<table>
<thead>
<tr>
<th>Proxy</th>
<th>SMBE N</th>
<th>SMBE Mean</th>
<th>SMISS N</th>
<th>SMISS Mean</th>
<th>Difference Mean</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DACC</td>
<td>218</td>
<td>0.0259</td>
<td>83</td>
<td>0.0011</td>
<td>0.0248</td>
<td>0.041**</td>
</tr>
<tr>
<td>REV1'</td>
<td>213</td>
<td>0.0036</td>
<td>86</td>
<td>−0.0070</td>
<td>0.0106</td>
<td>0.143</td>
</tr>
<tr>
<td>REV2'</td>
<td>116</td>
<td>0.0141</td>
<td>37</td>
<td>−0.0084</td>
<td>0.0225</td>
<td>0.104</td>
</tr>
<tr>
<td>DEXP'</td>
<td>215</td>
<td>0.0247</td>
<td>98</td>
<td>0.0081</td>
<td>0.0166</td>
<td>0.263</td>
</tr>
<tr>
<td>RND'</td>
<td>111</td>
<td>0.0014</td>
<td>37</td>
<td>−0.0029</td>
<td>0.0042</td>
<td>0.188</td>
</tr>
<tr>
<td>SGA'</td>
<td>200</td>
<td>0.0349</td>
<td>86</td>
<td>−0.0092</td>
<td>0.0441</td>
<td>0.120</td>
</tr>
<tr>
<td>PROD1'</td>
<td>220</td>
<td>0.0280</td>
<td>90</td>
<td>0.0371</td>
<td>−0.0091</td>
<td>0.611</td>
</tr>
<tr>
<td>PROD2'</td>
<td>206</td>
<td>−0.0112</td>
<td>83</td>
<td>0.0290</td>
<td>−0.0402</td>
<td>0.841</td>
</tr>
<tr>
<td>DGAIN'</td>
<td>74</td>
<td>0.0020</td>
<td>33</td>
<td>0.0059</td>
<td>−0.0039</td>
<td>0.929</td>
</tr>
</tbody>
</table>

Column SMBE in Panel A and B depicts the means of nine different earnings management proxy variables for firms with zero or slightly positive EPS (i.e., $0.000 \leq \text{EPS} < 0.224$). Columns SMISS (Panel B) and OTHERS (Panel A) depict the proxy means for firms with small negative EPS (i.e., $-0.224 \leq \text{EPS} < 0.000$) and all remaining firms (i.e., $-0.224 > \text{EPS} \geq 0.224$), respectively. **DACC** is a proxy variable for aggregate accruals management (Jones, 1991; Dechow et al., 2003), **REV1' and REV2'** for premature revenue recognition (Stubben, 2010), **DEXP', RND', and SGA'** for discretionary cost cutting, **PROD1' and PROD2'** for overproduction, and **DGAIN'** for abnormal gains from selling fixed assets (Roychowdhury, 2004, 2006; Gunny, 2010). **DEXP', RND', and SGA'** are multiplied with $-1$ so that all proxy variables have the same predicted signs. **P-values** are calculated using one-tailed mean comparison $t$-tests. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***) asterisks, respectively.
### Table 7.7
Techniques to Avoid Losses—Regression Results

| Model: $PROXY = \alpha + \beta_1 \text{SMBE}^{ZE} + \beta_2 \text{SMISS}^{ZE} + \beta_3 \text{SIZE} + \beta_4 \text{MTB} + \beta_5 \text{ROA} + \epsilon$ |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | Cost            | Cost            | Cost            | Cost            | Cost            |
|                 | $p$-value       | $p$-value       | $p$-value       | $p$-value       | $p$-value       |
|                 | $\alpha$ CONST  | $\beta_1$ SMBE$^{ZE}$ | $\beta_2$ SMISS$^{ZE}$ | $\beta_3$ SIZE | $\beta_4$ MTB  | $\beta_5$ ROA  |
| $\alpha$ CONST  | 0.0147          | 0.0179          | 0.0014          | -0.0027         | 0.0002         | 0.1150         |
| $\beta_1$ SMBE$^{ZE}$ | 0.0004        | 0.0013          | -0.0048         | 0.0003          | -0.0004        | 0.0437         |
| $\beta_2$ SMISS$^{ZE}$ | 0.0034        | 0.0123          | -0.0083         | 0.0003          | -0.0003        | 0.0016         |
| $\beta_3$ SIZE   | 0.516           | 0.105           | 0.345           | 0.546           | 0.626          | 0.506          |
| $\beta_4$ MTB    | -0.0122         | 0.0272          | 0.0152          | 0.0026          | -0.0011        | 0.0005         |
| $\beta_5$ ROA    | 0.437           | 0.432           | 0.312           | 0.258           | 0.218          | 0.432          |
| $R^2$            | 0.0004          | 0.0152          | 0.0026          | 0.0011          | 0.0001         | 0.0001         |
| $N$              | 1,320           | 1,264           | 738             | 1,345           | 686            |

The table summarizes the results of regressing nine proxy variables ($PROXY$) for accrual and real earnings management on dummy variables for firms that barely achieve/miss ($SMBE^{ZE}$/SMISS$^{ZE}$) the zero earnings benchmark and control variables. $p$-values on $SMBE^{ZE}$ are calculated using one-tailed tests; all others are based on two-tailed tests. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (*** ) asterisks, respectively.
7.4.3 The Earnings Changes Benchmark

To test for earnings management at the earnings changes benchmark, I compare mean earnings management proxy variables of small meet or beat observations (SMBE group) with those of just miss firms (SMISS group) and all remaining observations (OTHERS group). Under the maintained hypothesis of earnings management to avoid falling short of last year’s EPS, I expect the proxy variables to be significantly higher for the SMBE group. I use the interval widths from Section 6.4.3 to partition the overall sample into the three groups as follows: The SMBE group consists of all observations in the first interval above the earnings changes threshold (i.e., $0.000 \leq \Delta \text{EPS} < 0.124$), the SMISS group covers all observations in the first interval below the benchmark (i.e., $-0.124 \leq \Delta \text{EPS} < 0.000$), and all remaining observations are classified as OTHERS (i.e., $-0.124 > \Delta \text{EPS} \geq 0.124$). Univariate and multivariate results are summarized in Table 7.8 and Table 7.9, respectively.

The results in Panel A of Table 7.8 show that five of nine earnings management proxy variables ($DACC$, $\text{REV}^1$, $\text{REV}^2$, $\text{DEXP}'$, and $\text{SGA}'$) exhibit higher means when compared with all other firms. Of these five earnings management measures, however, only the proxies of discretionary revenues ($\text{REV}^1$, $\text{REV}^2$) and discretionary SG&A cutting ($\text{SGA}'$) exhibit means that are significantly greater at generally accepted levels. Panel B of Table 7.8 shows that five earnings management proxies are higher for SMBE firms when compared with the SMISS group ($\text{REV}^2$, $\text{DEXP}'$, $\text{RND}'$, $\text{SGA}'$, and $\text{DGAIN}'$), but only discretionary SG&A cutting ($\text{SGA}'$) is significant at generally accepted levels. Regression analysis provides slightly different results: After controlling for size, M/B-ratio, and firm performance, five of the nine $\beta_1$ coefficient estimates in Table 7.9 are positive ($DACC$, $\text{RE}^1$, $\text{REV}^2$, $\text{SGA}'$, and $\text{PROD}^1$) and suggest higher proxy means for the SMBE group when compared with all other firms. Of these coefficients, however, only the $\beta_1$ estimate in the $\text{REV}^2'$ specification is significantly greater than zero (10% level). Comparing the SMBE and the SMISS group (see line “$H_0: \beta_1 \leq \beta_2$”), reveals positive differences in five regression specifications ($\text{REV}^2'$, $\text{DEXP}'$, $\text{RND}'$, $\text{SGA}'$, and $\text{DGAIN}'$), but only one difference that is significant at generally accepted levels ($\text{SGA}'$).

The analyses provide conflicting results. Comparing discretionary revenues of the SMBE and OTHERS group reveals that the former are significantly higher. The same does, however, not hold when the SMBE group is compared with the SMISS group. Similarly, regression results reveal that the proxy for discretionary SG&A cutting is significantly greater for the SMBE group when compared with the SMISS group but not when compared with all other firms. Overall, the results do not provide convincing evidence of earnings management to avoid missing last
year's EPS. This confirms survey evidence reported in Nöldeke (2007b), but conflicts with the distributional results in Chapter 6.

### TABLE 7.8

**Techniques to Achieve last Year’s EPS—Univariate Results**

#### Panel A: Small Meet or Beat (SMBE) vs. All Other Firms (OTHERS)

<table>
<thead>
<tr>
<th>Proxy</th>
<th>SMBE</th>
<th>OTHERS</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>N</td>
</tr>
<tr>
<td>DACC</td>
<td>200</td>
<td>0.0063</td>
<td>1,061</td>
</tr>
<tr>
<td>REV1'</td>
<td>181</td>
<td>0.0053</td>
<td>979</td>
</tr>
<tr>
<td>REV2'</td>
<td>124</td>
<td>0.0199</td>
<td>565</td>
</tr>
<tr>
<td>DEXP'</td>
<td>201</td>
<td>0.0096</td>
<td>1,067</td>
</tr>
<tr>
<td>RND'</td>
<td>95</td>
<td>-0.0009</td>
<td>542</td>
</tr>
<tr>
<td>SGA'</td>
<td>189</td>
<td>0.0265</td>
<td>980</td>
</tr>
<tr>
<td>PROD1'</td>
<td>204</td>
<td>-0.0083</td>
<td>1,099</td>
</tr>
<tr>
<td>PROD2'</td>
<td>193</td>
<td>-0.0363</td>
<td>1,006</td>
</tr>
<tr>
<td>DGAIN'</td>
<td>66</td>
<td>-0.0004</td>
<td>575</td>
</tr>
</tbody>
</table>

#### Panel B: Small Meet or Beat (SMBE) vs. Small Miss Firms (SMISS)

<table>
<thead>
<tr>
<th>Proxy</th>
<th>SMBE</th>
<th>SMISS</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>N</td>
</tr>
<tr>
<td>DACC</td>
<td>200</td>
<td>0.0063</td>
<td>118</td>
</tr>
<tr>
<td>REV1'</td>
<td>181</td>
<td>0.0053</td>
<td>112</td>
</tr>
<tr>
<td>REV2'</td>
<td>124</td>
<td>0.0199</td>
<td>63</td>
</tr>
<tr>
<td>DEXP'</td>
<td>201</td>
<td>0.0096</td>
<td>108</td>
</tr>
<tr>
<td>RND'</td>
<td>95</td>
<td>-0.0009</td>
<td>58</td>
</tr>
<tr>
<td>SGA'</td>
<td>189</td>
<td>0.0265</td>
<td>101</td>
</tr>
<tr>
<td>PROD1'</td>
<td>204</td>
<td>-0.0083</td>
<td>114</td>
</tr>
<tr>
<td>PROD2'</td>
<td>193</td>
<td>-0.0363</td>
<td>106</td>
</tr>
<tr>
<td>DGAIN'</td>
<td>66</td>
<td>-0.0004</td>
<td>51</td>
</tr>
</tbody>
</table>

Column SMBE in Panel A and B depicts the means of nine different earnings management proxy variables for firms with zero or slightly positive EPS changes (i.e., 0.000 ≤ ΔEPS < 0.124). Columns SMISS (Panel B) and OTHERS (Panel A) depict the proxy means for firms with small negative EPS changes (i.e., −0.124 ≤ ΔEPS < 0.000) and all remaining firms (i.e., −0.124 > ΔEPS ≥ 0.124), respectively. DACC is a proxy variable for aggregate accruals management (Jones, 1991; Dechow et al., 2003), REV1' and REV2' for premature revenue recognition (Stubben, 2010), DEXP', RND', and SGA' for discretionary cost cutting, PROD1' and PROD2' for overproduction, and DGAIN' for abnormal gains from selling fixed assets (Roychowdhury, 2004, 2006; Gunny, 2010). DEXP', RND', and SGA' are multiplied with −1 so that all proxy variables have the same predicted signs. P-values are calculated using one-tailed mean comparison t-tests. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**) and three (***) asterisks, respectively.
### Table 7.9

Techniques to Achieve last Year’s EPS—Regression Results

Model: $\text{PROXY} = \alpha + \beta_1 \text{SMBE}^{\text{EC}} + \beta_2 \text{SMISSION}^{\text{EC}} + \beta_3 \text{SIZE} + \beta_4 \text{MTB} + \beta_5 \text{ROA} + \varepsilon$

<table>
<thead>
<tr>
<th></th>
<th>DACC</th>
<th>REV1'</th>
<th>REV2'</th>
<th>DEXP'</th>
<th>RND'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>p-value</td>
<td>Coeff.</td>
<td>p-value</td>
<td>Coeff.</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>CONST</td>
<td>0.0039</td>
<td>0.546</td>
<td>-0.0069</td>
<td>0.133</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>SMBE$^{\text{EC}}$</td>
<td>0.0016</td>
<td>0.419</td>
<td>0.0062</td>
<td>0.133</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>SMISSION$^{\text{EC}}$</td>
<td>0.0063</td>
<td>0.315</td>
<td>0.0125</td>
<td>0.039</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>SIZE</td>
<td>-0.0007</td>
<td>0.458</td>
<td>0.0012</td>
<td>0.063</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>MTB</td>
<td>0.0000</td>
<td>0.987</td>
<td>-0.0003</td>
<td>0.207</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>ROA</td>
<td>0.1245</td>
<td>0.000***</td>
<td>0.0374</td>
<td>0.000***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SGAD</th>
<th>PROD1'</th>
<th>PROD2'</th>
<th>DGAIN'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>p-value</td>
<td>Coeff.</td>
<td>p-value</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>CONST</td>
<td>0.0099</td>
<td>0.444</td>
<td>-0.0221</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>SMBE$^{\text{EC}}$</td>
<td>0.0208</td>
<td>0.148</td>
<td>0.0047</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>SMISSION$^{\text{EC}}$</td>
<td>-0.0189</td>
<td>0.225</td>
<td>0.0140</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>SIZE</td>
<td>-0.0015</td>
<td>0.467</td>
<td>0.0042</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>MTB</td>
<td>0.0002</td>
<td>0.676</td>
<td>-0.0022</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>ROA</td>
<td>0.0624</td>
<td>0.011**</td>
<td>-0.2473</td>
</tr>
</tbody>
</table>

| $H_0$: $\beta_1 \leq \beta_2$ | 0.051* | 0.645  | 0.875  | 0.436  |
| $R^2$ | 0.009  | 0.059  | 0.046  | 0.004  |
| N | 1,270  | 1,417  | 1,305  | 692  |

The table summarizes the results of regressing nine proxy variables (PROXY) for accrual and real earnings management on dummy variables for firms that barely achieve/miss (SMBE$^{\text{EC}}$/SMISSION$^{\text{EC}}$) last year’s EPS and control variables. P-values on SMBE$^{\text{EC}}$ are calculated using one-tailed tests; all others are based on two-tailed tests. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***) asterisks, respectively.
7.4.4 Further Analyses and Robustness Tests

This section provides additional analyses and robustness tests to extend and verify the results on earnings management techniques. Section 7.4.4.1 provides additional tests of earnings guidance based on the methodology proposed by Bartov et al. (2002). Section 7.4.4.2 tests the results’ sensitivity to alternative interval widths. In Section 7.4.4.3, I follow the advice of Kothari et al. (2005) and examine performance-adjusted and -matched discretionary accruals around the zero earnings benchmark. Section 7.4.4.4 analyzes the distribution of premanaged EPS. Eventually, potential specification problems are addressed in Section 7.4.4.5.

7.4.4.1 Alternative Expectations Management Measure

Nöldeke (2007b) reports that German firms use management forecasts to guide analysts towards achievable targets. The results in Section 7.4.1 do not suggest a significant relation between earnings guidance ($FC'$) and benchmark achievement ($SMBE^{FC}$). Lack of significance, however, may stem from a noisy measure of earnings guidance. I assume the proxy variable $FC'$ to be noisy because expected earnings changes are estimated based on prior year’s change in earnings (see Eq. (7.2.22) on p. 160). In contrast, Matsumoto (2002) and Burgstahler and Eames (2006) use quarterly specifications to estimate expected earnings changes. Obviously, their quarterly models yield more precise estimates of unexpected forecast errors. To assess whether a lack of significance is attributable to the non-existence of the earnings guidance phenomenon or a noisy measure of expectations management, this section covers alternative tests of earnings guidance based on the methodology proposed by Bartov et al. (2002).

To assess whether managers actively guide analysts towards beatable targets, Bartov et al. (2002) compare the distributions of earnings surprises and forecast errors. I define forecast error $FERR$ as the difference between reported annual EPS and the first consensus forecast after year $t - 1$’s earnings release. In other words, $FERR$ is the (hypothetical) earnings surprise absent any forecast revisions during the year. Similarly, $FERRQ4$ is the difference between reported annual EPS and the first consensus forecast after the end of the third quarter. Earnings surprise $SURP$ is the difference between reported annual EPS and the last consensus forecast prior to year $t$’s earnings announcement. Figure 7.4 illustrates variable definitions on a time line. I follow Bartov et al. (2002) and perform two tests of expectations management:
Fig. 7.4.—Forecast Errors and Earnings Surprises. The time line illustrates the measurement of forecast errors and earnings surprises for year $t$. $FERR$ is the difference between reported annual EPS and the first consensus forecast after year $t-1$’s earnings release. $FERRQ_4$ is the difference between reported annual EPS and the first consensus forecast after the end of the third quarter. $SURP$ is the difference between reported annual EPS and the last consensus forecast prior to year $t$’s earnings announcement.

Test I: Frequencies of Negative $FERR$ and Negative $SURP$. Under the null of no expectations management, I expect no systematic difference in relative frequencies of negative $FERR$, $FERRQ_4$, and $SURP$. If managers guide analysts downward, however, I expect more negative forecast errors ($FERR$) than negative forecast errors at the beginning of the fourth quarter ($FERRQ_4$) and more negative fourth quarter forecast errors than negative earnings surprises at year-end ($SURP$). Denoting relative frequencies as $f_r(\cdot)$, I expect

$$f_r(FERR < 0) > f_r(FERRQ_4 < 0) > f_r(SURP < 0).$$

Table 7.10 confirms the notion that managers guide analysts downward. The relative frequencies of negative $FERR$, $FERRQ_4$, and $SURP$ are 61.7%, 55.2%, and 47.0%, respectively. Hence and in support of earnings guidance, there are relatively more negative forecast errors ($FERR$, $FERRQ_4$) than negative earnings surprises ($SURP$). Furthermore, the difference in relative frequencies of negative $FERRQ_4$ and negative $SURP$ (8.2%) is larger than the difference in the relative frequencies of $FERR$ and $FERRQ_4$ (6.5%). Hence, downward revisions seem to be more prominent in the last quarter of the fiscal year than in the first three quarters.
Table 7.10

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FERR &lt; 0</td>
<td>N in %</td>
<td>N in %</td>
<td>N in %</td>
</tr>
<tr>
<td>730</td>
<td>61.7</td>
<td>653</td>
<td>55.2</td>
</tr>
</tbody>
</table>

(1) – (2): 77 or 6.5%
(2) – (3): 97 or 8.2%
(1) – (3): 174 or 14.7%

The table depicts absolute and relative frequencies of negative forecast errors (\( FERR, FERRQ_4 \)) and negative earnings surprises (\( SURP \)). \( FERR \) is the difference between reported annual EPS and the first consensus forecast after prior year’s earnings release. \( FERRQ_4 \) is the difference between reported annual EPS and the first consensus forecast after the end of the third quarter. \( SURP \) is the difference between reported annual EPS and the last consensus forecast prior to the year’s earnings announcement.

Test II: Frequency of Selected Expectation Paths. If managers do not engage in expectations management, I do not assume more or less observations that turn negative forecast errors into zero or positive earnings surprises than zero or positive forecast errors into negative earnings surprises. If managers engage in earnings guidance, however, I expect the proportion of firms turning negative forecast errors (\( FERR \) or \( FERRQ_4 \)) into zero or positive earnings surprises (\( SURP \)) to be significantly higher than the proportion of firms turning zero or positive forecast errors into negative earnings surprises. Formally, this expectation is formulated as

\[ f_r(SURP \geq 0 | FERR < 0) > f_r(SURP < 0 | FERR \geq 0), \]

and

\[ f_r(SURP \geq 0 | FERRQ_4 < 0) > f_r(SURP < 0 | FERRQ_4 \geq 0). \]

Table 7.11 compares the proportion of firm-years that turn negative forecast errors into zero or positive earnings surprises (suspect observations) with the proportion of firm-years that turn zero or positive forecast errors into negative earnings surprises (non-suspect observations). 34.9% of all negative \( FERR \) observations end up with a zero or positive earnings surprise at year-end. In contrast, only 17.9% of all zero or positive \( FERR \) observations fall short of the final consensus forecast at the end of the year. The difference in proportions
(17.0%) is significant at the 1% level. Similar inferences can be drawn for the analysis of downward guidance in the last quarter: 23.6% of all negative $FERRUQ4$ observations turn into zero or positive earnings surprises at year-end. In comparison, only 10.8% of all zero or positive $FERRUQ4$ observations turn into negative earnings surprises at the end of the year. The difference in proportions (12.8%) is highly significant. These results strongly support the expectations management hypothesis.

**Table 7.11**

<table>
<thead>
<tr>
<th>Observations Suspected to Engage in Expectations Management</th>
<th>Observations Not Suspected to Engage in Expectations Management</th>
<th>Difference in Relative Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>FERR &lt; 0</td>
<td>FERR ≥ 0</td>
<td>17.0%***</td>
</tr>
<tr>
<td>thereof SURP ≥ 0</td>
<td>thereof SURP &lt; 0</td>
<td>17.9%</td>
</tr>
<tr>
<td>Relative Frequency</td>
<td>Relative Frequency</td>
<td></td>
</tr>
<tr>
<td>34.9%</td>
<td>17.9%</td>
<td></td>
</tr>
<tr>
<td>FERRUQ4 &lt; 0</td>
<td>FERRUQ4 ≥ 0</td>
<td></td>
</tr>
<tr>
<td>thereof SURP ≥ 0</td>
<td>thereof SURP &lt; 0</td>
<td></td>
</tr>
<tr>
<td>Relative Frequency</td>
<td>Relative Frequency</td>
<td></td>
</tr>
<tr>
<td>23.6%</td>
<td>10.8%</td>
<td>12.8%***</td>
</tr>
</tbody>
</table>

The table depicts and compares the relative frequencies of observations that turn negative forecast errors into zero or positive earnings surprises (suspect observations) and the relative frequencies of observations that turn zero or positive forecast errors into negative earnings surprises (non-suspect observations). $FERR$ is the difference between reported annual EPS and the first consensus forecast after prior year’s earnings release. $FERRUQ4$ is the difference between reported annual EPS and the first consensus forecast after the end of the third quarter. $SURP$ is the difference between reported annual EPS and the last consensus forecast prior to the year’s earnings announcement. Statistical significance of differences in relative frequencies is tested using a test of proportions. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (****) asterisks, respectively.

The results in Section 7.4.1 are potentially affected by measurement error in expected forecasts. Using a different approach sheds new light on the earnings guidance hypothesis. Comparing the distributions of earnings surprises and forecast errors, I find strong evidence in support of forecast guidance. These results confirm previous German evidence presented in Nöldeke (2007b) and suggest that the irregularity in the distribution of earnings surprises is at least in part attributable to forecast guidance.

### 7.4.4.2 Alternative Interval Widths

This section tests the robustness of the results with respect to alternative interval widths. In the main analyses of this chapter, observations are classified as SMBE, SMISS, and OTHERS based
on the optimal interval widths calculated for the distributional tests in Chapter 6. Since interval width affects the composition of the three groups, the results in this chapter may be spurious due to an arbitrary choice of interval size. To test the results’ sensitivity to interval width, I repeat the regression analysis with four alternative interval widths. These are identical to those used in the robustness tests of the last chapter (see Section 6.4.4.1 for details). Table 7.12 summarizes the results for the discretionary accrual (DACC) specification at the zero earnings benchmark. In support of the main results, the coefficient $\beta_1$ on $SMBE^{ZE}$ is significantly greater than zero for all interval widths. The differences of the coefficient estimates on $SMBE^{ZE}$ ($\beta_1$) and $SMISS^{ZE}$ ($\beta_2$) are significantly positive for four of five interval widths. Taken together, the results do not suggest that the evidence of accrual-based earnings management at the zero earnings threshold is spurious. In unreported tests, I also examine the remaining proxy variables at all three benchmarks to rule out that their lack of significance is attributable to arbitrary binwidth choice. These analyses, however, do not indicate any type of earnings or expectations management for alternative interval widths.

Overall, the results presented in the main tests of this chapter remain qualitatively unchanged for alternative interval widths.

### Table 7.12

<table>
<thead>
<tr>
<th>Interval Width</th>
<th>0.184</th>
<th>0.224</th>
<th>0.241</th>
<th>0.264</th>
<th>0.304</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SMBE^{ZE}$</td>
<td>0.025</td>
<td>0.018</td>
<td>0.018</td>
<td>0.014</td>
<td>0.013</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.003***</td>
<td>0.014**</td>
<td>0.011**</td>
<td>0.030**</td>
<td>0.029**</td>
</tr>
<tr>
<td>$SMBE^{ZE} - SMISS^{ZE}$</td>
<td>0.020</td>
<td>0.016</td>
<td>0.015</td>
<td>0.016</td>
<td>0.015</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.085*</td>
<td>0.097*</td>
<td>0.104</td>
<td>0.094*</td>
<td>0.084*</td>
</tr>
</tbody>
</table>

The table shows the estimated regression coefficients on $SMBE^{ZE}$ ($\beta_1$) and the differences in coefficients on $SMBE^{ZE}$ ($\beta_1$) and $SMISS^{ZE}$ ($\beta_2$) for five alternative interval widths under the discretionary accruals specification (DACC). Alternative interval widths are identical to those used in the robustness tests of the previous chapter (see Section 6.4.4.1). $P$-values are one-tailed. Significance at the 10%, 5%, and 1% level is indicated by one (*), two (**), and three (***). asterisks, respectively.

### 7.4.4.3 Controlling for Firm Performance

The analysis in Section 7.4.2 provides evidence that managers use aggregate discretionary accruals to avoid reporting a loss. However, these results may be spurious and biased towards erroneously rejecting the null of no earnings management when benchmark achievement is correlated with firm performance (see Section 7.2.1.1.4 for details and references). In the regression
The table presents coefficient estimates for $\beta_1$ and $\beta_2$ with related $p$-values from fitting

$$DACC_{PA/PM} = \alpha + \beta_1 SMBEZE + \beta_2 SMISSZE + \beta_3 SIZE + \beta_4 MTB + \beta_5 ROA + \varepsilon$$

via pooled OLS regression and Huber (1967)-White (1980, 1982) heteroskedasticity-consistent sandwich estimators of standard errors. $DACC_{PA}$ and $DACC_{PM}$ are performance-adjusted and performance-matched discretionary accruals, respectively (Kothari et al., 2005). All other variables are defined as before. $p$-values are one-tailed for the coefficient on $SMBEZE$ and two-tailed for the coefficient on $SMISSZE$. $H_0: \beta_1 \leq \beta_2$ is tested using a test of differences in regression coefficients; related $p$-values are one-tailed. Significance at the 10%, 5%, and 1% level is indicated by one (**), two (**), and three (***). asterisks, respectively.

analyses, I control for a potential influence of firm performance by including return on total assets ($ROA_{it}$) as independent variable (see Eq. (7.1.1) on p. 141). Kothari et al. (2005) directly correct for firm performance when estimating discretionary accruals and find performance-adjusted and -matched accrual models to be superior when earnings management incentives and firm performance are correlated. To avoid spurious results, this section provides robustness tests based on performance-adjusted and performance-matched discretionary accruals. Following Kothari et al. (2005), I estimate performance-adjusted discretionary accruals ($DACC_{PA}$) as residuals from the modified Jones Model with lagged return on assets ($ROA_{t-1}$) as control variable (see Eq. (7.2.9) on 149). Performance-matched discretionary accruals ($DACC_{PM}$) are calculated as modified Jones Model discretionary accruals minus discretionary accruals of another firm matched on industry, year, and $ROA_{t-1}$.20

Regression results for performance-adjusted and performance-matched discretionary accruals are summarized in Table 7.13. In both specifications, the SMBEZE-coefficient ($\beta_1$) is significantly greater than zero at at least 5%. Furthermore, Wald tests of differences in coefficients

---

20Kothari et al. (2005) report that performance-adjusted and -matched discretionary accruals are better specified when firm performance is based on current ROA. For the zero earnings benchmark, however, Ayers et al. (2006) recommend to use lagged ROA.

21To improve performance-matching, firm-years are only matched when lagged ROA differs by less than 10 percent points.
indicate that $\beta_1$ is significantly greater than $\beta_2$ (at the 5% level in the $DACC^{PA}$ and at the 10% level in the $DACC^{PM}$ specification). Overall, the results show that the positive relation between the small profit group and discretionary accruals is not merely attributable to differences in firm performance and reinforce the reported evidence of accrual-based earnings management to avoid reporting a loss.

### 7.4.4.4 The Distribution of Unmanaged EPS

If accrual-based earnings management is responsible for the discontinuity in the distribution of EPS, then the distribution of premanaged EPS should be considerably smoother. Dechow et al. (2003) compare the distributions of unmanaged and managed earnings and find that unmanaged earnings are less kinky than reported earnings. This finding suggests earnings management as a reasonable explanation for a distributional discontinuity at the zero earnings benchmark. The frequency histogram in Figure 6.7 of Section 6.4.2 exhibits a considerable pile-up of observations in the vicinity of zero EPS. Evidence from the main tests in this chapter (see Section 7.4.2) suggests that at least some of this discontinuity is attributable to accrual-based earnings management. To reinforce this argument, I follow Dechow et al. (2003) and compare the distribution of EPS with the distribution of premanaged EPS.\(^{22}\)

Since premanaged EPS are unknown, I use Lagged Model estimates of discretionary accruals to proxy for premanaged earnings. Specifically, I deduct discretionary accruals per share from reported EPS to arrive at premanaged EPS.\(^{23}\) The results are presented in Figure 7.5. In support of the earnings management hypothesis, the distribution of premanaged EPS (in gray) is clearly smoother than the distribution of reported EPS (in black). For premanaged EPS, the two test statistics (see Panels B and C) neither indicate an underrepresentation of small loss observations in interval $-1$ ($\tau_{-1}^{BD} = 1.72$, $\tau_{-1}^{BP} = 1.85$) nor a highly significant overrepresentation of small profit observations in interval $+1$ ($\tau_{+1}^{BD} = 1.11$, $\tau_{+1}^{BP} = 1.38$). In comparison, both test statistics are clearly above the required significance levels for the distribution of reported EPS ($\tau_{-1}^{BD} = -4.12$, $\tau_{-1}^{BP} = 1.91$).

---

\(^{22}\)Dechow et al. (2003) show that an “unmixing” of distributions generally triggers distributional smoothness. As a result, a smoother distribution of premanaged EPS is a necessary but not a sufficient condition to confirm earnings management. Nevertheless, finding a smoother distribution of unmanaged earnings in combination with other evidence of earnings management decreases the chance of erroneously detecting earnings management activity.

\(^{23}\)That is, I calculate premanaged EPS as reported EPS (WS: 18193/WS: 05202) minus undeflated discretionary accruals ($DACC_t \times TA_{t-1}$) divided by the number of shares used to calculate basic EPS (WS: 05192).
$$\tau_{-1}^{BP} = -2.60; \tau_{+1}^{BD} = 6.09, \tau_{+1}^{BP} = 2.94).$$

Taken together, the results underpin previous evidence of accrual-based earnings management to avoid missing the zero EPS threshold.

**Fig. 7.5.—Distribution of Reported and Premanaged EPS.** Panel A displays the frequency histograms of reported (black) and premanaged EPS (gray). Premanaged EPS are calculated as reported EPS minus discretionary accruals per share (Lagged Model). Panels B and C depict the Burgstahler and Dichev (1997) (BD) and Bollen and Pool (2009) (BP) test statistics of distributional discontinuities, respectively. Kernel bandwidths are calculated using Silverman’s (1986) rule of thumb (Rule V). The binwidths of the histograms in Panel A correspond to the bandwidths of the kernel estimators.

Note that these statistics differ from those reported in Section 6.4.2 because I exclude all observations with no premanaged EPS available to ensure comparability.
7.4.4.5 Specification Issues

A common concern with the pooled regressions applied to panel data in Sections 7.4.1 to 7.4.3 is that coefficient estimates and/or their variability are biased.

- **Omitted Variables.** Coefficient bias arises when unobserved effects that are correlated with any of the explanatory variables are omitted (e.g., fixed industry or time effects).\(^{25}\) To test whether potential omitted industry and/or time fixed-effects affect my results, I rerun the regressions after including dummy variables for industry (two-digit SIC) and year.\(^{26}\) The results do not suggest that coefficients are biased due to endogeneity.

- **Within Cluster Correlation.** Correlated residuals in panel data sets bias estimated standard errors and coefficient variability. As a result, hypothesis tests are potentially unreliable.\(^{27}\) To rule out spurious inferences stemming from biased standard errors, I follow the advice in Petersen (2009) and rerun the regressions with clustered standard errors.\(^{28}\) The results do not indicate problems concerning clustered standard errors.

Overall, the results do not seem to be spurious due to omitted fixed-effects and/or within cluster correlation.


\(^{26}\)For details on identification and treatment of unobservable fixed-effects in panel regressions refer to, e.g., Wooldridge (2002, pp. 265–279).

\(^{27}\)Petersen (2009) and Gow et al. (2010) describe the problem of correlated residuals in panel data sets in detail and offer practical advice for the treatment of correlated residuals in the context of finance and accounting research.

\(^{28}\)Cluster-robust standard errors take into account that errors are correlated (e.g., serially or cross-sectionally) within but not between clusters and adjust for within cluster correlations. They are often referred to as Rogers standard errors. For a description of cluster-robust standard errors see, e.g., Petersen (2009) and Gow et al. (2010). I do not account for clustering within years because cluster-robust standard errors with few clusters (e.g., five years) are biased towards overrejecting the null hypothesis (Cameron et al., 2008).
7.5 Summary

In this chapter, I examined potential earnings management techniques applied by managers to achieve benchmarks and provided a rigid test of the distributional approach. The analysis yields two main results. First, I find significantly more negative forecast errors than negative earnings surprises and significantly more negative forecast errors that turn into positive earnings surprises than positive forecast errors that turn into negative earnings surprises. Both findings suggest that German managers guide analysts down to beatable targets and confirm previous empirical evidence and survey results in Nöldeke (2007b). Second, I show that small profit firms exhibit significantly higher discretionary accruals than small loss firms and all other sample firms. I interpret this as evidence of accrual-based earnings management to avoid reporting a loss. Apart from these findings, the results do not support the notion of earnings management to achieve earnings benchmarks. This is consistent with the survey results presented in Nöldeke (2007b), but contrasts with previous US evidence that documents accrual-based and real earnings management to avoid missing benchmarks (see, e.g., Lin et al., 2006; Roychowdhury, 2006; Gunny, 2010).

The interpretation of the relatively weak results in this chapter requires to recall that the analysis is a joint test of earnings management and my models’ power in correctly detecting discretionary activities (Dechow et al., 2003). Hence, failing to detect earnings management does not necessarily mean that benchmark-driven earnings management is a myth. Alternatively, earnings could be managed but my models are not able to detect it. However, taking into account the alleged high level of earnings management activity in Germany (see, e.g., Leuz et al., 2003) and the broad set of earnings management proxies applied in this study, the weakness of the results remains surprising. Hence, I have doubt that the previously documented discontinuities in the distributions of German earnings metrics (see, e.g., Glaum et al., 2004; Daske et al., 2006) are primarily attributable to managerial discretion and caution to treat the results of distributional tests as ipso facto evidence of earnings management.
Chapter 8

Conclusion

In 1998, the former Chairman of the SEC, Arthur Levitt, expressed concerns about a decrease in the quality of financial reporting in general and earnings in particular. In a speech at the New York University’s Center for Law and Business, he stated:

While the problem of earnings management is not new, it has swelled in a market that is unforgiving of companies that miss their estimates. I recently read of one major US company, that failed to meet its so-called “numbers” by one penny, and lost more than six percent of its stock value in one day. (Levitt, 1998)

Chairman Levitt’s concerns are reflected in a broad body of empirical literature on the importance of earnings targets and the means of achieving such targets. Earnings benchmarks include, e.g., the latest analyst consensus forecast, zero earnings, and last year’s earnings. Achieving these targets supposedly signals financial health and positive performance prospects to investors and other stakeholders. As a result, benchmark achievers are rewarded with a return premium (e.g., Bartov et al., 2002; Kasznik and McNichols, 2002; Lopez and Rees, 2002). Missing an important target, presumably, erodes investor confidence and potentially causes severe drops in stock price performance (e.g., Skinner and Sloan, 2002). These relations are assumed to trigger managerial incentives for discretionary activities to avoid so-called “earnings torpedoes”. Several empirical studies in the US suggest earnings management and earnings guidance to achieve benchmarks (e.g., Matsumoto, 2002; Roychowdhury, 2006; Gunny, 2010). Moreover, US survey evidence reported by Graham et al. (2005) suggests that managers are willing to accept decreases in long-term firm value to avoid missing important benchmarks. The consequences of this earnings game are potentially severe since managing earnings to targets increases market expectations
and forces executives to deliver even more unrealistic earnings figures. Fuller and Jensen (2002) and Jensen (2004) argue that this destructive circle causes real economic and social costs.

Most of the literature on earnings benchmarks is based on US or UK data. European evidence is relatively scarce in comparison. For Germany, only a few papers address the mechanics of the earnings game in some detail (see, e.g., Glaum et al., 2004; Holzapfel, 2004; Daske et al., 2006). This dissertation extends prior research and provides a comprehensive examination of the benchmark beating phenomenon in Germany. Three interrelated research questions form the framework of this study: 1) Do German investors reward benchmark achievement with a premium?, 2) Is benchmark beating a prevalent phenomenon in Germany?, and 3) What techniques are applied by managers to achieve earnings targets? The first question asks whether beating earnings benchmarks is regarded as a positive signal by investors and thus rewarded with a return premium. The second research question asks whether firms that would otherwise barely miss a benchmark engage in discretionary activities to achieve target. The last research question directly addresses the means of benchmark related manipulations. Potential candidates are accrual-based earnings management (e.g., drawing down on accounting reserves), real earnings management (e.g., discretionary cost cutting), and analyst guidance (e.g., issuing management forecasts).

The results shed light on benchmark related earnings management in Germany. In the first empirical part of this study, I use a multivariate regression model to test for the capital market incentives for earnings management. I find that achieving the latest analyst consensus forecast yields a positive abnormal return of 0.8% cumulated over a three day period centered on the earnings announcement date. This result is consistent with the US evidence documented Lopez and Rees (2002). Similarly, achieving last year’s EPS is associated with a cumulative abnormal return of 0.9%. Both results hold after controlling for the information content in current earnings. In comparison, achieving the zero earnings benchmark does not yield an incremental premium; a result inconsistent with capital market incentives for loss avoidance. Despite this incremental premium, benchmark beating also affects the relations of unexpected earnings and returns. At all three benchmarks, the market reaction to unexpected earnings (i.e., the earnings response coefficient) increases significantly when a threshold is achieved. Hence, benchmark-beating seems to affect the information content of earnings. In the second part of the empirical analysis, I examine the distributions of earnings surprises, EPS, and changes in EPS for typical patterns that indicate discretionary activities to achieve benchmarks. I document a significant overrepresentation of observations in the region just above the benchmarks. Though not significant in every specification of my test, I also find less observations than expected in the region just below the
in line with prior German studies (e.g., Glaum et al., 2004; Daske et al., 2006), these results suggest that executives engage in discretionary activities to push earnings across thresholds. Specifically, I document that a maximum of 22% of small loss firms manage earnings to avoid reporting an EPS loss. Similar, but considerably lower in magnitude, 10% of the firms that would otherwise miss the latest analyst consensus forecast are suspected to manage EPS or guide analysts to achieve expectations. Eventually, 8% of firms with otherwise small EPS declines are suspected to increase net income to avoid falling short of prior year’s EPS. The last part of the empirical analysis directly tests for ten different types of manipulations. I find that executives use forecast guidance to avoid missing the latest analyst forecast. None of the earnings management measures, however, suggests manipulative activities to meet market expectations. At the zero earnings benchmark, abnormal discretionary accruals of benchmark beaters reveal that German executives use accounting maneuvers to avoid reporting a loss. Real earnings management, in comparison, does not seem to be a frequent tool to achieve the zero earnings benchmark. Eventually, none of the earnings management measures indicates manipulation to avoid small earnings declines.

This study has several interesting implications. First, I use a rigid test of benchmark related earnings management based on measures of accrual-based and real earnings management and find relatively weak results. These call into question whether specific patterns of earnings distributions may be treated as ipso facto evidence of earnings management. I thus caution researchers to rely on distributional patterns to draw inferences on managerial discretion. Second, my results suggest strong capital market incentives for earnings management at the analyst forecast and the earnings changes benchmark. The empirical results in the last chapter of this study, however, do not indicate accrual-based or real earnings management to achieve these benchmarks. With regard to loss avoidance, managers seem to view zero earnings as an important benchmark, though its achievement does not yield market rewards. These results suggest other explanations for benchmark-driven earnings management than capital market incentives. Third, German managers seem to use forecast guidance rather than earnings management to achieve market expectations. This finding reinforces prior German evidence documented in Holzapfel (2004) and Nöldeke (2007a).

As for every empirical study, some caveats merit discussion. One is the relatively small sample size. Although, a small sample reduces the statistical power of my tests, I refrain from including observations before the year 2005 for two reasons. First, limiting the sample to consolidated financial statements of public firms after 2004 ensures that the results are not biased by a mixture of IFRS and German GAAP accounting. Second, manual checks reveal a significant
drop in data availability and data quality before 2005. Another caveat is a lack of data precision that potentially leads to spurious results. Generally, data quality heavily depends on the respective data items. I use extensive manual checks and rigid sample selection procedures to maximize data validity. Albeit these efforts, it is impossible to completely rule out that erroneous database entries bias the results. Eventually, the outcomes of this study depend on the accuracy of the applied econometric models. If my models are not able to measure earnings management with the necessary degree of precision, the results are spurious. I use a large set of earnings management measures and a simulation approach to minimize the risk of erroneous inferences due to severe measurement errors.

Though this study provides a comprehensive view on benchmark beating, some questions remain unanswered and provide nice areas for future research in the German setting. These include, for instance, a solid explanation for the salience of loss avoidance. Other interesting fields are the capital market consequences of managerial discretion and the effect of earnings management on subsequent firm performance. I leave these questions open for further investigations.
Appendices
## A.1 Accounting Discretion under IFRS

<table>
<thead>
<tr>
<th>Item/Topic</th>
<th>Standard</th>
<th>Areas of Discretion</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPE, Intangibles, and Investment Properties</td>
<td>IAS 16, IAS 17, IAS 36, IAS 38, IAS 40</td>
<td>Cost model vs. revaluation model, Fair value measurement, Depreciation (methods, useful life, residual value), Impairment (indication, recoverable amounts, cash generating units), Leases (classification, fair value measurement, interest rates), R&amp;D capitalization requirements, Recognition of investment properties at cost or fair value</td>
</tr>
<tr>
<td>Inventories and Construction Contracts</td>
<td>IAS 2, IAS 11</td>
<td>Initial measurement of manufactured goods (measurement and allocation of overheads), Measurement of inventory allowances, Cost formulas (first-in, first-out or weighted average cost), Construction contracts (method of revenue recognition, stage of completion, forecasted contract revenues and expenses)</td>
</tr>
<tr>
<td>Provisions and Pensions</td>
<td>IAS 19, IAS 37</td>
<td>Measurement (best estimate of expenditure to settle obligation), Classification provision/contingent liability, Treatment of actuarial gains/losses, Actuarial assumptions, Fair value measurement of plan assets</td>
</tr>
<tr>
<td>Financial Instruments</td>
<td>IAS 39, IFRS 9</td>
<td>Fair value measurement, Classification of financial instruments, Impairment (indication, measurement), Fair value option, Hedge accounting (hedge accounting requirements, hedge effectiveness)</td>
</tr>
<tr>
<td>First Time Adoption</td>
<td>IFRS 1</td>
<td>Exemptions from Business Combinations (IFRS 3), Exemptions from other IFRS (e.g., share-based payments, deemed costs, treatment of actuarial gains/losses, leases)</td>
</tr>
<tr>
<td>Consolidation and Business Combinations</td>
<td>IFRS 3, IAS 27, IAS 28, IAS 31</td>
<td>Proportionate consolidation or equity method for joint ventures, Purchase price allocation and goodwill measurement, Contingent considerations, Measurement of non-controlling interests</td>
</tr>
</tbody>
</table>
# A.2 Earnings Surprise Calculation based on I/B/E/S Data

Results on benchmark-driven earnings management are suspected to be biased if earnings or earnings surprises are based on split-adjusted I/B/E/S actuals and estimates (see Baber and Kang, 2002; Payne and Thomas, 2003). Payne and Thomas (2003) replicate the distributional study of Degeorge et al. (1999) and report that, especially for the analyst forecast benchmark, using split-adjusted I/B/E/S data causes a severe pile-up of observations at the threshold of interest. Previous German studies (see, e.g., Glaum et al., 2004; Daske et al., 2006) do not explicitly specify the type of underlying I/B/E/S data (i.e., split-adjusted or not). However, adjusted estimates were I/B/E/S-standard until 2006 (Robinson and Glushkov, 2006). As a result, previous German studies based on (adjusted) I/B/E/S data can be suspected to be biased. Since the bias may be severe, this appendix draws on Baber and Kang (2002) and Payne and Thomas (2003) to sketch the problems associated with adjusted I/B/E/S data.

When a stock split occurs, I/B/E/S immediately adjusts all current as well as historical forecasts and actuals for consistency in reflecting the current capitalization (Thomson Reuters, 2010). Generally, this procedure may turn historically large forecast errors into relatively small ones and vice versa. Consider, e.g., an earnings surprise of 2.00 per share in 2006. After a 5-for-1 stock split in 2007, the historically large surprise turns into 0.04 on the adjusted summary history file. However, if earnings surprises are deflated by (adjusted) share prices, these scaled surprises are not affected by the stock split adjustment. In the example, a historical share price of, say, 50.00 turns into 10.00 after the 5-for-1 stock split and the deflated earnings surprise would be 0.04 divided by five and reported as 0.004 in the adjusted summary history file. The rounding has severe consequences for earnings surprises: Although

\[ \frac{0.40}{10.00} = 0.004, \]

which is equal to the unadjusted deflated earnings surprise of \( \frac{2.00}{50.00} = 0.004 \). Unfortunately, the story is not as easy as it seems. Every time I/B/E/S adjusts current and prior database entries, the adjusted values are rounded to two decimals. Since every rounding is a loss of information, many earnings surprises are biased towards zero. To illustrate this point, consider the two examples in Figure A.2.1.

In Example A with a 5-for-1 split in 2009, the historical EPS of 1.98 are converted into \( \frac{1.98}{5} = 0.396 \). Due to the rounding performed in the I/B/E/S database, 0.396 turn into 0.04. Similarly, the historical forecast of 2.00 is divided by five and reported as 0.04 in the adjusted summary history file. The rounding has severe consequences for earnings surprises: Although

\[ \frac{0.40}{10.00} = 0.004, \]

which is equal to the unadjusted deflated earnings surprise of \( \frac{2.00}{50.00} = 0.004 \). Unfortunately, the story is not as easy as it seems. Every time I/B/E/S adjusts current and prior database entries, the adjusted values are rounded to two decimals. Since every rounding is a loss of information, many earnings surprises are biased towards zero. To illustrate this point, consider the two examples in Figure A.2.1.

---

1In other words, the split factor in the numerator and denominator cancel each other out.
## Example A

### Historical I/B/E/S database entries after a 5-for-1 stock split in 2009

<table>
<thead>
<tr>
<th></th>
<th>Unadj. 2008</th>
<th>Adj. 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual EPS</td>
<td>1.98</td>
<td>0.40</td>
</tr>
<tr>
<td>Forecasted EPS</td>
<td>2.00</td>
<td>0.40</td>
</tr>
<tr>
<td>Surprise</td>
<td>−0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>in % of EPS</td>
<td>1.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

## Example B

### Historical I/B/E/S database entries after a 50-for-1 stock split in 2009

<table>
<thead>
<tr>
<th></th>
<th>Unadj. 2008</th>
<th>Adj. 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual EPS</td>
<td>1.75</td>
<td>0.04</td>
</tr>
<tr>
<td>Forecasted EPS</td>
<td>2.00</td>
<td>0.04</td>
</tr>
<tr>
<td>Surprise</td>
<td>−0.25</td>
<td>0.00</td>
</tr>
<tr>
<td>in % of EPS</td>
<td>14.30</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The two examples illustrate the severe rounding problem that arises when earnings surprises are calculated based on stock split adjusted I/B/E/S data. Similar examples can be found in Payne and Thomas (2003).

The company reported a negative earnings surprise of $-0.02$ in 2008, the adjustment in 2009 turns it into a zero earnings surprise. The situation gets even worse in Example B with a 50-for-1 split in 2009. In this case, a negative surprise of $-0.25$ turns into a zero earnings surprise. In relation to unadjusted actual earnings, this is an error of 14.3%. The examples illustrate that the rounding procedure heavily biases earnings surprises towards zero. Furthermore, this bias increases with the split-ratio. Since my study primarily focuses on forecast errors in the vicinity of zero, using “unadjusted” data is strongly required.
A.3 Accrual Model Accuracy Test

Simulation Procedure

In the spirit of Kothari et al. (2005), I apply the following a simulation procedure to assess specification and power of several different accrual models in the German setting:

1. Draw a subsample of 50 random firms from the overall sample.

2. Artificially introduce discretion on total accruals and related accounts (50% expense manipulation, 50% revenue manipulation) for the subsample (as percentage of average total assets). For specification testing, the adjustment is zero (i.e., no earnings management activity). For power testing, the adjustment varies between $-8\%$ and $+8\%$ of total assets (i.e., earnings management between $-8\%$ and $+8\%$ of average total assets).

3. Estimate the industry-year model coefficients using the full sample excluding subsample firms.

4. Calculate discretionary accruals for all firms (including the subsample) as OLS predictions from the regressions in Step 3.

5. Calculate the mean estimate of discretionary accruals for subsample firms and test whether the mean is significantly greater than zero ($H_0$: $DACC_{mean} \leq 0$) or smaller than zero ($H_0$: $DACC_{mean} \geq 0$).

6. Repeat step 1 to 5 250 times for each adjustment level.

7. Summarize the rejection frequencies from simulations to assess model power and specification.

Sample

The initial sample is based on 2,100 firm-years (the extended sample for the zero earnings benchmark in Section 6.3.1). I exclude all firm-years that do not have enough data available to estimate all eleven accrual models. Furthermore, I exclude all industry/year combinations with less than 8 observations to ensure reliable coefficient estimates. From the remaining observations, I delete all observations in the 1st or 99th percentile of all variables required to estimate the accrual
models. Eventually, all industry/year combinations with less than 8 observations due to outlier deletion are dropped. The final sample covers 1,196 firm-years from 14 different industries.

**Tested Accrual Models**

All models are estimated cross-sectionally. The tested models include:

**Model 1:** Jones Model (Jones, 1991)

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{\Delta REV_{it}}{TA_{i,t-1}} + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \varepsilon_{it} \quad (A.3.1)
\]

**Model 2:** Modified Jones Model (Dechow et al., 1995)

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \left[ \frac{\Delta REV_{it} - \Delta REC_{it}}{TA_{i,t-1}} \right] + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \varepsilon_{it} \quad (A.3.2)
\]

**Model 3:** Performance-adjusted Jones Model (Kothari et al., 2005)

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{\Delta REV_{it}}{TA_{i,t-1}} + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \beta_3 ROA_{it} + \varepsilon_{it} \quad (A.3.3)
\]

**Model 4:** Performance-adjusted modified Jones Model (Kothari et al., 2005)

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \left[ \frac{\Delta REV_{it} - \Delta REC_{it}}{TA_{i,t-1}} \right] + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \beta_3 ROA_{it} + \varepsilon_{it} \quad (A.3.4)
\]
Model 5: Performance-adjusted lagged Jones Model (Kothari et al., 2005)

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{\Delta \text{REV}_{it}}{TA_{i,t-1}} + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \beta_3 \text{ROA}_{i,t-1} + \epsilon_{it}
\]  

(A.3.5)

Model 6: Performance-adjusted lagged modified Jones Model (Kothari et al., 2005)

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \left[ \frac{\Delta \text{REV}_{it} - \Delta \text{REC}_{it}}{TA_{i,t-1}} \right] + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \beta_3 \text{ROA}_{i,t-1} + \epsilon_{it}
\]  

(A.3.6)

Model 7: Adapted Model (Dechow et al., 2003)

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \left[ \frac{(1 + \hat{\xi}) \Delta \text{REV}_{it} - \Delta \text{REC}_{it}}{TA_{i,t-1}} \right] + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \epsilon_{it}
\]  

(A.3.7)

where \( \xi \) captures the increase in accounts receivable dependent on sales and is estimated as slope coefficient of regressing Eq. (A.3.8) for every industry-year combination:

\[
\frac{\Delta \text{REC}_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \left[ \frac{\Delta \text{REV}_{it}}{TA_{i,t-1}} \right] + \epsilon_{it}
\]  

(A.3.8)

Model 8: Lagged Model (Dechow et al., 2003)

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \left[ \frac{(1 + \hat{\xi}) \Delta \text{REV}_{it} - \Delta \text{REC}_{it}}{TA_{i,t-1}} \right] + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \beta_3 \frac{ACC_{i,t-1}}{TA_{i,t-1}} + \epsilon_{it}
\]  

(A.3.9)
Model 9: Forward-looking Model (Dechow et al., 2003)

\[
\frac{ACC_{it}}{TA_{i,t-1}} = \alpha_1 + \alpha_2 \frac{1}{TA_{i,t-1}} + \beta_1 \left[ \frac{(1 + \hat{\xi}) \Delta REV_{it} - \Delta REC_{it}}{TA_{i,t-1}} \right] + \beta_2 \frac{PPE_{it}}{TA_{i,t-1}} + \beta_3 \frac{\Delta REV_{it}}{REV_{i,t-1}} + \epsilon_{it}
\]

(A.3.10)

Model 10: Performance-matched modified Jones Model

Discretionary accruals of firm \( i \) are calculated as the residual \( \hat{\epsilon}_{it} \) from fitting Eq. (A.3.2) minus the residual \( \hat{\epsilon}_{jt} \) of firm \( j \) matched on industry, year, and \( ROA_t \).

Model 11: Performance-matched lagged modified Jones Model

Discretionary accruals of firm \( i \) are calculated as the residual \( \hat{\epsilon}_{it} \) from fitting Eq. (A.3.2) minus the residual \( \hat{\epsilon}_{jt} \) of firm \( j \) matched on industry, year, and \( ROA_{t-1} \).

Variable Definitions

Variables are defined as follows:

\( ACC \) Total accruals calculated as difference of net income including minority interests (WS: 01751 + 01501) and operating cash flow (WS: 04860),

\( TA \) Total assets as reported (WS: 18184 + 02999),

\( \Delta REV \) Change in revenues (WS: 01001),

\( \Delta REC \) Change in net trade receivables (WS: 18297),

\( PPE \) Net property, plant, and equipment (WS: 02501), and

\( ROA \) Return on assets calculated as the ratio of net income including minority interests (WS: 01751 + 01501) to reported total assets (WS: 18184 + 02999).
Simulation Results

Test results on specification are summarized in Table A.3.1. A well specified model detects no earnings management in situations where no earnings management has been exercised. In statistical terms, the well specified model does not (incorrectly) reject the null of no earnings management and thus minimizes Type I error. The interpretation of Table A.3.1 is straightforward: The closer rejection frequencies are to 0%, the better specified is the model. A rejection frequency of, e.g., 5%, indicates that the model incorrectly detects earnings management in 5% of all cases.

Test results on power are summarized in Table A.3.2. A powerful model detects earnings management in situations, when it is known to exist. Statistically spoken, the model (correctly) rejects the null of no earnings management in situations where earnings management has taken place and thus minimizes Type II error. Again, the interpretation of Table A.3.2 is straightforward: The closer rejection frequencies are to 100%, the more powerful is the model in detecting earnings management activity. A rejection frequency of, e.g., 80%, indicates that the model correctly detects earnings management in 80% of all cases.
**Table A.3.1**

Specification of Accrual Models


<table>
<thead>
<tr>
<th>Model</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0: DACC_{mean} \leq 0 )</td>
<td>5.6</td>
<td>8.8(^a)</td>
<td>6.0</td>
<td>8.8(^a)</td>
<td>6.8</td>
<td>6.4</td>
<td>7.2</td>
<td>5.2</td>
<td>11.2(^a)</td>
<td><strong>3.2</strong></td>
<td>6.0(^b)</td>
</tr>
<tr>
<td>( H_0: DACC_{mean} \geq 0 )</td>
<td>4.8</td>
<td>3.6</td>
<td>5.6</td>
<td>6.0</td>
<td>4.4</td>
<td>3.6</td>
<td>3.6</td>
<td><strong>2.0(^b)</strong></td>
<td>3.6</td>
<td>5.2</td>
<td>4.8</td>
</tr>
</tbody>
</table>

\(^1\)Model by Number:

The table compares the specification of eleven different abnormal accrual models. Due to random sampling, a well specified model is assumed to neither over nor underreject the null of no earnings management. All values are rejection frequencies of the respective hypothesis for 250 simulation runs in %. \(^a\) \((^b)\) indicate rejection rates significantly greater (smaller) than the specified test level of 5% based on one-tailed binomial tests. Bold values signify the lowest rejection rates.
### Table A.3.2
Power of Accrual Models

<table>
<thead>
<tr>
<th>Model by Number:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A:</strong> Income increasing earnings management. Rejection rates for the null $H_0$: $DACC_{mean} \leq 0$ at the 5% significance level.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj. by +1%</td>
<td>26.8</td>
<td>27.2</td>
<td>18.8</td>
<td>22.8</td>
<td>23.6</td>
<td>22.4</td>
<td><strong>28.0</strong></td>
<td>26.8</td>
<td>24.4</td>
<td>18.4</td>
<td>18.0</td>
</tr>
<tr>
<td>Adj. by +2%</td>
<td>50.4</td>
<td>50.8</td>
<td>32.4</td>
<td>40.0</td>
<td>40.0</td>
<td>48.0</td>
<td>46.8</td>
<td><strong>52.0</strong></td>
<td>48.8</td>
<td>47.6</td>
<td>39.6</td>
</tr>
<tr>
<td>Adj. by +3%</td>
<td>75.2</td>
<td>76.4</td>
<td>56.4</td>
<td>58.0</td>
<td>69.6</td>
<td>71.6</td>
<td><strong>77.6</strong></td>
<td>74.4</td>
<td>68.4</td>
<td>56.0</td>
<td>56.8</td>
</tr>
<tr>
<td>Adj. by +4%</td>
<td>90.4</td>
<td>90.4</td>
<td>74.8</td>
<td>78.0</td>
<td>85.2</td>
<td>87.6</td>
<td><strong>91.2</strong></td>
<td>89.2</td>
<td>84.8</td>
<td>78.8</td>
<td>76.0</td>
</tr>
<tr>
<td>Adj. by +6%</td>
<td><strong>99.2</strong></td>
<td><strong>99.2</strong></td>
<td>90.4</td>
<td>92.0</td>
<td>96.4</td>
<td>98.4</td>
<td><strong>99.2</strong></td>
<td>98.8</td>
<td>97.6</td>
<td>91.2</td>
<td>95.2</td>
</tr>
<tr>
<td>Adj. by +8%</td>
<td><strong>100.0</strong></td>
<td>99.6</td>
<td>98.4</td>
<td>99.2</td>
<td>98.8</td>
<td>99.6</td>
<td><strong>100.0</strong></td>
<td><strong>100.0</strong></td>
<td>99.6</td>
<td>97.6</td>
<td>99.6</td>
</tr>
</tbody>
</table>

| **Panel B:** Income decreasing earnings management. Rejection rates for the null $H_0$: $DACC_{mean} \geq 0$ at the 5% significance level. |
|------------------|---|---|---|---|---|---|---|---|---|----|----|
| Adj. by −1%      | 15.6 | **17.2** | 12.4 | 13.6 | 16.0 | 16.4 | 15.6 | 14.8 | 14.4 | 11.6 | 14.0 |
| Adj. by −2%      | 40.8 | 39.2 | 29.6 | 25.2 | 35.6 | **42.8** | 40.0 | 37.6 | 34.4 | 18.4 | 31.6 |
| Adj. by −3%      | 68.8 | 66.8 | 48.4 | 43.6 | 64.4 | 69.2 | 69.6 | **71.2** | 59.2 | 33.6 | 51.2 |
| Adj. by −4%      | 86.8 | 85.6 | 62.8 | 60.8 | 82.8 | 86.4 | 87.6 | **88.0** | 76.8 | 49.2 | 76.0 |
| Adj. by −6%      | **99.2** | 98.8 | 90.8 | 88.0 | 97.6 | 98.4 | 98.4 | **99.2** | 95.6 | 88.0 | 93.2 |
| Adj. by −8%      | **100.0** | 99.6 | 96.8 | 97.2 | **100.0** | **100.0** | **100.0** | **100.0** | 98.0 | 96.0 | 99.6 |

1 Model by Number:


The table compares the power of eleven different abnormal accrual models in detecting earnings management activity. To test for model power, income increasing (Panel A) and income decreasing (Panel B) earnings management is artificially induced for 50 random firm-years in each of 250 simulation runs. A powerful model rejects the null of no earnings management with a high frequency when earnings management is (artificially) induced. All values are rejection frequencies of the respective hypothesis for 250 simulation runs in %. Bold values signify the highest rejection rates.
A.4 Rules of Thumb for Bin- and Bandwidth Calculation

In Chapter 6, I suggest five different rules of thumb to calculate optimal histogram binwidth and kernel bandwidth. This Appendix sketches the basic concepts underlying these rules.

Rules of Thumb for Histogram Binwidth Calculation

An optimal binwidth minimizes the discrepancy between the estimated density $\hat{f}(x)$ and the true underlying density $f(x)$. Scott (1979) uses the (asymptotic) mean integrated squared error, defined as

$$MISE(h) = E \left[ \int_{-\infty}^{\infty} (\hat{f}_h(x) - f(x))^2 \right] dx$$ \hspace{1cm} (A.4.1)

as measure of discrepancy and derives the approximate formula for the optimal binwidth $h^*$ by minimizing $MISE(h)$. The optimal binwidth is given as

$$h^* = \left[ \frac{6}{N ||f'||_2^2} \right]^{1/3}$$ \hspace{1cm} (A.4.2)

with $||f'||_2^2 = \int_{-\infty}^{\infty} f'(x)^2 \, dx$. Appealing on the first sight, this formula bears a drastic caveat as the true density distribution and its derivates are unknown. To overcome this problem, Scott (1979) assumes a normal reference distribution with $f \sim N(\mu, \sigma^2)$. Thus, $||f'||_2^2$ becomes $\frac{1}{4\sqrt{\pi}\sigma^3}$ and

$$h^* = \left[ \frac{6}{N \frac{1}{4\sqrt{\pi}\sigma^3}} \right]^{1/3} = \left[ \frac{24\sqrt{\pi}\sigma^3}{N} \right]^{1/3} \approx 3.4908\sigma N^{-1/3}. \hspace{1cm} (A.4.3)$$

The theoretically optimal binwidth can now be calculated by replacing $\sigma$ with the sample’s standard deviation $\hat{\sigma}$:

$$h^* = 3.4908 \hat{\sigma} N^{-1/3}. \hspace{1cm} (A.4.4)$$
This formula—referred to as Scott’s Rule—yields the optimal binwidth in terms of asymptotic \( \text{MISE} \) when the underlying distribution is assumed to be approximately normal. Since \( \hat{\sigma} \) is sensitive to outliers, Freedman and Diaconis (1981) suggest the sample’s (standardized) interquartile range \( \text{IQR} \) as an alternative measure of spread, which is derived as ratio of the sample’s \( \text{IQR} \) to the interquartile range of the standard normal distribution (i.e., \( x_{0.75} - x_{0.25} = 1.34 \)):

\[
h^* \approx 3.4908 \frac{\text{IQR}}{1.34} N^{-\frac{1}{5}}.
\]

(A.4.5)

Though these rules are appealing, distributions of scaled earnings may be more complex in practice and this complexity usually requires to choose a narrower binwidth (Scott and Sain, 2005).

**Rules of Thumb for Kernel Bandwidth Calculation**

The optimal bandwidth \( w^* \) asymptotically minimizes the mean integrated squared error of the density estimate. In analogy to optimal binwidth of a histogram, it can be shown (see Silverman, 1986, pp. 40–48) that mean integrated squared error is minimized when

\[
w^* = \left[ \frac{||K||_2^2}{N \ ||f''||_2^2 k_2} \right]^{1/5},
\]

(A.4.6)

where \( ||K||_2^2 = \int_{-\infty}^{\infty} K^2(u) \, du \), \( ||f''||_2^2 = \int_{-\infty}^{\infty} f''(x)^2 \, dx \) and \( k_2 = \int_{-\infty}^{\infty} u^2 K(u) \, du \). One can easily see that the optimal bandwidth \( w^* \) depends on the applied kernel function \( K(u) \), sample size \( N \) and the higher derivates of the underlying density function \( f(x) \). While \( K \) depends on kernel choice (e.g., Gaussian or Epanechnikov kernel) and \( N \) is a known sample property, \( f(x) \) is the true distribution function to be estimated and thus unknown. As a practical way out of this dilemma, Silverman (1986, pp. 45–48) proposes to approximate \( f(x) \) by a normal distribution with mean \( \mu \) and variance \( \sigma^2 \). Doing so, the unknown \( ||f''||_2^2 \) may be replaced by (Silverman, 1986, p. 45)

\[
||f''||_2^2 = \sigma^{-5} \int_{-\infty}^{\infty} \phi''(x)^2 \, dx = \sigma^{-5} \frac{3}{8\sqrt{\pi}},
\]

(A.4.7)

with \( \phi''(x) \) denoting the second order derivation of the standard normal probability density func-
tion. Since we are estimating \( f(x) \) with a Gaussian kernel, \( ||K||_2^2 = \frac{1}{2\sqrt{\pi}} \) and \( k_2 = 1 \). Hence, using these values and Eq. A.4.7, the optimal bandwidth \( w^* \) is given as

\[
w^* = \left[ \frac{||K||_2^2}{N||f''||^2_2 k_2} \right]^{\frac{1}{5}} = \left[ \frac{\frac{1}{2\sqrt{\pi}}}{N\sigma^4 \frac{3}{8\sqrt{\pi}}} \right]^{\frac{1}{5}} = \left( \frac{4}{3} \right)^{\frac{1}{5}} N^{-\frac{1}{5}} \sigma \approx 1.059 \sigma N^{-\frac{1}{5}}. \tag{A.4.8}
\]

The calculation of \( w^* \) is now performed in analogy to the histogram binwidth. That is, \( \sigma \) is approximated either using the sample’s standard deviation \( \hat{\sigma} \) or the (standardized) interquartile range \( IQR \) (i.e., \( \frac{IQR}{1.34} \)). Because \( IQR \) is more robust concerning outliers and outperforms \( \sigma \) when the underlying distribution is long-tailed or skewed, it tends to oversmooth when the distribution exhibits bimodality, Silverman (1986, p. 47) suggests to use the smaller of \( \hat{\sigma} \) and \( IQR \). Formally, the optimal bandwidth is therefore derived as

\[
w^* \approx 1.059 \min(\hat{\sigma}, \frac{IQR}{1.34}) N^{-\frac{1}{5}}. \tag{A.4.9}
\]

Although Eq. (A.4.9) provides a good approximation of optimal bandwidth, Silverman (1986, pp. 47–48) recommends to decrease the factor 1.059 to 0.9 to further improve the results:

\[
w^* \approx 0.9 \min(\hat{\sigma}, \frac{IQR}{1.34}) N^{-\frac{1}{5}}. \tag{A.4.10}
\]

**Summary of Rules of Thumb**

The rules of thumb described in this Appendix are summarized in Table A.4.1:

<table>
<thead>
<tr>
<th>Rule</th>
<th>Author(s)</th>
<th>Type</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Scott (1979)</td>
<td>Binwidth</td>
<td>( 3.49 \cdot \hat{\sigma} \cdot N^{-\frac{1}{5}} )</td>
</tr>
<tr>
<td>II</td>
<td>Freedman and Diaconis (1981)</td>
<td>Binwidth</td>
<td>( 2 \cdot IQR \cdot N^{-\frac{1}{5}} )</td>
</tr>
<tr>
<td>III</td>
<td>Silverman (1986)</td>
<td>Bandwidth</td>
<td>( 1.06 \cdot \hat{\sigma} \cdot N^{-\frac{1}{5}} )</td>
</tr>
<tr>
<td>VI</td>
<td>Silverman (1986)</td>
<td>Bandwidth</td>
<td>( 0.79 \cdot \hat{\sigma} \cdot N^{-\frac{1}{5}} )</td>
</tr>
<tr>
<td>V</td>
<td>Silverman (1986)</td>
<td>Bandwidth</td>
<td>( 0.9 \cdot \min(\hat{\sigma}, \frac{IQR}{1.34}) \cdot N^{-\frac{1}{5}} )</td>
</tr>
</tbody>
</table>

The table summarizes rules of thumb for optimal histogram binwidth and kernel bandwidths. \( N \) is sample size, \( \hat{\sigma} \) is the sample’s standard deviation, and \( IQR \) stands for the interquartile range.
References


References


Dechow, P.M., Schrand, C.M., 2004. Earnings Quality. Research Foundation of the CFA Institute, Charlottesville, VA, USA.


References


Lin, S., Radhakrishnan, S., Su, L.N., 2006. Earnings management and guidance for meeting or beating analysts’ earnings forecasts. Unpublished Working Paper, University of Texas at Dallas, TX, USA.


——, 2007b. Expectations Guidance in Public Companies. An Examination of Management Forecast Disclosure in Germany and Switzerland. Versus, Zurich, Switzerland.


StataCorp LP, 2009. STATA Base Reference Manual (Release 11). Stata Press, College Station, TX, USA.


Curriculum Vitae

Alexander Eisele, born on October 8, 1978, in Munich, Germany

Education
2008–2011 University of St. Gallen (HSG), St. Gallen, Switzerland
   Doctoral Student
2002 University of Stellenbosch, Stellenbosch, South Africa
   Exchange Student in Business Administration and Management
1999–2004 Ludwig-Maximilians-University, Munich, Germany
   Student in Business Administration and Management

Professional Experience
2008–2011 University of St. Gallen (HSG), St. Gallen, Switzerland
   Research Associate
2004–2008 Ernst & Young AG, Munich, Germany
   Assurance Services
2000–2004 Internships in Banking, Auditing, Corporate Finance and Marketing