INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600
NONLINEARITY IN THE RETURNS-EARNINGS AND RETURNS-CASH FLOWS
RELATIONS: A STUDY OF LONG MEASUREMENT INTERVALS AND THE
OPERATING CYCLE

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy
in the Graduate School of The Ohio State University

By
Kirsten Lynn Anderson, B.S.

*****

The Ohio State University
1997

Dissertation Committee:
Professor Douglas A. Schroeder, Advisor
Professor Peter D. Easton
Professor Pervin K. Shroff
Professor Ralph A. Walkling

Approved by
Professor Douglas A. Schroeder
Advisor
Department of Accounting and
Management Information Systems
This study analyzes the relations between returns and earnings and returns and cash flows. The purpose is to determine how accounting information is associated with firm performance (market returns). These relations are studied over both short (one and five years) and long (ten years) measurement intervals. Models with varying degrees of distributional restrictions are employed to determine the nature of these relations over each interval. These are OLS, arctan (nonlinear), and a semiparametric (single index) model. Within each measurement interval, firms are partitioned into quartiles based on the length of the firm's operating cycle. The operating cycle measures how long it takes a firm to convert cash outflows to cash inflows from operations. It also measures the approximate length of time that it takes for changes in operations to result in changes in firm profitability. Nonlinearities likely occur in the returns-earnings relation in part because of the timing of recognition differences in the accrual system. Information is likely recognized by the market before it is in accounting. This differential in recognition results in different valuation coefficients on different components of accounting earnings. Certain components of earnings may have already been recognized by the market so a lower coefficient results. It is likely that extreme earnings contain a higher proportion of these types of components so nonlinearities result in the tails of the earnings distribution.
Since changes in operations for firms with shorter operating cycles likely impact the firm's profitability more quickly, I expect the returns-earnings relation to become linear more quickly for these firms. I hypothesize and results show that the returns-earnings relation is nonlinear over the shorter measurement intervals and become linear as the measurement interval is lengthened for firms with shorter operating cycles. The returns-cash flows relation is also nonlinear for most operating cycle partitions over shorter windows and approximates linearity for most operating cycle partitions over longer intervals. On average, earnings explain a higher proportion of returns than do cash flows for all measurement intervals. This result is even more pronounced when well-specified models are employed.
To my family
ACKNOWLEDGMENTS

I want to thank my adviser, Dr. Douglas A. Schroeder for his invaluable help and support throughout my career at Ohio State. I also appreciate valuable comments and suggestions from the other members of my committee, Drs. Peter D. Easton, Pervin K. Shroff, and Ralph A. Walkling. Finally, this paper has benefitted from suggestions by participants at workshops at Ohio State University, SUNY-Buffalo, Syracuse University, University of Notre Dame, Villanova University, University of Massachusetts, University of Maryland, and CUNY-Baruch College.

I also wish to thank my fellow PhD students in accounting and finance for their support during my five years in the doctoral program. I would especially like to thank Shane Corwin, Amy Edwards, Jeff Harris, Suzanne Holifield, Dave Hyland, Lee Pinkowitz, and Gary Taylor for being my best friends and making the doctoral program more enjoyable.

Finally, I would like to thank my family for their love and support. I thank my parents, Bob and Nancy, for always supporting me, teaching me the value of education, and instilling in me and the ambition I needed to complete my degree. I thank my brother Dave, my sister Jen, and my brother-in-law Joel for being supportive and my friends. Finally, I thank Mike Gerry for his support and caring.
VITA

November 29, 1967 ...................................... Born - Wilmington, Delaware

1989 ...................................................... B.S., University of Delaware,
Newark, Delaware

1989-1990 ..................................................... Financial Management Assistant,
Campbell Soup Company, Camden,
New Jersey

1990-1991 ..................................................... Credit Operations Specialist,
Campbell Soup Company, Camden,
New Jersey

1991-1992 ..................................................... Accountant, Accountemps, Columbus, Ohio

1992-1997 ..................................................... Graduate Teaching Associate, The Ohio
State University, Columbus, Ohio

FIELDS OF STUDY

Major Field: Accounting and Management Information Systems

Minor Fields: Economics, Statistics
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>VITA</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
</tbody>
</table>

## CHAPTER

1. **INTRODUCTION** ......................................................... 1

2. **REVIEW OF RELATED LITERATURE** ............................... 8
   2.1 The Returns-Earnings Relation ............................... 9
   2.2 Nonlinearity in the Returns-Earnings Relation ........ 11
   2.3 The Relationship between Returns, Cash Flows, and Accruals ................................. 14
   2.4 The Operating Cycle ............................................. 16
   2.5 Long Measurement Interval Studies ...................... 16
   2.6 Semiparametric Regression ................................. 18
   2.7 Summary ............................................................ 18

3. **HYPOTHESIS DEVELOPMENT** ........................................ 20
   3.1 Nonlinearity in the Returns-Earnings Relation .... 21
BIBLIOGRAPHY ................................................................. 78
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Descriptive Statistics: One-Year Measurement Interval</td>
</tr>
<tr>
<td>2</td>
<td>Descriptive Statistics: Five-Year Measurement Interval</td>
</tr>
<tr>
<td>3</td>
<td>Descriptive Statistics: Ten-Year Measurement Interval</td>
</tr>
<tr>
<td>4</td>
<td>Regression of Length of Operating Cycle on SIC Code</td>
</tr>
<tr>
<td>5</td>
<td>Correlation Between Length of Operating Cycle, Book-to-Market Ratio, and Size of Firm</td>
</tr>
<tr>
<td>6</td>
<td>Returns-Earnings Model Results: One-Year Interval</td>
</tr>
<tr>
<td>7</td>
<td>Returns-Cash Flows Model Results: One-Year Interval</td>
</tr>
<tr>
<td>8</td>
<td>Specification Tests for Returns-Earnings Model: One-Year Measurement Interval</td>
</tr>
<tr>
<td>10</td>
<td>Comparison of Explanatory Power between Adequate Models: One-Year Measurement Interval</td>
</tr>
<tr>
<td>11</td>
<td>Returns-Earnings Model Results: Five-Year Interval</td>
</tr>
<tr>
<td>12</td>
<td>Returns-Cash Flows Model Results: Five-Year Interval</td>
</tr>
</tbody>
</table>
13 Specification Tests for Returns-Earnings Model:
   Five-Year Measurement Interval ............................... 63
14 Specification Tests for Returns-Cash Flows Model:
   Five-Year Measurement Interval ............................... 64
15 Comparison of Explanatory Power between Adequate Models:
   Five-Year Measurement Interval ............................... 65
16 Returns-Earnings Model Results: Ten-Year Interval ......... 66
17 Returns-Cash Flows Model Results: Ten-Year Interval ....... 67
18 Specification Tests for Returns-Earnings Model:
   Ten-Year Measurement Interval ............................... 68
19 Specification Tests for Returns-Cash Flows Model:
   Ten-Year Measurement Interval ............................... 69
20 Comparison of Explanatory Power between Adequate Models:
   Ten-Year Measurement Interval ............................... 70
CHAPTER 1

INTRODUCTION

The purpose of this study is to explore the relationship between returns, earnings, and cash flows. It extends prior studies, in particular Anderson, Schroeder, and Shroff [1996] and Dechow [1994], to determine the nature of the relationships between returns and earnings and returns and cash flows. Specifically, it attempts to determine how accounting information is associated with the market measure of firm performance. It analyzes whether these relationships are nonlinear and if, after aggregating earnings and cash flows over time, these nonlinearities dampen. Operating cycle length is one firm characteristic analyzed to determine its effect on these relations. Parametric (OLS and arctangent) and semi-parametric regressions are employed in this study to determine the nature of these relationships.

The profitability of the firm is measured by its market performance (returns). Accounting also measures a firm’s performance using a different metric. This study attempts to determine whether and how these two performance measures are associated. Accounting numbers are expected to be associated with market performance because both are a measure of a firm’s profitability. Prior literature dating back to Ball and Brown
[1968] and Beaver [1968] show a relation between accounting information and market performance.

However, this association between accounting information and market performance is not high when returns and earnings are measured over short intervals. This low association is true even for measurement periods of a year. The primary reason for this low association over short measurement intervals is timing differences between the market and accounting in the recognition of events. The market typically leads accounting in the recognition of events and information that affects the firm's future profitability and performance. Therefore, by the time that accounting recognizes this information, the market has already impounded this information into its performance measure (returns). For example, the firm may announce that a reorganization is going to occur over the next six months. The market likely impounds the information about the reorganization upon announcement, while accounting recognizes the results of the reorganization at a later date.

This study considers three important factors in attempting to better understand the relationship between accounting information and market performance. They are the variable choice, the model choice, and the measurement interval. This study extends prior research because most prior studies consider only one or two of these factors, but not all three (e.g. Easton, Harris, Ohlson [1992]; Dechow [1994]).

In this study, two variables are considered in explaining returns. They are accounting earnings and operating cash flows. The accounting literature contains many studies of the relationship between returns and earnings dating back to Ball and Brown.
[1968] and Beaver [1968]. Also, several studies have examined the ability of cash flows versus accounting earnings to explain firm performance (e.g. Rayburn [1986], Bowen, Burgstahler, and Daley [1987], Wilson [1986, 1987], and Bernard and Stober [1989]). A comparison of explanatory power of returns for both cash flows and earnings is made. This helps to address the issue of whether accrual accounting earnings are associated with market performance or whether cash flows information does an adequate job in explaining firm performance from the financial reporting side.

The model choice has been less studied in the accounting literature. In this study different regression models are employed to determine adequate model specifications for these relationships. A misspecified model is a potential problem because it may fail to adequately describe or summarize the data. For example, Cheng, Hopwood, and McKeown [1992] document that the parameters of nonearnings disclosures are not stable when the structural relation is misspecified. Also, Das and Lev [1994] study nonlinearities in the returns-earnings and returns-cash flows relations over an annual measurement interval.

The models employed in this study impose varying degrees of restrictions on their structure. The most restricted is OLS which imposes a linear relationship between the dependent and independent variables. Next is the arctan model which allows for a nonlinear relationship but imposes a specific functional form. Finally, a semi-parametric or single index model is employed which relaxes distributional assumptions and does not restrict the data to follow any pre-described functional form. Each model is estimated for a given relation, such as returns-earnings or returns-cash flows. Model specification tests
are performed to determine which restricted model adequately describes the relationship being analyzed.

Finally, this study analyzes the nature of the returns-earnings relationship and the returns-cash flows relationship over several measurement intervals. Aggregation periods of one, five, and ten years are used. This is similar to Easton, Harris, and Ohlson [1992] which studied the aggregation of earnings. By aggregating earnings and cash flows, we can determine their relationship with returns over various measurement intervals.

Varying the measurement interval attempts to address the timing differences question. It is expected that as the measurement interval is extended, the timing differences diminish. In other words, the longer the measurement interval, the more likely it is that accounting and the market recognize the effects of an event in the same period. This recognition in the same period is likely to have implications for the relation between accounting and market performance. Easton, et. al. [1992] show that as the measurement interval is extended, the explanatory power of earnings for returns increases. This study hypothesizes that as the measurement interval is extended the relation between returns and earnings and returns and cash flows becomes linear. This linear relation is a result of the timing differences diminishing.

Anderson, et. al. [1997] analyze the nonlinearities in the returns-earnings relationship over both short and long measurement intervals. Semi- and nonparametric regression techniques are employed in addition to ordinary least squares (OLS) and arctangent (nonlinear) regressions to determine the relationship between returns and earnings. This relationship is determined for one-year, five-year, and ten-year
measurement intervals. Anderson, et. al. [1996] find that the relation between returns and earnings is nonlinear over shorter measurement intervals. They also find evidence that as the length of the measurement interval increases, the returns-earnings relation becomes more linear. In fact, over the ten-year measurement interval, linearity is an adequate characterization of the returns-earnings relationship relative to a very flexible benchmark.

Dechow [1994] concentrates on comparing the relative power of cash flows and earnings in explaining returns. She finds that earnings have more explanatory power relative to cash flows, but as the measurement interval increases (up to four years), the relative explanatory power of cash flows increases. In addition, Dechow reports that as the magnitude of accruals increase, the association of cash flows and stock returns declines. Finally, for firms with longer operating cycles, earnings are more strongly associated with returns than cash flows.

This study extends the Dechow and Anderson, et. al. studies by comparing the relationship between returns and cash flows, and returns and earnings in three ways. First, this study employs different regression models with varying degrees of distributional restrictions to determine the nature of the relationships. Second, comparisons are made as both the earnings and cash flows are aggregated over varying time intervals to determine the nature of the way in which both variables help to explain returns.

Third, within each measurement interval, firms are partitioned into quartiles based on the length of the firm’s operating cycle. The operating cycle is the time between the purchase of the inventory and the collection of cash from the customer. In other words,
the time it takes for a firm to take cash and convert it back into cash inflows from operations. The firm's operating cycle length may be an indicator of when a firm's change in operations will impact its profitability. For example, a change in operations for firms with shorter operating cycles will likely impact profitability relatively quickly. If this is the case, then the market may place a higher valuation coefficient on this information than information concerning a change in a firm's operations if the firm has a longer operating cycle. Therefore, nonlinearities may exist in the returns-earnings and returns-cash flows relations over longer measurement intervals for firms with longer operating cycles.

Results indicate that as the measurement interval increases, the relationship between both returns and earnings and returns and cash flows becomes linear. For firms with shorter operating cycles, the relation between returns and earnings approximates a linear one over a ten-year window. However, for firms with longer operating cycles, the returns-earnings relation is nonlinear even over a ten-year measurement interval. The returns-cash flows relation is approximately linear over the ten-year measurement interval regardless of operating cycle length. This is an important result because it gives insight into one explanation of nonlinearities in the returns-earnings relation and why these nonlinearities dampen over extended measurement intervals. A firm's operating cycle length appears to play a role in its returns-earnings relation.

In addition, on average, earnings explain a higher proportion of returns than do cash flows for all measurement intervals. This is the case when OLS is used to test the relationship and is similar to the results obtained by Dechow [1994]. This result is even
more pronounced when a less restricted model is employed. It is also interesting that the
higher explanatory power for earnings over cash flows occurs regardless of the length of
the operating cycle. Earnings still explain firm value and returns better than cash flows
even for firms which reverse their timing differences rather quickly. This is important for
understanding accounting’s role in measuring firm value. It also appears to support the
Statement of Financial Accounting Concepts Number 1 which states “information about
enterprise earnings and its components measured by accrual accounting generally
provides a better indication of enterprise performance than information about current cash
receipts and payments.” (FASB [1978]).

The remainder of this dissertation is as follows. Chapter 2 reviews the literature
related to this study. The hypotheses are developed in Chapter 3. Chapter 4 describes the
research design. Results are reported in Chapter 5. Conclusions are summarized in
Chapter 6.
CHAPTER 2

REVIEW OF RELATED LITERATURE

This chapter reviews previous literature related to this study. Section 2.1 begins the chapter by describing the accounting literature on the returns-earnings relation. The section begins by describing the early literature linking accounting information to the market performance of a firm. Then, components of earnings are discussed followed by literature on the explanatory power of earnings for returns. Section 2.2 reviews the literature relating to nonlinearity in the returns-earnings relation. Section 2.3 discusses literature that analyzes the relation between returns, accruals, and cash flows. Section 2.4 summarizes the literature relating to the impact of the operating cycle of a firm on the returns-earnings and returns-cash flows relations. Literature on the topic of long measurement interval studies is discussed in Section 2.5. These long measurement interval studies primarily measure earnings and returns over longer periods. Section 2.6 provides a brief description of the literature on semiparametric regression. A further discussion is provided in Appendix B. Finally, this chapter is concluded with a brief summary of the literature in Section 2.7.
2.1 The Returns-Earnings Relation

The relationship between a firm’s accounting earnings and its market returns has been studied extensively in the accounting literature. Some of the earliest empirical work includes Ball and Brown [1968] and Beaver [1968]. Ball and Brown [1968] shows a relationship between earnings and the market performance of the firm’s stock. They find that the market reaction is generally in the same direction as the difference between realized and expected earnings. Beaver [1968] studies both price and volume reactions to earnings announcements. He concludes that accounting earnings contain information based on his finding of significant price and volume reactions surrounding the earnings announcement.

Beaver, Clarke, and Wright [1979] attempt to determine if a relationship exists between returns and the magnitude of earnings forecast errors. This relationship was studied for both individual firms and portfolios of firms. They conclude that a strong positive correlation exists between earnings forecast errors and changes in prices.

Beaver, Lambert, and Morse [1980] analyze the returns-earnings relationship in a different way. They invert the traditional testing of the effect of the firm’s earnings on the firm’s change in price. Instead, Beaver, et. al. regress change in price on change in earnings to determine whether price helps to predict earnings. They do this by grouping the dependent variable. They conclude that earnings do not appear to follow a random walk process when the additional information contained in prices is incorporated into the prediction of future earnings.
Beaver, Lambert, and Ryan [1987] continue studying this type of returns-earnings relationship. Beaver, et. al. [1987] reverse the traditional returns-earnings regression and regress change in earnings on change in price. They conclude that their regression analysis is equivalent to the grouping of the dependent variable technique employed in Beaver, et. al. [1980]. This reverse regression technique has some advantages over the grouping technique. An important advantage is the ability to use multiple regression which includes lagged changes in price. When Beaver, et. al. [1987] add the most immediate lagged change in price variable, it is significant. This finding suggests that information is reflected in the firm's price sooner than the firm's earnings.

In addition to studying earnings as an explanatory variable for market performance, components of earnings have also been studied. Two components of earnings typically analyzed are the permanent component and the transitory component. The impact of each component's persistence on the price is often the measure of interest. For example, transitory components of earnings are not likely to persist so their effect on firm value is thought to be low.

Ohlson [1995] provides a theoretical framework for the way in which earnings components with differential persistence map into firm value. Empirical studies in this area include Lipe [1986] and Ramakrishnan and Thomas [1993].

Ramakrishnan and Thomas [1993] use three types of earnings components in their study: permanent, transitory, and price-irrelevant. They show that these components are a result of accounting rules and financial reporting discretion. In addition, Ramakrishnan
and Thomas conclude that the different components all have different response coefficients and that the price-irrelevant components are related over time.

Another issue in studying the returns-earnings relationship is the level of explanatory power of earnings for returns. Lev [1989] notes that in most major studies of the returns-earnings relation, earnings have a relatively low explanatory power for returns. He attributes this partly to low quality accounting earnings. His paper then suggests ways in which to improve the quality and explanatory power of accounting earnings. One of these is changing the measurement interval which is discussed in Section 2.5 and the empirical methods section.

For a further discussion of returns-earnings studies in accounting, see Brown [1994] and Bernard [1989]. In addition, Brennan [1995] provides a review of some literature concerning the effect of accounting information on stock prices.

2.2 Nonlinearity in the Returns-Earnings Relation

Some recent work in the returns-earnings relation literature has focused on returns and earnings having a nonlinear relationship. These studies attempt to determine the nature of any nonlinearities that exist in the returns-earnings relation and what model is the most accurate specification of this relation.

Freeman and Tse [1992] examine the relation between quarterly abnormal returns and unexpected earnings. They hypothesize this relation to be an “S”-shaped curve due mainly to transitory and price-irrelevant components of earnings. They choose to model the “S”-shaped curve using a nonlinear regression model with the arctangent as the
functional form. They find support for the arctangent model being a better descriptor of this quarterly returns-earnings relation than a linear function estimated using ordinary least squares regression.

Freeman and Tse [1994] extend this work using the arctangent relation by examining other issues relating to the returns and earnings relation. They reach three main conclusions. First, most of the price responses to quarterly earnings announcements occur before the actual announcement. Second, they find evidence of high measurement error in small earnings innovations. Finally, their results indicate that the marginal price response varies inversely with firm size for firms with high (in absolute terms) earnings surprises.

Other literature has examined this nonlinear relation between returns and earnings using a variety of other regression techniques. These include Beneish and Harvey [1991], Cheng, Hopwood, and McKeown [1992], Das and Lev [1994], and Anderson, Schroeder, and Shroff [1997].

Beneish and Harvey [1991] use several models (both parametric and nonparametric) to explore the relation between abnormal returns and unexpected earnings. They use these models in firm-specific time series regressions. The models include a linear model, a piecewise linear model, a quadratic model, a fourier flexible form model, and a nonparametric regression. They find that for some firms, a linear model is an adequate specification, but for other firms there are significant deviations from a linear relationship.
Using nonparametric regression, Das and Lev [1994] further explore nonlinearities in the returns-earnings relation for both annual earnings levels and changes. They find that nonlinearity exists for annual earnings even after the removal of extraordinary items and special items. Nonlinearity exists in the relation between returns and both earnings surprises and levels, but the nonlinearity is more pronounced for the relationship between returns and earnings surprises. Finally, there is some evidence of a nonlinear returns-cash flows relation.

Cheng, Hopwood, and McKeown [1992] explore nonlinearity in the returns-earnings relation. Specifically, they are concerned with the effects of a misspecified model. They conclude that if the returns-earnings model is misspecified, then the coefficients on both earnings and any other independent variables of interest are not reliable.

Anderson, Schroeder, and Shroff [1997] use several regression techniques to analyze the returns-earnings relation over several different measurement intervals. They also incorporate two firm characteristics, size and book-to-market ratio, into their analysis. Their models include linear, arctangent, semiparametric, nonparametric and various interaction models. These models are estimated for a one-, five-, and ten-year measurement interval. They find that significant nonlinearities exist in the one-year returns-earnings relation. However, the relation between returns and earnings over a ten-year measurement interval can be adequately described using a linear function. Finally, they find that adding the ‘control’ variables (size and book-to-market ratio) do not necessarily help in explaining returns if the returns-earnings model is adequately
specified. This is generally true if the 'control' variables are added as a continuous variable in the regression. Explanatory power can be gained by interacting the variables with earnings and using a model that allows for the existence of nonlinearities in the shorter measurement intervals. A linear specification with the interaction of the 'control' variables is adequate to gain explanatory power over the long measurement interval.

2.3 The Relationship between Returns, Cash Flows, and Accruals

The relationship between returns and earnings and returns and cash flows has also been studied extensively in the accounting literature. For the most part, these studies have attempted to determine the incremental explanatory power of accruals over cash flows or cash flows over accruals.

Rayburn [1986] uses a cross-sectional regression of cumulative abnormal earnings on two components of earnings: operating cash flows and accrual adjustments. Using this method, she tests whether accruals provide incremental information given operating cash flows and vice-versa. She then further decomposes accrual adjustments into three components: changes in working capital, depreciation, and deferred taxes. She finds a significant association between operating cash flows and returns conditional on accrual adjustments and a significant association between accrual adjustments and returns conditional on operating cash flows. She also shows that the current accruals contain information while the long-term accruals do not.

The incremental information content of cash flows and accruals is further investigated in Wilson [1986] and Wilson [1987]. Wilson [1986] finds that cash and
total accruals have incremental information compared to earnings. He also finds that the total accruals component of earnings has incremental information over the cash flow component of earnings. Finally, Wilson [1986] finds that long-term accruals provide no incremental information over current accruals.

Wilson [1987] further studies the incremental information content of accruals and cash flows. He finds a positive association between returns and the accruals component and operating cash component of earnings. The evidence is inconclusive as to whether or not working capital from operations provides incremental information. Cash flows and accruals appear to have incremental information beyond earnings itself. Wilson also concludes that the market reacts more favorably to larger cash flows and smaller accrual components of earnings.

Bernard and Stober [1989] extend Wilson’s [1986, 1987] work. Their first finding is that Wilson’s results do not generalize to their sample although they were able to replicate Wilson’s [1987] results using Wilson’s sample period. However, they were also not able to explain a large portion of stock behavior by partitioning the data according to states of the economy, time periods, or management’s expectations of future sales.

Bowen, Burgstahler, and Daley [1987] conclude that cash flows have incremental information relative to earnings. They also find evidence that cash flows have incremental information relative to earnings and working capital from operations. However, they also find evidence that earnings and working capital from operations (accrual measures) have incremental information relative to cash flows.
Dechow [1994] is a study of the returns-earnings and returns-cash flows relations. She regresses abnormal returns on aggregate earnings and aggregate cash flows for measurement periods of one quarter, one year, and four years. She finds that accounting income has higher explanatory power than cash flows for returns over all intervals.

More recently, Chambers, Jennings, and Thompson [1997] attempt to provide more evidence on the incremental information content of cash flows and earnings by looking at the case of capital expenditures. They find that firms which expense capital expenditures have earnings that explain less of their returns than firms which use depreciation.

2.4 The Operating Cycle

One of the firm characteristics that Dechow [1994] studies is the length of the firm's operating cycle's effect on the relation between returns and earnings and returns and cash flows. She finds that accrual income contains more information relative to cash flows for firms with longer operating cycles.

2.5 Long Measurement Interval Studies

Easton, Harris, Ohlson [1992] study the change in earnings' explanatory power of returns as the measurement interval is lengthened. They use aggregate returns regressed on aggregate earnings in their model. The measurement interval lengths are one year, two years, five years, and ten years. The one-year measurement interval produces $R^2$s in the range of 3-8% which is consistent with prior literature. Returns and earnings aggregated
over a ten-year measurement interval produced an $R^2$ of about 60%. Easton, et. al. find that as the measurement interval is extended, the explanatory power of earnings for returns is increased.

Shroff [1997] extends this work by attempting to determine why the explanatory power increases over the extended measurement intervals. She concludes that the difference in accounting recognition diminishes as the aggregation window is lengthened. This results in an increase in both the coefficient on earnings and the $R^2$. She also concludes that the positive correlation between current and future earnings helps to increase the $R^2$ over the ten-year interval.

Dechow [1994] also uses longer measurement intervals in her analysis of the returns-earnings and the returns-cash flows relations. She uses one quarter, one year and four year intervals to determine whether cash flows or earnings have higher explanatory power for returns and if their relative explanatory power changes under certain circumstances such as length of a firm's operating cycle. She concludes that earnings always have higher explanatory power, but the difference becomes less as the measurement interval is extended.

Guay and Sidhu [1996] extend her analysis by looking at short-term versus long-term accruals. They find that cash flows, short-term accruals, and long-term accruals are all economically significant with cash flows having the highest significance and long-term accruals the lowest. They also show that the usefulness of accruals differs for firms depending on the firm's performance, leverage, and growth options.
Anderson, Schroeder, and Shroff [1997] find evidence that as the measurement interval is extended, the relationship between returns and earnings becomes linear. They find that the relationship over shorter measurement intervals (one and five years) is a nonlinear one.

2.6 Semiparametric Regression

The semiparametric regression techniques used in this paper are outlined in Appendix B. The literature from which these techniques are derived are also discussed in Appendix B.

The primary source for the semiparametric regression technique employed is Powell, Stock, and Stoker [1989]. They outline the theory behind this model. Other sources of information on these regression models are Robinson [1988] and Craven and Wahba [1979].

In addition to these econometric papers, some of the accounting papers discussed in Section 2.2 have a good discussion of these techniques. See for example Das and Lev [1994].

2.7 Summary

The returns-earnings relation has been studied in the accounting literature extensively. The modern empirical studies date back to 1968 with Ball and Brown [1968] and Beaver [1968]. Many facets of this relation have been studied such as the components of earnings and the relatively low explanatory power of earnings for returns.
In addition to the returns-earnings relationship, the incremental explanatory power of earnings over cash flows and cash flows over earnings has been studied. The primary studies include Wilson [1986, 1987] and Bernard and Stober [1989]. Dechow [1994] compares the explanatory power of cash flows and earnings for returns.

Finally, the econometric literature as well as some accounting literature contains discussions of semiparametric regression. The Stoker [1991] monograph is the primary source for information on this regression technique.
CHAPTER 3

HYPOTHESES DEVELOPMENT

This chapter discusses the hypotheses and the research question. Section 3.1 describes the nonlinearity in the returns-earnings relation. In this section it is hypothesized that the relationship between returns and earnings is nonlinear in the shorter measurement intervals. Section 3.2 discusses the hypothesis for the returns-earnings relationship as the measurement interval is extended. It is expected that this relation will become linear as the interval over which returns and earnings are measured is lengthened. The hypothesis that nonlinearities exist in the returns-cash flows relation over shorter measurement intervals is outlined in Section 3.3. Also, it is hypothesized that earnings will have higher explanatory power for returns than cash flows over all measurement intervals. This explanatory power differential is expected to be even higher when comparing well-specified models. Section 3.4 hypothesizes that the returns-cash flows relation also becomes linear as the measurement interval is extended. Section 3.5 states the hypothesis that firms with shorter operating cycles are likely to have
a linear relation between returns and earnings and returns and cash flows over shorter measurement intervals than firms with longer operating cycles. Finally, these hypotheses are summarized in Section 3.6.

3.1 Nonlinearity in the Returns-Earnings Relation

It is hypothesized that the relationship between returns and earnings is nonlinear over short measurement intervals. This nonlinear relation is largely due to timing differences in the recognition of information between the market and accounting. The market most likely leads accounting in recognition of the effects of events. For example, if a new product is announced, the market likely reacts to the expected effects of the new product around the announcement of the new product. Accounting, on the other hand, recognizes results once the product is sold. These differences in when information about the firm is recognized is important in determining the relationship between accounting and market performance. It is likely to impact both the functional form of the relation and the level of association (i.e. the explanatory power of earnings or the $R^2$ in the regression of returns on earnings).

As a result of timing differences, there are two components of earnings that are likely to cause differences in valuation. These two components are transitory and price-irrelevant components. Transitory components are portions of earnings that are not likely to be sustained in the future, i.e. they are a one-time event. Price-irrelevant components are portions of earnings that are recognized in accounting in the current period, but the market has already recognized the effects of this information in a prior period. Extreme
earnings are likely to contain price-irrelevant components and transitory components that are either valued less by the market or not valued at all by the market in the period the earnings are released. This dampening of the valuation coefficient on extreme earnings leads to a nonlinear relationship between returns and earnings.

I hypothesize that accounting information is a measure of performance of the firm and will be associated with the market return. This association is hypothesized to be stronger if any nonlinearities in the returns-earnings relationship are allowed to be estimated.

Freeman and Tse [1992, 1994] argue that there are at least two reasons why a linear relationship between returns and earnings may not exist. First, earnings components with differential persistence may not affect the market the same way. Specifically, the market reacts less to transitory components of earnings than to permanent components of earnings. Second, earnings may contain price-irrelevant components. Therefore, the market reaction will be dampened for high and low earnings causing a nonlinear function. Freeman and Tse model this as an arctan function.

In addition to Freeman and Tse’s use of an arctan function to evaluate the returns-earnings relation, other studies have explored the nonlinearities in this relation. For example, Beneish and Harvey [1991] and Lyon and Schroeder [1992] use semi- and nonparametric regressions to relax restrictions on the nature of the nonlinearity.

Anderson, Schroeder, and Shroff [1996] explore the relationship between returns and earnings by using different regression techniques. There are three models employed in their study, each with different degrees of restrictions on the data. The first is ordinary
least squares (OLS) which restricts the relationship between returns and earnings to be linear. The second is a nonlinear regression using the arctangent (arctan) function. This is less restrictive than OLS in that the relationship between returns and earnings can have a nonlinear form. However, the nonlinear form is restricted to be an arctan function. Finally, a semiparametric or single index regression model is employed. This is the most general model used because it imposes no restrictions on the functional form describing the relation between returns and earnings.

3.2 The Returns-Earnings Relation over Long Intervals

It is hypothesized that as the measurement interval is extended, the relation between returns and earnings will become linear. As the measurement interval is lengthened, timing differences in the recognition of events diminishes. It is likely that with a long interval the period of recognition in accounting as well as the market is the same. Therefore, differential valuation coefficients on portions of earnings are less likely (i.e. nonlinearities in the returns-earnings relation dampen). Price-irrelevant components are not as likely to be a factor over longer measurement intervals since they are likely recognized by the market and accounting in the same period.

Easton, Harris, and Ohlson [1992] use OLS to describe the returns-earnings relation and compare $R^2$ values for a one-year, two-year, five-year, and ten-year intervals. Earnings are aggregated over the measurement interval in their analysis. The $R^2$ values for the one-year window average approximately 5% while the ten-year window $R^2$ value is approximately 60%. This work shows that the association between market
performance and accounting information increases as the measurement interval is extended. I hypothesize that as the measurement interval is extended, not only will the explanatory power or level of association between returns and earnings change, but the nature of the relationship will change.

Anderson, et. al. [1997] extends this work by analyzing the nature of the nonlinear relation as earnings are aggregated over one-year, five-year, and ten-year windows. They find that the nonlinearities in the one-year window are such that neither OLS nor arctan regressions are adequate descriptors of the relation. Only the single index model captures the nonlinearities that exist in the short window. As the measurement interval is lengthened, the nonlinearities dampen to the extent that by the ten-year window OLS is the most restricted adequate descriptor of the returns-earnings relationship. That is, the relation is adequately described as linear over a ten-year measurement interval.

When performing future studies such as valuation studies or testing the effect of nonearnings disclosures, knowing the length of time it takes accruals to cancel out (timing differences to reverse) helps to design the empirical model. Using OLS for longer window studies appears to be appropriate in that this model specification allows for reliable inferences. For example, Haw and Lustgarten [1988] and Ohlson and Penman [1992] rely on OLS while aggregating accounting information over long windows. Based on results of Anderson, et. al. [1996] this model seems appropriate. Model specification is important because a misspecified model may not adequately describe and summarize the data. If this occurs, the descriptive statistics of the model, such as the coefficients, may not provide reliable inferences.
3.3 Nonlinearity in the Returns-Cash Flows Relation

Dechow [1994] studies the relation between cash flows and returns. She finds a significant relationship between cash flows and returns in both shorter and longer measurement intervals.¹ Much like the returns-earnings relationship, cash flows and returns are positively related to each other. In addition, the $R^2$ values of the regression of returns on cash flows increase as the measurement interval increases. However, these $R^2$ values are less than the $R^2$ values of the returns on earnings regressions.

Dechow analyzes the returns-cash flows relation using OLS. Using OLS imposes the restriction that the relationship between returns and cash flows be linear. This is most likely not the case, especially over shorter measurement intervals. In order to understand firm valuation, it is important to determine the true relationship between cash flows and firm value. With a misspecified model, the association of cash flows and firm value may not be reliably captured.

I expect that the relation between returns and cash flows is best described by a nonlinear function in the shorter measurement intervals. Cash flows are likely to have a higher variance than earnings and consequently more extreme observations. That is, firms are more likely to report either very high cash flows or very low cash flows for a given shorter measurement interval. These extreme cash flows are likely to contain a relatively high transitory component much as extreme earnings over a short measurement interval do.

¹ Dechow uses measurement intervals of one quarter, one year, and four years.
I choose to analyze the returns-cash flows relation for nonlinearities for two primary reasons. First, I expect cash flows to be at least as variable as earnings resulting in a nonlinear relation between returns and cash flows (Dechow [1994]). Since Anderson, et. al. [1996] demonstrate that the relation between returns and earnings over short intervals is nonlinear, I apply the same techniques to the returns-cash flows relation in order to determine the nature of this relationship. Furthermore, Dechow [1994] shows a weaker measurement of linear fit (lower $R^2$) between cash flows and returns than earnings and returns.

It is likely that nonlinearities in the relation between cash flows and returns occur in the extreme tails of the cash flows distribution as they do in the returns-earnings relation. In the high cash flows region, it may be the case that there is a dampening in the market reaction relative to a more moderate cash flow report. When using operating cash flows as the measurement, then the market may not react as positively to an extremely high cash flow report. Perhaps the firm is collecting revenue in advance or it was slow to collect previous periods' revenue. Neither of these would signal that future cash flows will continue at this level.

Operating cash flows may not have a dampening effect for negative cash flows because negative operating cash flows are not likely to indicate future profitability. However, it is possible that there are cases that the market's reaction is not as negative to negative operating cash flows. For example, a firm may report negative cash flows for the period, but still report high earnings and have the potential for future profitability. This could occur if the firm recognizes a large amount of revenue late in the year but has
not collected the cash yet. In this case, the negative cash flow trend most likely is not expected to continue in the future. Therefore, the market's reaction is dampened.

I also hypothesize that earnings will provide a higher explanatory power for returns than cash flows. Cash flows are expected to provide a measure of firm performance that is associated with the firm's market performance. However, it is expected that earnings provides more information about the firm performance than just cash flows since the matching of revenues and expenses approximate 'normalized earnings'. This higher explanatory power of earnings is expected to be even higher for models that allow for nonlinearities where nonlinearities exist in the relationships (i.e. well-specified models).

3.4 The Returns-Cash Flows Relation over Long Intervals

I expect that as the measurement interval increases, nonlinearities in the returns-cash flows relation are dampened. Dechow [1994] demonstrates that as the measurement interval is increased, the explanatory power of cash flows (both operating and net) increases. However, the \( R^2 \) values of aggregated cash flows do not reach the \( R^2 \) values of aggregated earnings when explaining aggregated returns over the longest measurement interval used which is four years. Over the lifetime of the firm, accruals are equal to cash flows so the relationship between these two variables and returns should become similar as the measurement interval is increased.

---

2 See Dechow [1994] for a further discussion of why the matching of revenues and expenses helps to provide a higher explanatory power for a firm's market performance.
3.5 The Operating Cycle

In addition to manipulating the length of the interval over which returns, earnings, and cash flows are measured, the length of the firm’s operating cycle is also analyzed.

Firms with shorter operating cycles are expected to have a returns-earnings relationship that becomes linear more quickly than firms with longer operating cycles. All information that is relevant to firm performance is naturally impounded into the market return. However, for firms with shorter operating cycles, information about the future is more likely to be impounded with a higher valuation coefficient since it is expected that these future events will occur in the near future resulting in an increase in the firm’s profitability relatively quickly. For firms with longer operating cycles, information about the future may not be valued as highly since it takes longer for these the new information to create profit.

For example, a retailer has a relatively short operating cycle so that news about changes in operations is likely to affect the profitability of the firm in the short-run. Therefore, the market values this information relatively higher than firms with a longer operating cycle because the impact on the firm’s profitability should occur in the near future. This is because it will not take the shorter operating cycle firm as long to begin realizing profits from the sale of the new operations since it does not take long to convert the cash outflows from operations into cash inflows from operations. However, an airplane manufacturer has a long operating cycle and news that impacts its operations may not affect profitability until long into the future so the market places less value on
this information. This is because when an airplane manufacturer invests cash flows into
the operations, it takes a much longer time to realize the cash inflows from these
operations.

Based on Dechow's [1994] study, accruals are not as incrementally valuation
relevant, given cash flows, for measuring firm performance for firms with shorter
operating cycles. This is because earnings and cash flows are more similar for firms with
shorter operating cycles. For this reason, it is expected that the explanatory power of
earnings and cash flows is closer for firms with shorter operating cycles than firms with
longer operating cycles.

3.6 Summary

In summary, this paper explores the returns-earnings and return-cash flows
relations over short and long measurement intervals. The effect of the length of the
firm's operating cycle is also examined. It is hypothesized that over shorter measurement
intervals, the relation between returns and earnings and returns and cash flows will be
nonlinear. However, as the measurement interval is lengthened, it is expected that these
relations will become linear. This is especially true for firms with shorter operating
cycles.
CHAPTER 4

DATA AND METHODS OF TESTING

The purpose of this chapter is to describe the data that is used in this study as well as the regression techniques employed. A more detailed discussion of semiparametric regression is found in Appendix B. Section 4.1 describes where the data is obtained. Section 4.2 shows the method used to partition the data based on the length of the firm’s operating cycle. Section 4.3 describes the regression models used in this study. The model specification test employed is described in Section 4.4. A description of the data is given in Section 4.5. Finally, this chapter is summarized in Section 4.6.

4.1 Data

All accounting information data are obtained from the annual current and research COMPUSTAT files covering the New York and American Stock Exchanges and Over-The-Counter traded firms. Prices are also taken from the COMPUSTAT tapes. The COMPUSTAT coverage is 1950-1994. All per share data are adjusted for stock splits and stock dividends using the COMPUSTAT adjustment factor (item #27).

Price per share employed in this analysis is the price at the firm’s fiscal year-end
(COMPUSTAT item #199). Earnings per share including extraordinary items
(COMPUSTAT item #53) is used in all returns-earnings analysis.3

For firms using a Funds Flow Statement,4 operating cash flows is calculated as
funds from operations plus changes in working capital.5 If a firm reports a Statement of
Cash Flows, operating cash flows is taken directly from the financial statement.6

The length of a firm's operating cycle is measured as the sum of the days goods
are in inventory and the days receivables are outstanding. Trade receivables
(COMPUSTAT item #151), revenue (COMPUSTAT item #12), cost of goods sold
(COMPUSTAT item #41), and total inventory (COMPUSTAT item #3) are used in these
calculations.7

The first ten-year measurement interval for which there is sufficient data is 1976-
1985. The last ten-year measurement interval is 1985-1994. Each firm is randomly
assigned to one ten-year interval. This interval is used for all ten-year analysis. For the

3 When aggregating earnings, as I do in my analysis, it is important to use earnings per
share including extraordinary items. This means that total earnings are being summed across

4 The COMPUSTAT variable format code (Item #318) is used to determine which
statement a firm reported. Typically, the Funds Flow Statement was used prior to 1987 and the
Statement of Cash Flows is used since 1987.

5 COMPUSTAT Item #110 + Item #236.

6 COMPUSTAT Item #308.

7 The formula for calculating length of operating cycle is explicitly documented in
Section 4.2.
five-year analysis, each ten-year interval is divided into two parts, the first five years and
the second five years. For the one-year analysis, each ten-year interval is divided into ten
one-year periods.

4.2 Operating Cycle Partitions

The first step in this analysis is to determine the length of each firm's operating
cycle so that partitions can be formed. By partitioning firms based on the length of their
operating cycle, it can be determined whether the length of the cash-to-cash cycle plays a
role in the relations between returns and earnings and returns and cash flows.
Specifically, firms which have shorter operating cycles and therefore shorter time periods
over which timing differences reverse should have a returns-earnings relation that
becomes linear more quickly than firms with longer operating cycles.

The length of the operating cycle is computed at the middle of the ten-year
measurement interval for each firm. The fifth year of the ten-year measurement interval
is used in the calculation which is described in the next paragraph. Once the firm is put
into a specific operating cycle partition based on the ten-year measurement interval, it
remains in that partition for the analysis of all three measurement intervals.
The length of a firm's operating cycle consists of two components, the time between the purchase and sale of inventory and the collection of cash from the sale. Operating cycle length is calculated as follows (similar to Dechow [1994]).

\[
\text{Operating Cycle Length} = \frac{(365 \times \text{Inventory}) + (365 \times \text{Accounts Receivable})}{\text{Cost of Goods Sold}} - \frac{\text{Sales Revenue}}{365}
\]

After the length of each firm's operating cycle is computed using the ten-year measurement interval, portfolios are formed based on this measurement. The operating cycles are ranked by length and quartiles are calculated. The firm's operating cycle partition remains the same throughout all analyses. The first quartile contains the firms with the shortest operating cycles. Analysis is then performed on each quartile. By using portfolios, it can be determined whether the returns-earnings relation and the returns-cash flows relation becomes linear more quickly for firms with shorter operating cycles.

4.3 Model Description

This section gives an overview of the regression models employed in the analysis of returns, earnings, and cash flows. For a more detailed discussion of the semiparametric regression models employed in this study, see Appendix B.

Three regression models are employed in the analysis. These are ordinary least squares (OLS), arctan (nonlinear), and single index (semiparametric; Powell, Stock, and Stoker [1989]) regression models.

---

* Inventory and accounts receivable balances are the ending balances as of the Year 5 balance sheet. Sales revenue and cost of goods sold are for the period ending Year 5.
In general, the OLS model is described as follows:

\[ E[y|x] = \gamma'x. \]

Specifically in this study, the OLS or linear model that is employed is described as follows:

 Aggregate Returns = \gamma_0 + \gamma_1 \times \text{Aggregate Earnings} + \epsilon

and

 Aggregate Returns = \gamma_0 + \gamma_1 \times \text{Aggregate Cash Flows} + \epsilon.

Aggregate returns are defined as:

\[
\frac{P_t - P_{t-T} + d_{t-T}}{P_{t-T}}
\]

where \( P \) is the price per share (adjusted for stock splits and stock dividends) at fiscal year-end, \( d \) is the dividends per share (adjusted for stock splits and stock dividends), and \( T \) takes the value of 1, 5, or 10.

Aggregate earnings are defined as:

\[ eps_{t, T} \]

where \( eps \) is the primary earnings per share after extraordinary items (adjusted for stock splits and stock dividends) and \( T \) takes the value of 1, 5, or 10.
Aggregate cash flows are defined as:

\[ \text{cfps}_{i,t} \]

where \text{cfps} is either operating cash flows per share (adjusted for stock splits and stock dividends) as calculated in Section 4.1 or free cash flows per share (adjusted for stock splits and stock dividends) as calculated in Section 4.1.

The OLS model is the linear model that is used in the study. This regression imposes a linear relation between the dependent and the independent variable. It is the most restrictive model employed in this study. OLS is chosen as the linear model used because it is the most commonly employed model in returns-earnings and returns-cash flows studies, and it is easy to interpret.

In general, the arctangent (arctan) model is described as follows:

\[ E[y_{ix}] = \delta_0 + \delta_1 \arctan(\delta_2'x). \]

Specifically in this study, the arctan model that is employed is described as follows:

Aggregate Returns = \( \delta_0 + \delta_1 \arctan(\delta_2 \text{ Aggregate Earnings}) + \epsilon \)

and

Aggregate Returns = \( \delta_0 + \delta_1 \arctan(\delta_2 \text{ Aggregate Cash Flows}) + \epsilon \).

All variables in this regression are defined the same as in the OLS regression.

The arctan model is used because it is an 'S'-shaped function that has been used in prior literature (Freeman and Tse [1992, 1994]). The 'S'-shaped function is the theoretical relation between returns and earnings. It also is reasonable to assume this
could be the relation between returns and cash flows for similar reasons that it is considered the theoretical relation between returns and cash flows.

The primary reasons for the 'S'-shaped function are (1) that transitory components of earnings do not impact price as much as permanent components and (2) price-irrelevant components of earnings. These components of earnings also may be components of cash flows as well and therefore cause the 'S'-shaped curve to exist. Also, there are timing differences between accruals and the market that may cause the curvature in the relation.

In general, the semiparametric (single index) model is described as follows:

$$E[y|x] = g(\alpha'x).$$

Specifically in this study, the single index model that is employed is described as follows:

Aggregate Returns = g (\alpha'Aggregate Earnings)

and

Aggregate Returns = g (\alpha'Aggregate Cash Flows).

where g is some function (perhaps nonlinear) and all other variables as are defined in the OLS regression model.

The single index model is the most general model employed because it places the fewest restrictions on the relationship studied. This model allows for deviations from the theoretical 'S'-shaped relation that the arctan models.

All three regression models are run on the length of operating cycle partitions for each of the three measurement intervals. The models are also run without the data first

---

9 For one regressor, the single index model and nonparametric regression are the same.
being partitioned on operating cycle length.

4.4 Specification Test

The first step in the analysis of the returns-earnings relation and the returns-cash flows relation is to estimate each of the models described in Section 4.3. After the estimations are complete, then the models are evaluated against each other to determine the most restricted model that adequately describes the relation in question. The most restricted model that adequately describes the relation is the best one to use because more restricted models are easier to interpret and less costly to estimate.

The models are evaluated against each other in an iterative fashion for each length of operating cycle partition within each measurement interval. Restricted models are compared to the less restricted models. The comparison is performed by regressing the residuals of the more restricted model on the predictions of the less restricted model. For example, OLS can be compared to the arctan model by regressing the OLS residuals on the arctan predictions. The specification test in general is written as follows

\[ R_t - E_t [R_t | X_t] = \beta E_g [R_t | X_t] + \nu_t \]

where \( R_t \) = returns, \( X_t \) = earnings or cash flows, \( \nu_t \) = residuals, \( E_t [ \cdot ] \) is the conditional expectation estimated via the more restricted model, and \( E_g [ \cdot ] \) is the conditional expectation estimated via the more general model (see Stoker [1991]).

---

10 This specification test is for comparing nested models. Although the OLS model is not nested in the arctan model, over the range of my data the arctan and the OLS model will both estimate the same line if a linear relation is the correct specification.
If the coefficient in the specification test, \( \beta \), is significantly positive\(^{11} \), it is an indication of model specification inadequacy. Again, for a more detailed discussion of model specification tests, see Appendix B.

### 4.5 Description of Data

Table 1 shows descriptive statistics for the one-year measurement interval. Firms across all quartiles of operating cycle length appear to comparable to each other when size is measured as either total market value of common equity, total book value of common equity, or total assets. Also, returns appear to relatively similar across operating cycle partitions. Operating cycle quartile 3 does show a lower average earnings per share number than the other quartiles, however, the median earnings per share is similar to the other quartiles. Quartile three has a higher average cash flow per share than the other quartiles, but again, the median is comparable to the remaining partitions.

Table 2 reports descriptive statistics for the five-year measurement interval. As in the one-year descriptive statistics, size is comparable across all partitions of operating cycle length. Quartile 3 still displays a lower than average aggregate earnings per share and a higher aggregate cash flows per share than the other quartiles. However, the medians are similar to the remaining partitions.

Descriptive statistics for the ten-year measurement interval are shown in Table 3. The ten-year univariate statistics show the same pattern as the one-year interval and five-year interval statistics.

\(^{11} \) The critical value for all specification tests is a one-tailed p-value of 0.025.
Table 4 reports the results of a regression of the length of operating cycle on SIC codes. A dummy variable is constructed for each one-digit SIC code. There does not appear to be a significant relation between the length of the operating cycle and industry membership as defined by the first digit of the SIC codes. None of the coefficients are significant.

Pearson correlations between the length of a firm's operating cycle and its book-to-market ratio and size of its market value of common equity are reported in Table 5. There appears to be no significant correlation between either the length of operating cycle and book-to-market ratio or the length of operating cycle and firm size. This is expected from the results of Tables 1, 2, and 3 which show that the firms are comparable in size across the operating cycle partitions as measured in both market value of equity and book value of equity.

4.6 Summary

This chapter summarizes the data and methods of testing. Three regression models are employed in this study or the returns-earnings and returns-cash flows relation: ordinary least squares, an arctangent model, and semiparametric regression. Firms are partitioned into quartiles based on the length of their operating cycles. Then, each model is estimated for each partition over a one year period, a five year period, and a ten year period. Model specification tests are then run to determine the most restricted model that adequately describes both the returns-earnings and the returns-cash flows relations over each measurement interval.
CHAPTER 5

EMPIRICAL ANALYSIS

This chapter reports the results from the regression models for all measurement intervals and all partitions of the operating cycle. Section 5.1 reports the results for the one-year measurement interval. In general, both the returns-earnings and returns-cash flows relations are nonlinear for this short measurement interval for all partitions of operating cycle length. The explanatory power of earnings for returns is also higher than cash flows when comparing well-specified models. Five-year measurement interval results are reported in Section 5.2. It appears that some of the nonlinearities that exist in the short interval begin to dampen over five years in some of the shorter operating cycle partitions. Again, the explanatory power of earnings is higher than cash flows especially when comparing well-specified models. Section 5.3 reports that firms with shorter operating cycles have a linear returns-earnings relation over a ten-year measurement interval. It also reports that the returns-cash flows relation is primarily linear over this measurement interval. The higher explanatory power of earnings result is also present in this measurement interval.
5.1 One-Year Interval

Table 5 shows a summary of model results for all three regression models for the one-year returns-earnings relation. Model results for the one-year returns-cash flows relation are found in Table 6. The explanatory power is higher for earnings than for cash flows for most partitions in the OLS models. This can be seen from comparing the MSEs and R²'s for each partition. Results for the arctan model show that the explanatory power of earnings is higher than for cash flows in all partitions. Results from the single index model show that the explanatory power of earnings is greater than that of cash flows for some partitions, and the difference is even more pronounced than it is in the OLS or arctan comparison.

Results from the specification tests for the one-year returns-earnings models are found in Table 7. In general, the arctan model is a better descriptor of the relationship than the OLS model. This can be seen from the results of the regression of residuals from the OLS model on the predictions of the arctan model. The significant p-values reveal that OLS is not as good a descriptor of the returns-earnings relation as the arctan model. There is one partition where this is not the case. In the fourth operating cycle quartile the OLS model is not inadequate when compared to the arctan model.

_{Note that all of these results are reported using the calculation of returns described in Section 4.3. This return calculation is defined as the sum of dividends plus the change in price scaled by beginning price. The results do not materially differ if returns are calculated assuming dividends are reinvested and earn a return of 10% (see Easton, Harris, Ohlson [1992]). Results of this return calculation are not reported but are available from the author._

_{All specification tests use a one-tailed critical value of 0.025._
However, for firms in all partitions except one (short operating cycle), the single index model is the best descriptor of the returns-earnings relation over a short measurement period. This is as expected since it is hypothesized and empirically shown by Freeman and Tse [1992, 1994] among others that nonlinearities exist in the returns-earnings relation over short intervals. However, the arctan model employed by Freeman and Tse places more distributional restrictions on the relationship than the single index model. According to the results of the specification tests, the arctan model is not an adequate model specification relative to the less restrictive single index model.

Table 8 reports the results of the one-year returns-cash flows models' specification tests. These tests show that the relationship between cash flows and returns over a short measurement interval is nonlinear in the second and third quartiles. It also shows that in the fourth quartile that the relation between cash flows and returns can best be described as linear since OLS is an adequate model. This is also the case for firms with shorter operating cycles (Quartile 1).

A comparison of explanatory power of returns for earnings and cash flows for the one-year measurement interval are reported in Table 9. $R^2$s and MSEs are reported for the most restricted adequate model for the relations. Earnings consistently has higher explanatory power than cash flows. This is expected and is consistent with Dechow's [1994] results.
5.2 The Five-Year Measurement Interval Results

Returns-earnings model summaries for the five-year measurement interval are found in Table 10. Table 11 describes the model results for the five-year returns-cash flows relation. Again, comparing R^2s and MSEs for the OLS models reveals that in general earnings explains more of the variation in returns than cash flows does. This is true for both the arctan models and the single index models also.

For the five-year measurement interval, for all but one partition (third quartile), OLS and arctan models are adequate descriptors relative to each other of the returns-earnings relationship. This can be seen from the model specification tests reported in Table 12. For all but one partition (quartile 4), the single index is the most adequate descriptor of the returns-earnings relationship in the five-year measurement interval. Therefore, even after aggregating earnings over five years, the nonlinearities in the returns-earnings relation do not dampen significantly.

Model specification tests for the returns-cash flows relation are reported in Table 13. In this relation, there is some evidence that nonlinearities in the returns-cash flows relation dampen as the measurement window is extended. This is apparent because OLS is an adequate descriptor for two of the partitions. These partitions are the first and fourth quartiles of operating cycle length.

The relative explanatory power of earnings to cash flows is particularly high in comparing the well-specified models to each other. This result is exhibited in Table 14. This is an interesting result because it shows that both variable choice and model choice are important factors in explaining returns.
5.3 The Ten-Year Measurement Interval Results

Ten-year measurement interval model summaries are found in Table 15 for the returns-earnings relation and Table 16 for the returns-cash flows relation. Consistent with Dechow [1994], the explanatory power of earnings is higher than the explanatory power of cash flows. This is true even after aggregating both variables over ten years. As stated earlier, over the lifetime of a firm, these variables would be equal causing the explanatory power to be equal. However, a ten year period is not long enough to capture this. In addition, only operating cash flows are employed in the study instead of total cash flows.

Specification tests of the ten-year returns-earnings relation are found in Table 17. Note again that OLS is an adequate model for all partitions when compared to the arctan model. It is interesting to note here that the single index model best describes the returns-earnings relation for the third operating cycle partition. This means that even after the ten-year measurement period, the returns-earnings relation is not linear for all firms. The three partitions that show a linear relation (OLS being the most restricted model that adequately describes the relation) are firms with shorter operating cycles. This is as expected. However, it is also the case that firms with very long operating cycles also show a linear relation (quartile 4).

These results are slightly different from those of Anderson, et. al. [1996]. They show that OLS is the most restricted adequate model to describe the ten-year returns
earnings relation. However, their study does not partition the data on length of operating cycle.

Finally, as can be seen from Table 18, the returns-cash flows relation becomes linear as the cash flows are aggregated over a ten-year period. OLS is the most restricted adequate model describing the relation in all partitions except quartile 1.

Comparing the MSEs of the single index model in the returns-earnings relation with the MSEs of the OLS model in the returns-cash flows relation shows again that even after accounting for model specification, the earnings variable has a higher explanatory power than the cash flows variable.

Table 19 shows a comparison between earnings and cash flows of explanatory power for returns for the ten-year measurement interval. The explanatory power of earnings is still higher than cash flows, consistent with Dechow [1994]. However, when comparing well-specified models, the explanatory power level of earnings and cash flows does not converge to each other over the ten-year interval. Dechow finds after a four-year measurement period that the explanatory power becomes similar between earnings and cash flows. However, she only compares linear models.
CHAPTER 6

CONCLUSION

This study analyzes three choices that are necessary in attempting to explain returns. First, variable choice is studied, and two are employed. These are accounting earnings and operating cash flows. Second, model choice is studied. Three models are employed, each with differing levels of distributional restrictions. Finally, the measurement interval is analyzed. Earnings and cash flows are aggregated over one-, five-, and ten-year periods.

These three facets are interacted with each other in order to explain returns. The main results are as follows. First, the relation between returns and earnings is nonlinear over short measurement intervals. For firms with shorter operating cycles, the returns-earnings relation becomes linear over a ten-year aggregation period. Second, in most cases, the returns-cash flows relation is nonlinear in the shorter intervals. This relation becomes linear as the measurement interval is increased to ten years. Third, for an adequately specified model, the explanatory power of earnings is generally higher than the explanatory power of cash flows.

These results are important for studies of firm valuation and the effect of
non-earnings disclosures. Variable choice, model choice, and measurement interval are all important aspects of research design. A misspecified model may fail to adequately describe or summarize the data which causes problems when relying on the model statistics such as the coefficients. For future valuation and disclosure studies which analyze shorter measurement intervals, a less restricted model should be employed to determine if nonlinearities exist in the relation being analyzed. OLS appears to be a good model to employ in studies of long window relations between firm value and earnings or cash flows.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Operating Cycle Length Quartile 1</th>
<th>Operating Cycle Length Quartile 2</th>
<th>Operating Cycle Length Quartile 3</th>
<th>Operating Cycle Length Quartile 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Market Value of Common Equity at Beginning of Interval ($ in millions)</td>
<td>669.436</td>
<td>599.520</td>
<td>418.505</td>
<td>662.600</td>
</tr>
<tr>
<td>Total Book Value of Common Equity at Beginning of Interval ($ in millions)</td>
<td>91.069</td>
<td>64.282</td>
<td>65.175</td>
<td>95.965</td>
</tr>
<tr>
<td>Total Assets at Beginning of Interval ($ in millions)</td>
<td>2,013.676</td>
<td>3,185.857</td>
<td>1,235.079</td>
<td>1,833.578</td>
</tr>
<tr>
<td>Total Assets at Beginning of Interval ($ in millions)</td>
<td>1,098.875</td>
<td>696.423</td>
<td>658.428</td>
<td>1,056.583</td>
</tr>
<tr>
<td>Total Assets at Beginning of Interval ($ in millions)</td>
<td>127.146</td>
<td>94.284</td>
<td>91.019</td>
<td>125.181</td>
</tr>
<tr>
<td>Total Assets at Beginning of Interval ($ in millions)</td>
<td>3,288.488</td>
<td>2,896.119</td>
<td>2,897.130</td>
<td>3,066.039</td>
</tr>
<tr>
<td>Return over Interval</td>
<td></td>
<td>0.176</td>
<td>0.222</td>
<td>0.208</td>
</tr>
<tr>
<td>Sum of EPS (Including Extraordinary Items) over Interval</td>
<td></td>
<td>0.662</td>
<td>0.665</td>
<td>0.395</td>
</tr>
<tr>
<td>Sum of Cash Flows per Share over Interval</td>
<td></td>
<td>2.226</td>
<td>2.872</td>
<td>36.097</td>
</tr>
<tr>
<td>Sum of EPS over Interval Scaled by Price at Beginning of Interval</td>
<td></td>
<td>2.178</td>
<td>1.930</td>
<td>5.888</td>
</tr>
<tr>
<td>Sum of EPS over Interval Scaled by Price at Beginning of Interval</td>
<td></td>
<td>1.465</td>
<td>1.343</td>
<td>1.271</td>
</tr>
<tr>
<td>Sum of EPS over Interval Scaled by Price at Beginning of Interval</td>
<td></td>
<td>3.078</td>
<td>3.084</td>
<td>54.716</td>
</tr>
<tr>
<td>Sum of Cash Flows per Share scaled by Price at Beginning of Interval</td>
<td></td>
<td>0.056</td>
<td>0.066</td>
<td>0.058</td>
</tr>
<tr>
<td>Sum of Cash Flows per Share scaled by Price at Beginning of Interval</td>
<td></td>
<td>0.078</td>
<td>0.083</td>
<td>0.078</td>
</tr>
<tr>
<td>Sum of Cash Flows per Share scaled by Price at Beginning of Interval</td>
<td></td>
<td>0.193</td>
<td>0.214</td>
<td>0.275</td>
</tr>
<tr>
<td>Sum of Cash Flows per Share scaled by Price at Beginning of Interval</td>
<td></td>
<td>0.198</td>
<td>0.207</td>
<td>0.195</td>
</tr>
<tr>
<td>Sum of Cash Flows per Share scaled by Price at Beginning of Interval</td>
<td></td>
<td>0.160</td>
<td>0.163</td>
<td>0.155</td>
</tr>
<tr>
<td>Sum of Cash Flows per Share scaled by Price at Beginning of Interval</td>
<td></td>
<td>0.252</td>
<td>0.312</td>
<td>0.338</td>
</tr>
</tbody>
</table>

*Mean, *Median, *Standard Deviation

Table 1
Descriptive Statistics
One-Year Measurement Interval
(After Adjusting for Stock Splits and Stock Dividends; N=1,280 in each quartile)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Operating Cycle Length Quartile 1</th>
<th>Operating Cycle Length Quartile 2</th>
<th>Operating Cycle Length Quartile 3</th>
<th>Operating Cycle Length Quartile 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Market Value of Common Equity at Beginning of Interval ($ in millions)</td>
<td>528.930</td>
<td>424.115</td>
<td>368.763</td>
<td>548.462</td>
</tr>
<tr>
<td></td>
<td>82.581</td>
<td>52.441</td>
<td>46.731</td>
<td>78.571</td>
</tr>
<tr>
<td></td>
<td>1,652.966</td>
<td>1,772.524</td>
<td>1,206.055</td>
<td>1,544.844</td>
</tr>
<tr>
<td>Total Book Value of Common Equity at Beginning of Interval ($ in millions)</td>
<td>342.542</td>
<td>224.131</td>
<td>267.863</td>
<td>359.159</td>
</tr>
<tr>
<td></td>
<td>58.787</td>
<td>35.638</td>
<td>35.056</td>
<td>50.100</td>
</tr>
<tr>
<td></td>
<td>1,023.632</td>
<td>733.715</td>
<td>1,229.600</td>
<td>967.081</td>
</tr>
<tr>
<td>Total Assets at Beginning of Interval ($ in millions)</td>
<td>933.810</td>
<td>528.039</td>
<td>601.299</td>
<td>911.170</td>
</tr>
<tr>
<td></td>
<td>113.496</td>
<td>79.259</td>
<td>73.508</td>
<td>94.694</td>
</tr>
<tr>
<td></td>
<td>2,897.293</td>
<td>1,783.124</td>
<td>3,005.361</td>
<td>2,772.742</td>
</tr>
<tr>
<td>Return over Interval</td>
<td>1.242</td>
<td>1.627</td>
<td>1.206</td>
<td>1.416</td>
</tr>
<tr>
<td></td>
<td>0.650</td>
<td>0.892</td>
<td>0.621</td>
<td>0.732</td>
</tr>
<tr>
<td></td>
<td>3.618</td>
<td>2.492</td>
<td>1.879</td>
<td>2.406</td>
</tr>
<tr>
<td>Sum of EPS (Including Extraordinary Items) over Interval</td>
<td>3.311</td>
<td>3.326</td>
<td>1.977</td>
<td>3.667</td>
</tr>
<tr>
<td></td>
<td>3.116</td>
<td>2.640</td>
<td>2.762</td>
<td>2.515</td>
</tr>
<tr>
<td></td>
<td>6.900</td>
<td>10.093</td>
<td>67.146</td>
<td>12.495</td>
</tr>
<tr>
<td>Sum of Cash Flows per Share over Interval</td>
<td>10.891</td>
<td>9.649</td>
<td>29.440</td>
<td>12.216</td>
</tr>
<tr>
<td></td>
<td>11.020</td>
<td>12.387</td>
<td>229.490</td>
<td>20.151</td>
</tr>
<tr>
<td>Sum of EPS over Interval Scaled by Price at Beginning of Interval</td>
<td>0.478</td>
<td>0.648</td>
<td>0.504</td>
<td>0.512</td>
</tr>
<tr>
<td></td>
<td>0.435</td>
<td>0.533</td>
<td>0.489</td>
<td>0.445</td>
</tr>
<tr>
<td></td>
<td>0.670</td>
<td>0.883</td>
<td>0.979</td>
<td>0.697</td>
</tr>
<tr>
<td>Sum of Cash Flows Per Share Scaled by Price at Beginning of Interval</td>
<td>1.252</td>
<td>1.582</td>
<td>1.329</td>
<td>1.297</td>
</tr>
<tr>
<td></td>
<td>0.924</td>
<td>1.110</td>
<td>0.981</td>
<td>1.023</td>
</tr>
<tr>
<td></td>
<td>1.236</td>
<td>1.846</td>
<td>1.461</td>
<td>1.282</td>
</tr>
</tbody>
</table>

*Mean, *Median, *Standard Deviation

**Table 2**

Descriptive Statistics

Five-Year Measurement Interval

(After Adjusting for Stock Splits and Stock Dividends; N=256 in each quartile)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Operating Cycle Length Quartile 1</th>
<th>Operating Cycle Length Quartile 2</th>
<th>Operating Cycle Length Quartile 3</th>
<th>Operating Cycle Length Quartile 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Market Value of Common Equity at Beginning of Interval ($ in millions)</td>
<td>331.221</td>
<td>257.833</td>
<td>313.827</td>
<td>376.399</td>
</tr>
<tr>
<td></td>
<td>63.970</td>
<td>34.393</td>
<td>34.913</td>
<td>47.799</td>
</tr>
<tr>
<td></td>
<td>887.431</td>
<td>1,007.982</td>
<td>1,286.331</td>
<td>1,007.391</td>
</tr>
<tr>
<td>Total Book Value of Common Equity at Beginning of Interval ($ in millions)</td>
<td>251.798</td>
<td>156.795</td>
<td>229.568</td>
<td>280.808</td>
</tr>
<tr>
<td></td>
<td>44.648</td>
<td>30.212</td>
<td>28.810</td>
<td>31.233</td>
</tr>
<tr>
<td></td>
<td>792.912</td>
<td>478.995</td>
<td>1,158.844</td>
<td>769.706</td>
</tr>
<tr>
<td>Total Assets at Beginning of Interval ($ in millions)</td>
<td>700.163</td>
<td>362.857</td>
<td>464.549</td>
<td>659.254</td>
</tr>
<tr>
<td></td>
<td>89.226</td>
<td>67.567</td>
<td>58.356</td>
<td>71.745</td>
</tr>
<tr>
<td></td>
<td>2,248.503</td>
<td>1,158.28</td>
<td>2,427.527</td>
<td>1,872.942</td>
</tr>
<tr>
<td>Return over Interval</td>
<td>2.624</td>
<td>5.373</td>
<td>3.703</td>
<td>3.997</td>
</tr>
<tr>
<td></td>
<td>1.725</td>
<td>2.163</td>
<td>1.899</td>
<td>1.874</td>
</tr>
<tr>
<td></td>
<td>3.166</td>
<td>9.026</td>
<td>6.158</td>
<td>6.559</td>
</tr>
<tr>
<td>Sum of EPS (Including Extraordinary Items) over Interval</td>
<td>7.120</td>
<td>6.347</td>
<td>3.695</td>
<td>7.391</td>
</tr>
<tr>
<td></td>
<td>5.880</td>
<td>5.084</td>
<td>5.247</td>
<td>5.573</td>
</tr>
<tr>
<td></td>
<td>11.914</td>
<td>9.901</td>
<td>23.705</td>
<td>13.900</td>
</tr>
<tr>
<td>Sum of Cash Flows per Share over Interval</td>
<td>22.065</td>
<td>19.244</td>
<td>58.678</td>
<td>24.703</td>
</tr>
<tr>
<td></td>
<td>21.345</td>
<td>21.808</td>
<td>455.342</td>
<td>39.458</td>
</tr>
<tr>
<td>Sum of EPS over Interval Scaled by Price at Beginning of Interval</td>
<td>1.208</td>
<td>2.176</td>
<td>1.518</td>
<td>1.551</td>
</tr>
<tr>
<td></td>
<td>1.072</td>
<td>1.337</td>
<td>1.237</td>
<td>1.284</td>
</tr>
<tr>
<td></td>
<td>1.275</td>
<td>2.858</td>
<td>2.532</td>
<td>1.863</td>
</tr>
<tr>
<td>Sum of Cash Flows Per Share Scaled by Price at Beginning of Interval</td>
<td>3.070</td>
<td>5.030</td>
<td>3.588</td>
<td>3.729</td>
</tr>
<tr>
<td></td>
<td>2.545</td>
<td>3.281</td>
<td>2.800</td>
<td>2.817</td>
</tr>
<tr>
<td></td>
<td>2.722</td>
<td>6.026</td>
<td>3.393</td>
<td>3.121</td>
</tr>
<tr>
<td>Operating Cycle Length (in days)</td>
<td>58.500</td>
<td>109.729</td>
<td>155.363</td>
<td>337.759</td>
</tr>
<tr>
<td></td>
<td>60.548</td>
<td>110.188</td>
<td>154.259</td>
<td>230.639</td>
</tr>
<tr>
<td></td>
<td>18.965</td>
<td>13.300</td>
<td>13.985</td>
<td>983.458</td>
</tr>
</tbody>
</table>

*Mean, *Median, *Standard Deviation

Table 3
Descriptive Statistics
Ten-Year Measurement Interval
(After Adjusting for Stock Splits and Stock Dividends; N=128 in each quartile)
<table>
<thead>
<tr>
<th>First Digit of SIC Code</th>
<th>Parameter (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>9</td>
</tr>
</tbody>
</table>

Table 4

Regression Results of the Length of the Operating Cycle on the SIC Code

A dummy variable is used for each one digit SIC code in the regression.
<table>
<thead>
<tr>
<th></th>
<th>Book-to-Market Ratio</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Length of Operating Cycle</strong></td>
<td>0.0203 (0.6494)</td>
<td>0.0352 (0.4301)</td>
</tr>
</tbody>
</table>

Table 5
Correlation between Length of Operating Cycle and the Book-to-Market Ratio and Size of Firm
Pearson Correlation Coefficients
P-Values in Parentheses

Book-to-Market Ratio is the book value of the common stockholders' equity of the firm divided by the market value of the common equity of the firm; Size is the market value of the common shareholders' equity of the firm; Length of operating cycle is the number of days inventory is on hand plus the number of days accounts receivable are outstanding.
**Table 6**

Summary of Returns-Earnings Model Results - One-Year Measurement Interval

Operating Cycle Quartile Comparison

<table>
<thead>
<tr>
<th>Operating Cycle Length Quartiles</th>
<th>OLS</th>
<th>Arctan</th>
<th>Single Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept (p-value)</td>
<td>AESLP (p-value)</td>
<td>MSE R²</td>
</tr>
<tr>
<td>Quartile 1</td>
<td>0.1392 (0.0001)</td>
<td>0.6587 (0.0001)</td>
<td>0.188 0.078</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>0.1779 (0.0001)</td>
<td>0.6623 (0.0001)</td>
<td>0.224 0.081</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>0.1839 (0.0001)</td>
<td>0.4206 (0.0001)</td>
<td>0.316 0.040</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>0.1608 (0.0001)</td>
<td>0.8304 (0.0001)</td>
<td>0.277 0.081</td>
</tr>
</tbody>
</table>

*For Arctan and Single Index models, this is a pseudo-R².*

N=1,280 in each quartile
where:

\[
\text{Return} = \frac{\text{Change in price over the measurement interval plus the sum of dividends over the measurement interval all scaled by the price at the beginning of the measurement interval}}{	ext{Price at the beginning of the measurement interval}}
\]

\[
\text{AESLP (Aggregate Earnings Scaled by Lag Price)} = \frac{\text{Sum of primary earnings per share including extraordinary items over the measurement interval scaled by the price at the beginning of the measurement interval}}{	ext{Price at the beginning of the measurement interval}}
\]

\[
\text{BW} = \text{Bandwidth}
\]

Models:

\[
\text{OLS: } \text{Return} = \gamma_0 + \gamma_1 \times \text{AESLP}
\]

\[
\text{Arctan: } \text{Return} = \delta_0 + \delta_1 \times \arctan (\delta_2 \times \text{AESLP})
\]

\[
\text{Single Index: } \text{Return} = g (\alpha' \times \text{AESLP})
\]

Firms are partitioned into quartiles based on the length of the firm’s operating cycle. The first operating cycle quartile contains firms with the shortest operating cycle. Length of operating cycle is the number of days inventory is on hand plus the number of days accounts receivable are outstanding.
Table 7
Summary of Returns-Cash Flows Model Results -One-Year Measurement Interval
Operating Cycle Quartile Comparison

N=1,280 in each quartile
where:

\[
\text{Return} = \text{Change in price over the measurement interval plus the sum of dividends over the measurement interval all scaled by the price at the beginning of the measurement interval}
\]

\[
\text{ACFSLP (Aggregate Cash Flows Scaled by Lag Price)} = \text{Sum of operating cash flows per share over the measurement interval scaled by the price at the beginning of the measurement interval}
\]

\[
\text{BW} = \text{Bandwidth}
\]

Models:

\[
\text{OLS: } \text{Return} = \gamma_0 + \gamma_1 \cdot \text{ACFSLP}
\]

\[
\text{Arctan: } \text{Return} = \delta_0 + \delta_1 \cdot \arctan (\delta_2 \cdot \text{ACFSLP})
\]

\[
\text{Single Index: } \text{Return} = g (\alpha \cdot \text{ACFSLP})
\]

Firms are partitioned into quartiles based on the length of the firm's operating cycle. The first operating cycle quartile contains firms with the shortest operating cycle. Length of operating cycle is the number of days inventory is on hand plus the number of days accounts receivable are outstanding.
<table>
<thead>
<tr>
<th>Model Comparisons</th>
<th>Coeff. (p-value)</th>
<th>Coeff. (p-value)</th>
<th>Coeff. (p-value)</th>
<th>Coeff. (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Compared to Arctan</td>
<td>0.3134 (0.0001)</td>
<td>0.3181 (0.0001)</td>
<td>0.6350 (0.0001)</td>
<td>0.00064 (0.4745)</td>
</tr>
<tr>
<td>OLS Compared to Single Index</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.4010 (0.0001)</td>
</tr>
<tr>
<td>Arctan Compared to Single Index</td>
<td>0.0310 (0.3201)</td>
<td>0.2562 (0.0001)</td>
<td>0.2489 (0.0001)</td>
<td>0.4273 (0.0001)</td>
</tr>
</tbody>
</table>

* One-tailed p-values

Table 8
Specification Tests - Returns-Earnings Model
One-Year Measurement Interval - Operating Cycle Quartile Comparison
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(p-value)</td>
<td>(p-value)</td>
<td>(p-value)</td>
<td>(p-value)</td>
</tr>
<tr>
<td>OLS Compared to Arctan</td>
<td>0.0022</td>
<td>0.1224</td>
<td>0.3654</td>
<td>0.0612</td>
</tr>
<tr>
<td></td>
<td>(0.4929)</td>
<td>(0.0449)</td>
<td>(0.0004)</td>
<td>(0.2532)</td>
</tr>
<tr>
<td>OLS Compared to Single Index</td>
<td>0.1553</td>
<td>0.1689</td>
<td>N/A</td>
<td>0.0479</td>
</tr>
<tr>
<td></td>
<td>(0.0856)</td>
<td>(0.0058)</td>
<td></td>
<td>(0.3123)</td>
</tr>
<tr>
<td>Arctan Compared to Single Index</td>
<td>0.2587</td>
<td>0.0585</td>
<td>0.2385</td>
<td>0.0266</td>
</tr>
<tr>
<td></td>
<td>(0.0113)</td>
<td>(0.1887)</td>
<td>(0.0076)</td>
<td>(0.3926)</td>
</tr>
</tbody>
</table>

*One-tailed p-values*

Table 9
Specification Tests - Returns-Cash Flows Model
One-Year Measurement Interval - Operating Cycle Quartile Comparison
<table>
<thead>
<tr>
<th>Operating Cycle Length</th>
<th>Returns-Earnings</th>
<th>Returns-Cash Flows</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE R²</td>
<td>MSE R²</td>
</tr>
<tr>
<td>Quartile 1</td>
<td>0.176 0.139</td>
<td>0.191 0.064</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>0.196 0.203</td>
<td>0.212 0.134</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>0.283 0.150</td>
<td>0.308 0.075</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>0.264 0.129</td>
<td>0.278 0.078</td>
</tr>
</tbody>
</table>

Table 10

Comparison of Explanatory Power Between Adequate Models for the Returns-Earnings and Returns-Cash Flows Models

One-Year Measurement Interval
<table>
<thead>
<tr>
<th>Operating Cycle Length Quartiles</th>
<th>OLS</th>
<th>Arctan</th>
<th>Single Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>AESLP</td>
<td>MSE R²</td>
</tr>
<tr>
<td>Quartile 1</td>
<td>-0.2183</td>
<td>3.2417</td>
<td>8.824</td>
</tr>
<tr>
<td></td>
<td>(0.3656)</td>
<td>(0.0001)</td>
<td>0.326</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>0.3154</td>
<td>2.0112</td>
<td>2.921</td>
</tr>
<tr>
<td></td>
<td>(0.0214)</td>
<td>(0.0001)</td>
<td>0.530</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>0.7542</td>
<td>0.8973</td>
<td>2.741</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>0.224</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>0.3349</td>
<td>2.1203</td>
<td>3.697</td>
</tr>
<tr>
<td></td>
<td>(0.0345)</td>
<td>(0.0001)</td>
<td>0.361</td>
</tr>
</tbody>
</table>

*For Arctan and Single Index models, this is a pseudo-R².

Table 11
Summary of Returns-Earnings Model Results - Five-Year Measurement Interval
Operating Cycle Quartile Comparison

N=256 in each quartile
<table>
<thead>
<tr>
<th>Operating Cycle Length Quartiles</th>
<th>OLS</th>
<th>Arctan (δ₂) *</th>
<th>Single Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept (p-value)</td>
<td>ACFSLP (p-value)</td>
<td>MSE R²</td>
</tr>
<tr>
<td>Quartile 1</td>
<td>-0.2354 (0.4752)</td>
<td>1.2565 (0.0001)</td>
<td>11.290</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>0.3857 (0.0251)</td>
<td>0.7731 (0.0001)</td>
<td>4.083</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>0.4296 (0.0031)</td>
<td>0.5797 (0.0001)</td>
<td>2.799</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>0.2000 (0.3007)</td>
<td>0.9539 (0.0001)</td>
<td>4.311</td>
</tr>
</tbody>
</table>

*For Arctan and Single Index models, this is a pseudo-R².

Table 12
Summary of Returns-Cash Flows Model Results - Five-Year Measurement Interval
Operating Cycle Quartile Comparison

N=256 in each quartile
### Operating Cycle Length

<table>
<thead>
<tr>
<th>Model Comparisons</th>
<th>First Quartile</th>
<th>Second Quartile</th>
<th>Third Quartile</th>
<th>Fourth Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Compared to</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arctan</td>
<td>-0.0011</td>
<td>0.0072</td>
<td>0.2774</td>
<td>-0.0006</td>
</tr>
<tr>
<td></td>
<td>(0.5039)</td>
<td>(0.4524)</td>
<td>(0.0017)</td>
<td>(0.4973)</td>
</tr>
<tr>
<td>OLS Compared to</td>
<td>-0.3228</td>
<td>0.0409</td>
<td>N/A</td>
<td>0.1704</td>
</tr>
<tr>
<td>Single Index</td>
<td>(0.9138)</td>
<td>(0.2460)</td>
<td></td>
<td>(0.0257)</td>
</tr>
<tr>
<td>Arctan Compared to</td>
<td>0.0478</td>
<td>0.0461</td>
<td>0.1521</td>
<td>0.2086</td>
</tr>
<tr>
<td>Single Index</td>
<td>(0.4206)</td>
<td>(0.2180)</td>
<td>(0.0422)</td>
<td>(0.0085)</td>
</tr>
</tbody>
</table>

* One-tailed p-values

Table 13  
Specification Tests - Returns-Earnings Model  
Five-Year Measurement Interval - Operating Cycle Quartile Comparison
### Operating Cycle Length

<table>
<thead>
<tr>
<th>Model Comparisons</th>
<th>Coeff. (p-value)</th>
<th>Coeff. (p-value)</th>
<th>Coeff. (p-value)</th>
<th>Coeff. (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Compared to Arctan</td>
<td>0.0020 (0.4951)</td>
<td>0.0527 (0.2719)</td>
<td>0.1549 (0.0820)</td>
<td>0.0066 (0.4767)</td>
</tr>
<tr>
<td>OLS Compared to Single Index</td>
<td>-0.5546 (0.9688)</td>
<td>0.1526 (0.0360)</td>
<td>0.2163 (0.0219)</td>
<td>0.0115 (0.4638)</td>
</tr>
<tr>
<td>Arctan Compared to Single Index</td>
<td>-0.5288 (0.9622)</td>
<td>0.0871 (0.1488)</td>
<td>0.0642 (0.2688)</td>
<td>0.0178 (0.4440)</td>
</tr>
</tbody>
</table>

* One-tailed p-values

Table 14
Specification Tests - Returns-Cash Flows Model
Five-Year Measurement Interval - Operating Cycle Quartile Comparison
<table>
<thead>
<tr>
<th>Operating Cycle Length</th>
<th>MSE</th>
<th>R²</th>
<th>Operating Cycle Length</th>
<th>MSE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quartile 1</td>
<td>8.824</td>
<td>0.326</td>
<td>Quartile 1</td>
<td>11.290</td>
<td>0.137</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>2.921</td>
<td>0.530</td>
<td>Quartile 2</td>
<td>4.083</td>
<td>0.343</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>2.323</td>
<td>0.348</td>
<td>Quartile 3</td>
<td>2.632</td>
<td>0.261</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>3.542</td>
<td>0.393</td>
<td>Quartile 4</td>
<td>2.632</td>
<td>0.261</td>
</tr>
</tbody>
</table>

Table 15

Comparison of Explanatory Power Between Adequate Models for the Returns-Earnings and Returns-Cash Flows Models

Five-Year Measurement Interval
<table>
<thead>
<tr>
<th>Operating Cycle Length Quartiles</th>
<th>OLS</th>
<th></th>
<th></th>
<th>Arctan(δ₂*)</th>
<th>AESLP</th>
<th>Arctan(δ₂)</th>
<th>AESLP</th>
<th>MSE</th>
<th>R²</th>
<th>AESLP (p-value)</th>
<th>MSE</th>
<th>R²</th>
<th>BW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept (p-value)</td>
<td>AESLP (p-value)</td>
<td>MSE R²</td>
<td>Intercept (p-value)</td>
<td>Arctan(δ₂*) (p-value)</td>
<td>AESLP(δ₂) (p-value)</td>
<td>MSE R²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quartile 1</td>
<td>0.8887 (0.0081)</td>
<td>1.4376 (0.0001)</td>
<td>6.968</td>
<td>0.954 (0.0254)</td>
<td>1,233.9845 (1.0000)</td>
<td>0.0011 (1.0000)</td>
<td>15.652</td>
<td>0.537</td>
<td>1.8671 (0.0000)</td>
<td>7.047</td>
<td>0.318</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quartile 2</td>
<td>-0.5534 (0.2648)</td>
<td>2.7404 (0.0001)</td>
<td>19.933</td>
<td>0.0762 (0.9047)</td>
<td>212.7594 (0.9284)</td>
<td>0.0098 (0.9284)</td>
<td>24.294</td>
<td>0.703</td>
<td>2.2028 (0.0000)</td>
<td>21.805</td>
<td>0.729</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>Quartile 3</td>
<td>1.1237 (0.0156)</td>
<td>1.6537 (0.0001)</td>
<td>19.800</td>
<td>0.6184 (0.2778)</td>
<td>23.1537 (0.0134)</td>
<td>0.0889 (0.0584)</td>
<td>19.173</td>
<td>0.489</td>
<td>1.8748 (0.0000)</td>
<td>16.658</td>
<td>0.549</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>Quartile 4</td>
<td>0.2234 (0.6847)</td>
<td>2.4424 (0.0001)</td>
<td>22.358</td>
<td>0.3030 (0.6110)</td>
<td>143.3289 (0.9602)</td>
<td>0.0150 (0.9602)</td>
<td>23.018</td>
<td>0.475</td>
<td>2.3884 (0.0000)</td>
<td>22.395</td>
<td>0.481</td>
<td>1.75</td>
</tr>
</tbody>
</table>

*For Arctan and Single Index models, this is a pseudo-$R^2$.

Table 16
Summary of Returns-Earnings Model Results - Ten-Year Measurement Interval
Operating Cycle Quartile Comparison

N=128 in each quartile
### Table 17
Summary of Returns-Cash Flows Model Results - Ten-Year Measurement Interval
Operating Cycle Quartile Comparison

<table>
<thead>
<tr>
<th>Operating Cycle Length Quartiles</th>
<th>OLS</th>
<th>Arctan</th>
<th>Single Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept (p-value)</td>
<td>ACFSLP (p-value)</td>
<td>MSE R²</td>
</tr>
<tr>
<td>Quartile 1</td>
<td>0.8647 (0.0277)</td>
<td>0.5650 (0.0001)</td>
<td>7.970</td>
</tr>
<tr>
<td></td>
<td>0.0213 (0.9763)</td>
<td>1.0711 (0.0001)</td>
<td>39.364</td>
</tr>
<tr>
<td></td>
<td>-0.7383 (0.2065)</td>
<td>1.2246 (0.0001)</td>
<td>20.295</td>
</tr>
<tr>
<td></td>
<td>-0.0972 (0.9003)</td>
<td>1.1077 (0.0001)</td>
<td>31.200</td>
</tr>
</tbody>
</table>

*a*For Arctan and Single Index models, this is a pseudo-R².

N=128 in each quartile
<table>
<thead>
<tr>
<th>Model Comparisons</th>
<th>Coeff. (p-value)(^a)</th>
<th>Coeff. (p-value)</th>
<th>Coeff. (p-value)</th>
<th>Coeff. (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Compared to Arctan</td>
<td>0.0000 (0.5000)</td>
<td>-0.0004 (0.5021)</td>
<td>0.0302 (0.3727)</td>
<td>-0.0002 (0.5008)</td>
</tr>
<tr>
<td>OLS Compared to Single Index</td>
<td>0.0773 (0.3012)</td>
<td>0.0295 (0.2929)</td>
<td>0.2426 (0.0004)</td>
<td>0.0714 (0.2105)</td>
</tr>
<tr>
<td>Arctan Compared to Single Index</td>
<td>0.1544 (0.1489)</td>
<td>0.2741 (0.0001)</td>
<td>0.2320 (0.0006)</td>
<td>0.1806 (0.0295)</td>
</tr>
</tbody>
</table>

\(^a\) One-tailed p-values

Table 18
Specification Tests - Returns-Earnings Model
Ten-Year Measurement Interval - Operating Cycle Quartile Comparison
### Operating Cycle Length

<table>
<thead>
<tr>
<th>Model Comparisons</th>
<th>First Quartile</th>
<th>Second Quartile</th>
<th>Third Quartile</th>
<th>Fourth Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Compared to Arctan</td>
<td>-0.0023</td>
<td>0.0185</td>
<td>-0.0017</td>
<td>-0.0010</td>
</tr>
<tr>
<td></td>
<td>(0.5059)</td>
<td>(0.4150)</td>
<td>(0.5042)</td>
<td>(0.5021)</td>
</tr>
<tr>
<td>OLS Compared to Single Index</td>
<td>0.2894</td>
<td>-0.0229</td>
<td>0.1470</td>
<td>-0.1903</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.6126)</td>
<td>(0.2033)</td>
<td>(0.8164)</td>
</tr>
<tr>
<td>Arctan Compared to Single Index</td>
<td>0.6260</td>
<td>-0.0539</td>
<td>0.7442</td>
<td>0.1467</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.7523)</td>
<td>(0.0001)</td>
<td>(0.2460)</td>
</tr>
</tbody>
</table>

*a One-tailed p-values

Table 19
Specification Tests - Returns-Cash Flows Model
Ten-Year Measurement Interval - Operating Cycle Quartile Comparison
### Table 20

Comparison of Explanatory Power Between Adequate Models for the Returns-Earnings and Returns-Cash Flows Models

<table>
<thead>
<tr>
<th></th>
<th>Returns-Earnings MSE</th>
<th>Returns-Cash Flows MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>R²</td>
</tr>
<tr>
<td>Operating Cycle Length</td>
<td>6.968</td>
<td>7.770</td>
</tr>
<tr>
<td>Quartile 1</td>
<td>0.326</td>
<td>0.768</td>
</tr>
<tr>
<td>Operating Cycle Length</td>
<td>19.933</td>
<td>39.364</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>0.752</td>
<td>0.510</td>
</tr>
<tr>
<td>Operating Cycle Length</td>
<td>16.658</td>
<td>20.295</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>0.549</td>
<td>0.450</td>
</tr>
<tr>
<td>Operating Cycle Length</td>
<td>22.358</td>
<td>31.200</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>0.482</td>
<td>0.277</td>
</tr>
</tbody>
</table>
APPENDIX B

SEMIPARAMETRIC REGRESSION DISCUSSION
This appendix has been borrowed with permission from Anderson, Schroeder, and Shroff [1997]. It is intended to provide background into semiparametric regression and the techniques described in the dissertation.

This appendix synthesizes some of the literature on semiparametric and nonparametric kernel regression and its application to accounting data. Since estimation programs are not widely available, this appendix is intended to clarify the approach and aid in replication of the analysis.

Frequently in econometric analysis of accounting data, one is concerned with departures from standard parametric distributional model assumptions. Semi- and nonparametric methods provide an alternative means (to parametric models) to characterize data and assess parametric model robustness. Here, we focus on regression. That is, we examine the conditional relation between $y$ and $x$. The most flexible 'fit' of this conditional relation is nonparametric regression. The discussion that follows is a synthesis of numerous econometric studies as applied in this paper.

Nonparametric (kernel) regression

A nonparametric (kernel) regression can be represented as follows (Hardie [1990]).

$$E[y|x] = m(x)$$

$$m(x) = \frac{n^{-1}h^{-d} \sum_{i=1}^{n} K(\frac{x - X_i}{h}) Y_i}{n^{-1}h^{-d} \sum_{i=1}^{n} K(\frac{x - X_i}{h})}, \quad Y_i(X_i) \text{ is the ith observation for } y(x), \quad n \text{ is the}$$

---

16 Flexible fit refers to the imposition of few distributional restrictions on the data in estimating the relationship.

15 Nonparametric regression is motivated by at least the following four objectives: (1) it provides a versatile method for exploring a general relation between variables, (2) it give predictions without reference to a fixed parametric model, (3) it provides a tool for identifying spurious observations, and (4) it provides a method for 'fixing' missing values or interpolating between regressor values (see Hardle [1990, pp 6-7]).
number of observations, d is the dimension (number of regressors) of x, K() is any well-defined (multivariate) kernel, and h is the smoothing parameter or bandwidth.

A multivariate kernel is constructed, row by row, by computing the product of marginal densities for each variable in the design matrix x. That is,

\[ h^{-d} K\left( \frac{x - X_j}{h} \right) = \prod_{j=1}^{d} h^{-1} K\left( \frac{X_{j, \text{th column vector in the design matrix}}}{h} \right), \]

where our kernels are constructed on a 'leave-one-out' basis. That is, the current observation is excluded in the kernel construction to avoid overfitting. Since nonparametric regression simply exploits the explanatory variables to devise a weighting scheme for y, assigning no weight to the current observation on y is an intuitively appealing means of avoiding overfitting.

The bandwidth is determined by generalized cross-validation (Craven and Wahba [1979]) by choosing the bandwidth that minimizes the mean squared errors (MSE), for the 'leave-one-out' nonparametric regression estimator,

\[ \min_{h} \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[ Y_i - m(X_i) \right]^2 \right\}. \]

Since the properties of this statistic are data specific, convergence at a uniform rate cannot be assured. Thus, one evaluates a dense grid of values for h and derives the minimum MSE numerically.

Nonparametric (kernel) regression is the most flexible model that we employ and forms the basis for our other estimators. While nonparametric regression models provide a very flexible fit of the relation between y and x, this does not come at zero cost. In particular, it is more difficult to succinctly describe this relation, especially when x is a

---

16 We use the standard normal density in all of the applications. Hardle [1990, pp. 165-173] presents evidence that the choice of the kernel density has little impact on the adequacy of nonparametric regression models but that the bandwidth choice is critical.

17 By convention, whenever the numerator and denominator are zero, the predicted value for \( m(x) \) is defined to be zero.

18 Since we estimate one bandwidth for all regressors, the variables are first comparably scaled by their estimated standard deviation.

19 'Optimal' bandwidths are determined by trading off a 'good approximation' to the regression function (reduction in bias) and a 'good reduction' of observational noise (reduction in noise). The mean squared error is composed of squared bias and variance. The former (latter) is increasing (decreasing) in the bandwidth (Hardle [1990, pp. 29-30, 149]).
high dimension matrix. We next turn to models that retain some of the flexibility of nonparametric regression but enhance interpretability.

Partial linear models

Frequently, one is concerned about the relation between $y$ and $x$ but troubled that the analysis is plagued by omitted, correlated variables. One difficulty is that we do not know the form of the relation between our variables of interest and these other variables. Provided that we can observe these 'control' variables, Robinson [1988] suggests a two-stage approach which we call partial linear regression. Partial linear regression models involve nonparametrically fitting the relation between the dependent variable, $y$, and the control variables, $z$, and also the experimental regressors of interest, $x$, and the control variables. The predicted values are then subtracted from the observed dependent and experimental regressors to form their conditional mean differenced variables, $y - E[y|z]$ and $x - E[x|z]$.

Next, one simply employs no-intercept OLS regression of the differenced dependent variable on the differenced experimental regressor variables,

\[
y - E[y|z] = \beta (x - E[x|z]).
\]

The parameter estimator for $\beta$ fully captures the influence of the otherwise omitted, control variables and is accordingly, asymptotically consistent. Of course, we now have parameters to succinctly describe the relation between $y$ and $x$ conditional on $z$.

---

20 Another cost is that nonparametric regressions typically do not achieve parametric rates of convergence (i.e., square root $n$). Rates of convergence are a more technical subject than is intended for this note but under certain regulatory conditions, it can be shown that optimal rates of convergence for nonparametric models are $n^r$, $0 < r < 1/2$; more specifically, $r = (p+\beta-k)/(2[p+\beta]-d)$, where $p$ is the number of times the smoothing function is differentiable, $k$ is the order of the derivative of the particular estimate of interest ($k*p$), $\beta$ is the characteristic or exponent for the smoothness class, and $d$ is the order of the regressors (Hardle [1990, p. 93]).

21 Robinson demonstrates that this estimator converges at the parametric rate $r = n^{-1/2}$. 74
Single-index regression

The partial linear model discussed above imposes distributional restrictions on the relation between y and x in the second stage. One (semiparametric) approach for relaxing this restriction and retaining ease of interpretability is single-index regression. Single-index regression follows from a remarkable result. The average derivative of a general function with respect to the regressor is proportional to the parameters of the index.

Suppose that \( E[y|x] = G(x\beta) \), then define \( \delta = \frac{\partial E[y|x]}{\partial x} = dG/d(x\beta) \beta = \gamma(x\beta) \beta \). Thus, the derivative with respect to x is proportional to \( \beta \) for all x, and likewise the average derivative \( E[dG/d(x\beta)]\beta = \gamma \beta \), for \( \gamma \neq 0 \), is proportional to \( \beta \).

Our applications employ the density-weighted average derivative single-index model of Powell, Stock, and Stoker [1989]. That is,

\[
\hat{\delta} = -2n^{-1} \sum_{i=1}^{n} \frac{\partial E(x_i)}{\partial x_i} y_i = -2[n(n-1)]^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} h^{-d(x\beta)} K\left( \frac{x_i - x_j}{h} \right) \left( y_i - y_j \right).
\]

For a Gaussian kernel, \( K(u) \), notice that \( K'(u) = -u K(u) \). Thus,

\[
\hat{\delta} = 2[n(n-1)]^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} h^{-d(x\beta)} K\left( \frac{x_i - x_j}{h} \right) \left( y_i - y_j \right),
\]

where \( K(u) = (2\pi)^{-1/2} \exp\{-u^2/2\} \).

The asymptotic covariance matrix for the parameters \( \Sigma_\delta \) is estimated as

\[
\hat{\Sigma}_\delta = 4n^{-1} \sum_{i=1}^{n} \hat{r}(z_i) \hat{r}(z_i)^T - 4\hat{\delta} \hat{\delta}^T,
\]

where

\[
\hat{r}(z_i) = -(n-1)^{-1} \sum_{j=1}^{n} h^{-d(x\beta)} K\left( \frac{x_i - x_j}{h} \right) \left( y_i - y_j \right).
\]

The above estimator is proportional to the index parameters. Powell et al also proposed a properly scaled instrumental variable version of the density-weighted average derivative. We refer to this estimator as \( \hat{d} = \delta^{-1} \hat{\delta} \), where

---

21 Powell et al.'s description of the asymptotic properties of their average derivative estimator exploits a 'leave-one-out' approach, as discussed for nonparametric regression above. This estimator also achieves the parametric rate of convergence, \( r = n^{-1} \).

22 We thank Berwin Turlach for his assistance in pointing out the computational efficiency of exploiting the 'U-statistic' structure in this fashion.
The asymptotic covariance estimator for the parameters $\Sigma_\delta$ is estimated as

$$\hat{\Sigma}_d = 4n^{-1} \sum_{i=1}^{n} \hat{r}_d(z_i)\hat{r}_d(z_i)'$$

where

$$\hat{r}_d(z_i) = \hat{\delta}_x^{-1} \left\{ -(n-1)^{-1} \sum_{j=1}^{n} h^{-(d+1)} K\left( \frac{x_i - x_j}{h} \right) \left( \frac{x_i - x_j}{h} \right) (\hat{u}_i - \hat{u}_j) \right\}, \hat{u}_i = y_i - x_i \hat{\delta}.$$

The instrumental variables version is reported in our results. The optimal bandwidth is estimated similarly to that described above for nonparametric regression. First, $\hat{d}$ (and its covariance matrix) is estimated. Then, nonparametric regression predictions of $y$ on $x \hat{d}$ are estimated; this is the counterpart to regressing $y$ on $x$ in fully nonparametric regression. This provides a readily interpretable, flexibly fit index parameter, the counterpart to the slope parameter in OLS regression.

**Partial index regression models**

Now, we put together the last two sets of ideas. That is, nonparametric estimates for potentially omitted, correlated (control) variables as in the partial linear model are combined with single index model parameter estimates for the experimental regressors. Following Stoker [1991], these are called partial index models. As with partial linear models, the relation between $y$ and $z$ (the control variables) and the relation between $x$ and $z$ are estimated via nonparametric regression. Separate bandwidths are employed for the regression of $y$ and $x$ on $z$. Again, their conditional mean differences are computed, $(y - E[y|x])$ and $(x - E[x|z])$. Now, single index regression of $(y - E[y|x])$ on $(x - E[x|z])$ completes the partial index regression. Notice, that a third round of bandwidth selection is involved in the second stage.
**Specification testing of restricted models against a nonparametric regression benchmark**

The appendix is concluded with a brief discussion of the use of the above models in specification testing. In particular, specification testing of restricted models against a general nonparametric regression benchmark (Stoker[1991], Rodriguez and Stoker [1992]). Stoker suggests that specification tests of the restricted models can be conducted in the following simple and intuitive manner. Residuals from the restricted model (r) are regressed against predictions from a more general model (g) via standard OLS regression, $y - E[y|x] = \alpha + \lambda E[y|x] + \varepsilon.$ If standard t tests against a null hypothesis of $\lambda \leq 0$ indicate that $\lambda$ is positive, then the restricted model is rejected. One can evaluate a variety of restricted models against a variety of more general models. An adequate restricted model is preferred for ease of interpretation.24

---

24 If one knew the proper specification, it would be the model of choice on the grounds of statistical efficiency.
BIBLIOGRAPHY


Freeman, R.N. and S.Y. Tse, 1994, Price responses to quarterly earnings: Return interval, forecast source, and firm size, working paper, University of Texas-Austin.


Rodriguez, D. And T.M. Stoker, 1992, Semiparametric analysis of environmental effects, working paper, Sloan School of Management, MIT.

