Estimating the Amount of Estimation in Accruals

Jason V. Chen
University of Michigan
Ross School of Business
jvchen@umich.edu

Feng Li
University of Michigan
Ross School of Business
feng@umich.edu

November 5, 2012

Preliminary. Comments welcome.

Abstract

This paper examines the link between the amount of estimation needed during the accrual generating process and the persistence of the accruals portion of earnings. We measure the amount of estimation needed during the accrual generating process using the number of estimation-related linguistic cues in the notes to the financial statements. Consistent with the conjectures in Sloan [1996], we find that accruals that needed more estimation are less predictive of future earnings. We also find that such accruals map less into the past, current, or future cash flows in the sense of Dechow and Dichev [2002]. When we disaggregate the number of estimation linguistic cues into a component that is due to the existence of specific accruals accounts and a component that is due to the within-accounts variations, we find that our results are driven by both components. Lastly, we find mixed evidence as to whether the amount of estimation in accruals is systematically associated with the accrual anomaly. Overall, our results suggest that the estimation needed during the accrual generating process plays an important role in understanding the persistence of accruals.

We thank Matt DeAngelis, Ilia Dichev, Scott Richardson, Lakshmanan Shivakumar, Irem Tuna, and the workshop participants at the SUNY-Buffalo, London Business School, University of Michigan, Michigan State University, and Rice University for their comments. We gratefully acknowledge the support of the Ross School of Business and the financial support from the Paton Accounting Fellowship and the Harry Jones Endowment.

1. Introduction

We propose a new approach to measure the quality of accruals based on the amount of estimation embedded in the accruals, calculated as the number of estimation-related linguistic cues in the notes to the financial statements. The notes to the financial statements contain detailed descriptions of the nature of the accruals and how they are generated and are therefore likely to be informative about accruals quality.

We examine the association between this new measure of accruals quality and how predictive the accruals are of future earnings. Sloan [1996] and Richardson et al. [2005] argue that greater estimation needed during the accrual generating process explains why accruals are less persistent than the cash portion of earnings (cash flows). Unlike cash flows which need little estimation, accruals incorporate estimates of future cash flows, cash flow deferrals, depreciation and amortization, and fair value estimates. Richardson et al. [2005] argues that the greater estimation suggests that accruals are recognized with lower precision and therefore that is why they are less predictive of the future earnings.

While this argument is a well-accepted conjecture few studies have explicitly examined how the estimation involved during the accrual generating process relates to the persistence of accruals (Dechow et al. [2010]). Rather, prior research has primarily examined how the different components that comprise accruals explain its persistence. For example, Dechow and Ge [2006] finds that low accruals exhibit lower persistence when they contain special items. Richardson et al. [2005] examines the difference in persistence of working capital, non-current operating, and financial asset accruals, as these different types of accruals inherently require different levels of estimation. We complement the evidence in these studies by linking accruals persistence directly to the characteristics of the accruals generating process.

We start out by reading numerous notes to the financial statements to find linguistic cues that convey that some estimation was needed. We find that in general, these linguistic cues can

be broken down into three general types: 1. An estimation action targets some object (e.g. "we *estimated* receivables"), 2. An estimation object is the target of a use action (e.g. "we used *estimates*"), 3. An estimation word is an adjective to an object (e.g. "*estimated* costs") (Appendix 1 & 2). We then use statistical parsing techniques to automate the search for such linguistic cues in the footnotes section of firms' 10-K filings (Appendix 1).

Our textual analysis approach is used to capture the amount of estimation needed when the accruals are recognized. We focus on the notes to the financial statements section of the 10-K because it provides information specific to the accounting process. This section of the 10-K provides a wealth of information not found in other sections of the 10-K filing (Merkeley [2011], Riedl and Srinivasan [2010]). More importantly, this section provides information pertaining to the estimations made and the assumptions needed by management during the accrual generating process. While other sections of the 10-K, such as the Management Discussion and Analysis, may also provide some information pertaining to the accrual generating process it is typically minute and usually embedded within other pieces of information unrelated to the accounting process. Moreover, when there is accounting related information it is typically reiterated in the notes to the financial statements.

We first hypothesize that accruals that require more estimation are less predictive of future earnings. Consistent with this hypothesis, we find that accruals are significantly less persistent when there is greater estimation conveyed in the company's footnotes. In contrast, we find that the amount of estimation conveyed in the company's footnotes is not informative of the persistence of cash flows. These findings support the conjecture that the amount of estimation needed in accruals partially explains their lower persistence (Sloan [1996]).

We also find that accruals that involve more estimation also have lower quality measured in the sense of Dechow and Dichev [2002], i.e., these accruals map less into past, current, or future cash flows. This is consistent with the hypothesis that more highly estimated accruals are

less precise, and therefore have greater errors, and provides further evidence that the estimation needed during the accrual generating process drives the lower persistence.

A natural question arises whether the number of estimation cues simply captures the types of accruals accounts. For instance, firms that have defined benefits pension plans or employee stock options are likely to have more estimation cues in the footnote because pension and stock option calculations need to estimate more parameters. We therefore decompose the number of estimation cues in the footnotes into two components: a component that is due to the existence of specific accruals accounts and a component that is due to the within-accounts variations. Specifically, we identify 49 common footnote items (see footnote 5) and regress the number of estimation cues on the item fixed effects. The predicted value from this model captures the number of estimation cues due to the existence of different accruals accounts and the residual captures the within-accounts variations. We find that the persistence accruals and the quality of accruals are driven by both components of estimation cues.

Lastly, we examine the accrual anomaly as a function of the estimation cues in the footnotes. We find mixed evidence as to whether the market reacts as if it does not incorporate the amount of estimation in accruals into its valuation of the firm in a timely manner. Specifically, we find some evidence that the accrual anomaly documented in prior studies (Sloan [1996], Xie [2001], Mashruwala et al. [2006]) is more significant when more estimation is needed during the accrual generating process.[1] This is consistent with the hypothesis that the market more greatly overvalues (undervalues) highly estimated positive (negative) accruals in the short term. However, the robustness tests of our findings using the Carhart Four-Factor Alpha model provide mixed results.

This paper makes several contributions to the literature. First, this study incorporates the estimation needed during the accrual generating process as reflected in a firm's notes to the

[1] Our study uses one-year abnormal returns which begin 5 days after the 10-K filing.

financial statements into our understanding of the persistence of accruals. Many prior studies have ignored this important aspect of accruals and have simply focused on the accrual numbers themselves to explain the persistence of accruals.[2]

Second, our findings strengthen Sloan's argument that the estimation involved in accruals explains the difference in the persistence of the cash portion of earnings and the accruals portion of earnings. Some prior studies argue that the difference in the persistence of accruals and cash flows documented by Sloan is driven by omitted fundamental differences such as growth. Our findings suggest that estimation does partially explain the lower persistence of accruals.

Finally, this study contributes to the textual analysis accounting literature by using grammatical relationships to extract meaning from qualitative financial information. These relationships provide structure to the qualitative information and allow us to better infer meaning from the text. Additionally, this study adds to a growing field of textual analysis studies which suggest that using qualitative accounting information, in conjunction with quantitative accounting information, helps to provide a richer understanding of firms and their accounting process (Li [2011]).

The remainder of the paper proceeds as follows. Section 2 provides a discussion of prior literature and motivation for our hypotheses. In Section 3 we discuss how the sample of 10-K footnotes and financial information was prepared. In Section 4 we present the research design and main results. Lastly, Section 5 concludes the paper.

2. Prior Literature and Hypotheses

Sloan [1996] finds that the accruals portion of earnings is less predictive of future earnings than the cash portion earnings (i.e. accruals are less persistent than cash flows). He argues that the difference between the persistence of accruals and cash flows is due to the greater

[2] Richardson et al. [2005] indirectly test the association between accruals estimation and persistence.

estimation needed when deriving accruals. Accruals incorporate estimates of future cash flows, depreciation and allocations, deferrals, and valuations that are not needed when recording cash flows (Richardson et al. [2005]). These estimations make accruals less precise and therefore less persistent than cash flows.

Richardson et al. [2005] expands upon the hypothesis in Sloan [1996] and formally models accruals estimation as an error-in-variables problem. Their model assumes that if accruals can be measured without error, there is no need to correct for accrual errors in future periods. However, actual recorded accruals are measured with error since managerial estimation is needed during the accrual generating process. This error reduces the association between accruals and future earnings.

The estimation needed in the accruals portion of earnings is a function of the process and estimations needed by management when recording the accrual and not necessarily a reflection of the magnitude of the accruals. Even if two companies have the same total amount of accruals, the amount of estimation in the accruals of these two companies may be vastly different. Simply examining the total amount of accruals will not provide information about the firm's accrual generating process nor will it provide information about the estimation of accruals. For example, one company's total accruals may contain a large amount of estimated fair value accruals while another may contain a large amount of financial accruals involving less estimation. Even if the total amount of accruals is the same, the degree of estimation between the two company's accruals may be vastly different.

Prior studies have examined how the specific components that comprise accruals affect how well accruals predict future earnings. For instance, Dechow and Ge [2006] examines the persistence of low accruals when the firm has special items. Consistent with their hypothesis, they find that accruals are less persistent when the firm has special items. Richardson et al. [2005] disaggregates accruals into financing accruals, working capital accruals, and non-current

operating accruals. They posit that the accruals in each of the three categories have different degrees of estimation. Financial accruals require less estimation than working capital or non-current operating accruals because their terms are typically contractually defined. Therefore estimates of future cash flows are well defined and require a lower degree of estimation. On the other hand, estimates of future cash flows, valuations, and other estimates are needed when recording working capital and non-current operating accruals. The greater estimation in these accruals implies that these accruals are less likely to be realized in the cash flows and therefore will be less informative of future earnings. Consistent with their hypothesis, they find that financing accruals are more persistent than working capital and non-current operating accruals.

We fill the gap in the literature by explicitly measuring and examining the implications of accruals estimation on accruals persistence. Even though prior studies have provided interesting insight into the persistence of accruals, few have directly examined this important aspect of accruals. Accordingly, our first hypothesis follows the conjectures of Sloan [1996] and Richardson et al. [2005] and is as follows:

Prediction 1: *Accruals that involve more estimation are less persistent.*

Dechow and Dichev [2002] finds that firms that exhibit a lower mapping of accruals into past, current, and future cash flows also exhibit lower earnings persistence. They posit that if accruals map less into these cash flows then accrual errors must be greater (i.e. the accruals are recorded with low precision). Thus, accruals will be less predictive of future earnings.

If there is greater estimation in accruals then these earnings are likely to be recorded with lower precision (i.e. accruals map less into realized cash flows). Greater estimation in accruals implies that managers need to make more estimates during the accrual generating process. If managers make a large number of estimations when recording accruals then the range of possible

errors in the recorded accruals is greater. When accruals have greater error, they are less realized as cash in prior, current, or subsequent periods. Following this reasoning, greater estimation during the accrual generating process will be associated with a lower mapping of cash flows into the accruals portion of earnings.

Prediction 2: *Accruals that involve more estimation map less into the firm's past, current, or future cash flows.*

Prior studies have found that the lower persistence of accruals is not quickly incorporated by investors in their valuation of the firm (Sloan [1996], Hanlon [2005], Richardson et al. [2005]). One explanation for this finding is that investors fixate on total earnings thereby disregarding the affect of the lower persistence of accruals on how predictive current earnings are of future earnings (Sloan [1996], Kraft et al. [2006]). Accordingly, Sloan [1996] finds that the future abnormal returns of firms are negatively associated with the magnitude of firms' accruals. This finding is consistent with his hypothesis that investors misinterpret the persistence of accruals. If investors fixate on total earnings and ignore the accruals portion of earnings, which could be calculated from the statement of cash flows or from the balance sheet, then investors may not incorporate the estimation information in the footnotes in a timely manner. Therefore, in the short term, investors should more greatly undervalue firms with more estimated income reducing accruals (negative accruals) and over value firms with more estimated income increasing accruals (positive accruals). Hence, we would find that the estimation information found in the company's footnotes is informative of future returns.

On the other hand, investors may quickly incorporate the amount of estimation in accruals into their valuation of the firm since this information is readily available in the firm's disclosures. More specifically, information provided in a firm's footnote disclosures has been

shown to be incorporated by both investors and analysts (De Franco et al. [2011]). If the amount of estimation involved in the accruals portion of earnings can be found in the notes to the financial statements then investors may become informed of the lower persistence of these earnings upon the filing of the 10-K. If so, then investors may quickly incorporate this estimation information into their valuation of the firm and thus the estimation information found in the footnotes will not be associated with the future long term abnormal returns of the firm.

Prediction 3: *The market reacts as if it does not incorporate the amount of estimation in accruals in their valuation of the firm in a timely manner.*

3. Data Preparation

3.1 Extracting the Footnotes to the Financial Statements

We download all 10-K documents filed with the SEC for fiscal years between 1995 and 2010 from the SEC EDGAR Website.[3] We then extract the notes to the financial statements from each of the 10-K filings using Perl. The extracted footnotes were stripped of all HTML tags and tables. To mitigate any data issues related to extracting the notes to the financial statements from each 10-K filing we truncate our sample of notes to the financial statements by the total number of words at the 1% and 99% level. We also eliminate any filings which are not explicitly identified as either a "10-K" or "10-K405" in the filing's header.

3.2 Measuring Accruals Estimation

We read numerous notes to the financial statements to identify linguistic cues commonly used to denote that some estimation was needed. The first linguistic cue is when an estimation action targets some object. For example, the phrase "we estimated receivables" contains the estimation action "estimated" which targets the object "receivables". This cue denotes that

[3] Companies began filing using EDGAR beginning in 1994-1995.

receivables were estimated. Another linguistic cue is when a "use action" targets an estimate object. An example of this is the phrase "we used estimates" where the action "used" targets the object "estimates". Lastly, the use of an estimate adjective to modify some object also conveys that something was an estimate. An example of this is "estimated costs"; here the object "costs" is being modified by the adjective "estimated" thereby conveying that the costs are estimates.

We automate the search for these linguistic cues by first parsing each of the sentences in our sample using the Stanford open source statistical parser (Marneffe et al. [2006]). The parse of each sentence identifies its noun modifiers, direct object modifiers, adjective modifier, etc. (i.e. its grammatical relationships). Deconstructing sentences in this manner not only provides us with a map of the qualitative information but, more importantly, allows us to utilize the grammatical relationships between the words in the each sentence. Using the grammatical relationships allows us to more correctly identify linguistic cues that convey that some estimation was needed by management (Klein and Manning [2003], see Appendix 1).

Next, we construct fours dictionaries to help us extract meaning from the sentence parses. The first dictionary contains Estimation Actions. Words in the Estimation Actions dictionary convey that an estimation action was performed – this dictionary includes words such as "Estimate," "Anticipate," and "Approximate." The second dictionary is of Estimation Objects (Nouns). This dictionary contains estimation related objects and contains words such as "Belief," "Estimates," and "Approximations." The dictionary of Estimation Objects is used in conjunction with a word from our Use Words dictionary, our third dictionary. This dictionary includes words that denote that management used or needed some object and includes words such as "Make," "Use," and "Include." Our fourth and final dictionary is Estimation Adjectives and contains estimation words that are used to modify some object - these words include "Likely," "Estimated," and "Anticipated".

Finally, using our sentence parses and dictionaries we examine each sentence in our sample of footnotes for the linguistic cues that we identified as conveying that estimation was needed or used by management (see Appendix 2).[4] The number of these linguistic cues is used as our measure of the amount of estimation needed by management during the accrual generating process.

3.2.1 Measuring Between-Accounts Estimations and Within-Accounts Estimations

We also examine whether the number of estimation cues simply captures specific types of accruals accounts. For instance, a firm could have more estimation cues simply because it has defined benefits pension plans or employee stock options, which involve more estimations than other accruals. On the other hand, it is also possible that there are different levels of estimations for the same types of accounts or transactions across companies. That is, the receivables of one company may require more estimation compared with those for another company.

We decompose the number of estimation cues in the footnotes into two components, a component that is explained by the existence of specific accruals accounts ("Between-Accounts Estimations" or BAE) and a component that is due to the within-accounts variations ("Within-Account Estimations" or WAE), and examine whether they have different implications for accruals quality. We measure BAE by calculating the expected amount of estimation needed given the specific items in the company's notes to the financial statements. Specifically, we extract all footnote headers from our sample of 10-K filings and sort them based on their frequency. Starting with the most frequent footnote headers, we categorize approximately one thousand unique footnote headers by hand and find that there are approximately 49 unique

[4] We look for direct objects, nominal subjects, noun compound modifiers, adjectival modifiers, and quantifier phrase modifiers which convey that an estimation was made or used.

footnote items.[5][6] Using this list of unique footnote headers we then find which items appear in each company's notes to the financial statements.

We then regress the number of estimation cues on the footnote item fixed effects:

$$Estimation_{f,t} = \Sigma\beta_j Footnote_{j,f,t} + \epsilon_{f,t} \quad (1)$$

Where $Footnote_{j,f,t}$is an indicator which equals 1 if the company's notes to the financial statements contain the specific footnote item. The predicted (residual) value from this model captures BAE (WAE), or the number of estimation cues explained by the existence of different accruals accounts (the number of estimation cues due to the within-accounts variations). We posit that both BAE and WAE explain the lower persistence of accruals in comparison to cash flows.

3.3 Sample Preparation

We merge the estimation count data with annual financial information from the Wharton Research Data Services (WRDS) Compustat database and equity market information from the Center for Research in Security Prices (CRSP). For a handful of the footnotes in our sample were unable to find corresponding financial data or market information. The main reason for many of these stemmed from not being able to find an appropriate GVKEY for the CIK specified in the header of the 10-K filing. We then eliminate financial institutions from our sample due to the

[5] The list of footnote headers are: Taxes, Accounting Policies, Commitments, Contingencies, Affiliates, Stock, Long-term Debt, Subsequent Events, PP&E, Inventory, Pension and Retirement, Mergers and Acquisitions, Financial Instruments, Earnings Per Share, Segment Information, Leases, Financial Data, Discontinued Operations, Investments, Stock Options, Payables, Cash, Intangibles, Stock Compensation, Business, Cash Flows, Other Assets, Receivables, Credit Arrangements, Regulatory, Derivatives, Going Concern, Credit Risk, Fair Value, Comprehensive Income, Significant Customers, Accounting Changes, Restructuring, Allowance, Parent Company, Restatement, Shareholder Rights, Loan, Dividends, Real Estate, Other Expenses, Joint Ventures, Supplemental Information, and Reinsurance.

[6] The 1,000 hand categorized footnote headers directly account for approximately 70% of all footnote headers from our sample. This hand categorized sample was use to seed our Perl script which searched for footnote headers.

potential idiosyncratic nature of their accruals and disclosures.[7] This leaves us with a sample size of 64,510 firm year observations.

Future long window abnormal returns of the firm are calculated as the compounded returns of the firm minus the compounded returns of the market over the same window. Specifically, we calculated 1 year compounded returns beginning five days after the filing of the 10-K. We also calculated 1 year value weighted compounded market returns beginning five days after the 10-K filing for each of the firms in our sample. Compounded abnormal returns for each firm are the calculated by subtracting the 1 year value weighed compounded market returns from the 1 year compounded returns of the firm.

3.3.1 Estimation Summary

Table 1 presents the average estimation and the average total number of words found in the notes to the financial statements. Consistent with prior studies, we find that the average length of the footnotes has steadily increased over time (Li [2008]). On average, the length of the notes to the financial statements has doubled in size over our sample period going from an average of 5,673 words for the fiscal period 1995 to 13,164 words in 2010. This finding is also consistent with prior studies and anecdotal evidence which suggest that firms' financial disclosures have been increasing in complexity (Radin [2010]). We also find that estimation has increased monotonically during our sample period. Specifically, the number of linguistic cues which convey estimation increased from an average 33 in the fiscal period 1995 to an average of 98 in 2010.

Figure 1 plots the trend of estimation and of the length of the footnotes (as measured by the total number of words in the footnotes) and reiterates our description of the trend in both footnotes length and estimation. Overall, the length of the footnotes has been growing over the

[7] Financial firms are identified as those firms having SIC codes between 6000 and 6999.

years in our sample. Moreover, the estimation count has been growing as well. There appears to be a slight leveling off in the growth of the footnotes in the later periods of our sample.

Table 2 shows the average estimation and the average number of words in the notes to the financial statements by industry. Industry appears to play an important role in the amount of estimation. Agricultural production crops, automotive repair, building materials, home furniture, and leather products are the 5 industries with the least amount of estimation having an average of approximately 40 estimation linguistic cues in their footnotes. On the other hand. the industries of coal mining, electric and gas, petroleum, communications, and oil and gas constitute the 5 industries with the most amount of accruals estimation and have an average of 82 estimation related linguistic cues in the footnotes.

4. Research Design and Results

4.1 Determinants of Estimation

Cross-sectional tests are performed to examine the associations between certain characteristics of the firm and estimation. For this study we examine the following set of determinants of accruals quality indentify in prior studies (Dechow and Dichev [2002], Francis et al. [2005]):

Size

Larger firms typically have more operational complexity than smaller firms. Therefore greater estimation is needed to convey the activities of the firm through accruals.

Negative Earnings

Following the accounting conservatism literature, positive earnings need a greater degree of certainty to be recognized (lower estimation) while negative earnings may be recognized with lower certainty therefore firms with negative earnings are more likely to have more estimation.

Operating Cycle

Longer operating cycles imply greater uncertainty about the turnover of many accruals. Therefore more estimates may be needed by management when estimating future cash receipts.

Standard Deviation of Cash flows

Less stable cash flows suggest that greater estimation is needed to properly measure the cash flows associated with accruals.

Standard Deviation of Sales

A larger standard deviation of past sales suggests that the firm is potentially operating in a volatile market. In this type of environment managers are likely to need more estimates and assumptions when recording accruals.

We perform our cross-sectional test using the following tobit model left censored at 0 to examine the relationship between the determinants and estimation:

$$Estimation_{f,t} = \beta_0 + \beta_1 \text{Size}_{f,t} + \beta_2 OperatingCycle_{f,t} \quad (2)$$

$$+\beta_3 stdev(Sales)_{f,t} + \beta_4 stdev(Cash\ Flows)_{f,t} + \beta_5 NEGEARN_{f,t} + \epsilon_{f,t}$$

Where $Estimation_{f,t}$ is defined earlier. $Size_{f,t}$ is the log of the market value of the firm's equity. $Operating\_Cycle_{f,t}$ is the operating cycle of the firm calculated as $\log(\frac{invt}{cogs} * 360 + \frac{rect}{sales} * 360)$. $stdev(Sales)_{f,t}$ is the standard deviation of the firms sales over the past five years. $stdev(Cash\ Flows)_{f,t}$ is the standard deviation of the firms cash flows over the past five years. $NEGEARN_{f,t}$ is the number of years that the firm had negative earnings over the past 5 years.

4.1.1 Determinants of Estimation Findings

Table 5 presents the results for the determinants of estimation. Several of the characteristics including the size of the firm, the standard deviation of sales, operating cycle, and

the number of the past five years in which the firm has negative earnings are consistent with our hypotheses. One of the determinants of accruals quality, standard deviation of cash flows, is statistically significant but the coefficient loads is in the opposite direction of our prediction. One explanation for why the standard deviation of cash flows is negatively associated with estimation is that managers simply have not booked accruals since they are very uncertain of future cash flows due to their high volatility.

4.2 Estimation and the Persistence of Accruals

To test our first prediction (P1), we follow prior literature and measure the persistence of earnings by regressing the following year's earnings on the current year's earnings (Sloan [1996], Li [2008]). Fundamentally, this regression measures how predictive current earnings are of future earnings. If the estimated coefficient on current earnings is high then we would conclude that current earnings are highly persistent since they are highly associated with future earnings and vice versa when the estimated coefficient on earnings in the regression is low.

First, we examine the marginal effect of estimation on the persistence of total earnings. We include the interaction between the current year's earnings and the amount of estimation to measure the impact of estimation on the persistence of earnings. For a given level of earnings, how much more (or less) persistent are they for a given level of estimation. If our hypothesis is correct then we should find a negative coefficient on this interaction term.

$$Earnings_{f,t+1} = \beta_0 + \beta_1 Earnings_{f,t} + \beta_2 Estimation_{f,t} \quad (3)$$
$$+\beta_3 Earnings_f * Estimation_{f,t} + \Sigma\beta_i Controls_{f,t} + \Sigma\beta_j Earnings_{f,t} * Controls_{f,t}$$
$$+ AuditorFE_f + YearFE_t + \text{IndustryFE}_\text{f} + \epsilon_{f,t}$$

Next, we disaggregate earnings into a cash flows component and an accruals component and interact each component with our measure of estimation. We follow the recommendations made in Hribar and Collins [2002] and calculate accruals using the statement of cash flows. If greater estimation lowers the association between the current year's accruals the following year's earnings then the interaction between estimation and the accruals portion of earnings will be negative. Ideally, if the number of linguistic cues which convey estimation in the footnotes does not capture the precision of the cash portion of earnings the interaction between cash flows and estimation should be statistically insignificant.[8]

$$Earnings_{f,t+1} = \beta_0 + \beta_1 Cash_{f,t} + \beta_2 Accruals_{f,t} + \beta_3 Estimation_{f,t} \quad (4)$$
$$+ \beta_4 Cash_{f,t} * Estimation_{f,t} + \beta_5 Accruals_{f,t} * Estimation_{f,t} + \Sigma\beta_i Controls_{f,t}$$
$$+ \Sigma\beta_i Accruals_{f,t} * Controls_{f,t} + \Sigma\beta_i Cash_{f,t} * Controls_{f,t}$$
$$+ AuditorFE_f + YearFE_t + \text{IndustryFE}_\text{f} + \epsilon_{f,t}$$

where $Estimation_{f,t}$ is defined earlier; $Earnings$ is income before extraordinary items; $Cash_{f,t}$ is the portion of total earnings due to operating cash flows; A$ccruals_{f,t}$is the portion of total earnings due to accruals. The controls include size, operating cycle, standard deviation of sales, standard deviation of operating cash flows, the number of years over the past 5 years in which the firm had negative earnings, and the total length of the footnotes.[9] We also include

[8] The measure of estimation may also capture overall uncertainty about the firm. If so then the coefficient on the interaction between cash flows and estimation will also be negative statistically significant.
[9] We include the total length of the footnotes since longer footnotes are more likely to have more estimation related linguistic cues.

interactions between all control variables and earnings and auditor, year, and industry fixed effects.[10] All continuous variables are scaled by average total assets.

4.2.1 Estimation and the Persistence of Accruals Findings

Table 6 presents the results for the regression of next year's earnings on current year's earnings, estimation, the interaction between estimation and the current year's earnings, and our controls (4). The coefficient on the interaction between estimation and earnings in the current year is negative and statistically significant at 1% which suggests that earnings which needed more estimation are less persistent than those which required less estimation. The economic significance of the effect of estimation on the persistence of earnings when going from the 25th percentile of estimation to the 75th is approximately -0.0663 (-0.0013 * (84 – 33)); this translates into a percentage difference of approximately 10% to 15% when compared to the baseline persistence of accruals (Dechow et al. [2006]). Overall, these findings suggest that the amount of estimation needed during the accrual generating process is associated with an economically significant decrease in the persistence of total earnings.

The 3rd Column of Table 6 shows the interactions between BAE and WAE and earnings. Both interaction terms are statistically significant at the 1% level. The magnitude of the difference in the persistence of earnings when going from the 25th percentile of BAE to the 75th percentile of BAE is -0.0585 (-0.0013 * (84 – 39)). The economic magnitude when going from the 25th to the 75th percentile of WAE is -0.0406 (-0.0014 * (13 + 16)).

As discussed before, accruals are one component of total earnings and the measure of the amount of estimation should only pertain to the accruals portion of earnings and not the cash flows portion. Table 7 shows the results when we interact each component of earnings separately

[10] We include the interaction between the control variables and earnings since we want to control for the marginal impact of the control variable on the persistence of earnings in addition to the control variables impact on future performance.

with our measure of estimation.[11] As predicted the interaction between accruals and estimation is negative and statistically significant, which is consistent with our prediction that accruals which needed more estimation exhibit lower persistence. The difference in the persistence of accruals between the 25th percentile of estimation and the 75th percentile is approximately -0.0561 (-0.0011 * (84 - 33)). Next, the results show that the coefficient on the interaction between cash flows and the amount of estimation is statistically insignificant. This finding suggests that our measure of estimation captures some characteristic of accruals but not cash flows.

The 5th Column of Table 7 shows the interaction between the amount of BAE and WAE and both accruals and cash flows. Our findings suggest that the lower persistence of accruals in comparison to cash flows is driven by both components of estimation. The interaction between the amount of BAE and earnings and the interaction between WAE and earnings are both statistically significant. The difference in the persistence of earnings when going from the 25th percentile to the 75th percentile of BAE is approximately -0.054 (-0.0012 * (84 – 39)) while the difference in the persistence of accruals is -0.0232 (-0.0008 * (13 + 16)). An F-test of the coefficients on the interaction between BAE and accruals and WAE and accruals yields a p-value of 0.3949 thereby suggesting that the coefficients are not statistically different.

Table 8 presents the accruals and cash flows persistence test (4) including three characteristics of accruals that have been shown in prior studies to be associated with a lower persistence of accruals. Specifically, we include the absolute value of the magnitude of accruals, the standard deviation of the Dechow and Dichev residual, and special items into our tests of accruals persistence. Table 8 Column 4 of Panel A shows that all three measures are statistically significant at the 1% level and negatively associated with the persistence of accruals. Moreover, the standard deviation of the Dechow and Dichev residual and Special Items are associated with

[11] In untabluated results, consistent with prior research we find that the persistence of accruals is less than that of cash earning and that the magnitudes of the coefficients are similar to those found in prior research.

a lower persistence of cash flows which suggests that these measures pickup uncertainty about the cash flows of the firm as well. Even after including these alternate measures of estimation we still see that our measure of estimation is still associated with a lower persistence of accruals at the 1% level of significance. However, the economic significant of our measure has decreased to -0.0306, a reduction of approximately 45% from the economic significance found in Table 7. Even so, these findings suggest that our measure is informative about some aspect of accruals persistence not found in these other measures.

Panel B shows the results when we disaggregate estimation into BAE and WAE. BAE remains statistically significant at the 10% level and is still negatively associated with the persistence of accruals. On the other hand, the within-accounts portion of estimation is no longer statistically significant but the sign of the coefficient is still negative.

4.3 Estimation and the Mapping of Accruals into Cash Flow

Prediction 2 (P2) suggests that when greater estimation is needed during the accrual generating process accruals are less likely to be realized as cash. We use the measure of how well accruals map into cash flows developed by Dechow and Dichev [2002] (hereafter DD) to capture this effect. This model captures accruals quality by estimating how well working capital accruals map to into realized operating cash flows. The model is based on the premise that accruals are a way to shift the recognition of cash flows.[12] If the realized cash flows of the firm map well into the accruals of the firm then the firm's accruals are deemed to be of high quality. DD operationalize their theory by regressing current period working capital accruals on prior period, current period, and next periods operating cash flows. The standard deviation of the residual from this model is the measure of how well the firm's accruals map into cash flows.

[12] This model does not distinguish between managed earnings or those which arise due to unintentional errors or management uncertainty.

The specification of the DD model is shown in equation (1). We include the change in revenues and Property, Plant and Equipment (PPE) in the model as proposed in McNichols [2002]. We estimate the model by industry and year and use the standard deviation of the residual from the model over the past 5 years for each firm as our measure of how well the accruals of the firm map into cash flows of the firm.

$$TCACC_{f,t} = \beta_0 + \beta_1 CFO_{f,t-1} + \beta_2 CFO_{f,t} + \beta_3 CFO_{f,t+1} \quad (5)$$
$$+\beta_4 \Delta Rev_{f,t} + \beta_5 PPE_{f,t} + \epsilon_{f,t}$$

Where CFO are the operating cash flows of the firm. $TCACC_{f,t}$ is the total working capital accruals of firm. $\Delta Rev_{f,t}$ is the change in sales from the prior year. $PPE_{f,t}$ is the total property plant and equipment for the current fiscal period. All continuous variables are scaled by average total assets.

4.3.1 Estimation and the Mapping of Accruals into Cash Flow Findings

Table 9 presents the results for how estimation affects how well accruals map into cash flows (P2). Column 2 of Table 9 shows the regression of the determinants of accruals quality as identified in Francis et al. [2005]. All of the determinants of accruals quality are statistically significant and load in the same direction as found in prior studies (Francis et al. [2005], Dechow and Dichev [2002]).

The 3rd Column of Table 9 includes our measure of estimation into the model and we find that estimation is statistically significant at the 1% level and is positively associated with the standard deviation of the Dechow and Dichev residual. This finding is consistent with our hypothesis that the amount of estimation in accruals is associated with greater accrual errors and therefore associated with a lower mapping of cash flows into accruals.

The 4th Column of Table 9 presents the results when we decompose estimation into BAE and WAE. Once again we see that both components of estimation are statistically significant at the 1% level and positively associated with a lower mapping of accruals into cash flows. An F-test of the equality of the coefficients on BAE and WAE yields a p-value of 0.6768 therefore we can't reject the null. This result suggests that both components of estimation affect the mapping of accruals into cash flows similarly. Overall our findings suggest that when there is a greater amount of estimation during the accrual generating process there is a lower mapping of accruals into cash flows.

4.4 Estimation and Future Abnormal Returns

For our test of P3 we follow the research design of Sloan [1996] and Richardson et al. [2005] to determine whether the market reacts as if it quickly incorporates the estimation information found in the footnotes. Sloan [1996] regresses future abnormal returns on total accruals and finds a negative association between the two. Therefore, positive accruals are associated with negative future abnormal returns. On the other hand, negative accruals are associated with positive future abnormal returns. These findings are consistent with his hypothesis that the market over value the persistence of accruals.

We make several small but important modifications to their research design for our study. Since we are interested in the incremental effect of the amount of estimation on the persistence of accruals we include the interaction between the amount of estimation and total accruals into the model. The interaction term models the marginal effect of the amount of estimation on the association between current accruals and future abnormal returns. If the interaction effect is negative then this suggests that the market overvalues more highly estimated positive accruals and vice versa for negative accruals.

Next, we make two small changes to the specification of the model to better coincide with our research design. First, Sloan [1996] calculates future abnormal returns beginning four

months after the end of the firm's fiscal period. In contrast, our abnormal returns accumulation begins 5 days after firm files their 10-K form with the SEC. The information about the estimation of accruals used in this study is found in the firm's 10-K filing. Therefore, we need to ensure that the estimation information found in the footnotes to the financial statements is available to the market before we can assess whether the market incorporated the information in a timely manner. Of course, some of the estimation information may have been released prior to the filing of the 10-K but this would only bias results away from our prediction since the market would have had more time to incorporate the information. Second, rather than using a decile ranking of accruals we use the raw amount of accruals. One of the purposes of Sloan [1996] was to show that a trading strategy could be implemented by purchasing stock in firms with extreme low accruals (this in the lowest decile of accruals) and shorting those with extreme high accruals (this in the highest decile of accruals). The purpose of this study isn't to implement a trading strategy but rather the provide evidence that the markets appear to not quickly incorporate the estimation information found in the footnotes. Therefore, to preserve more of the information in accruals, we use the raw accruals amount rather than the decile ranking of the amount of accruals.

$$Abnreturns_{f,t} = \beta_0 + \beta_1 Accruals_{f,t} + \beta_2 Estimation_{f,t} \quad (6)$$
$$+\beta_3 Estimation_{f,t} * Accruals_{f,t} + \beta_4 Size_{f,t} + \beta_5 BTM_{f,t}$$
$$+\beta_6 ETP_{f,t} + \beta_7 Beta_{f,t} + \epsilon_{f,t}$$

Where $Estimation_{f,t}$ is defined earlier. A$ccruals_{f,t}$ is income before extraordinary items minus operating cash flows. S$ize_{f,t}$ is the log of the market value of the firms equity. $BTM_{f,t}$ is the book to market ratio of the firm. $ETP_{f,t}$ is the firms earnings to price ratio. B$eta_{f,t}$if the market

beta of the firm for fiscal period t. $Abnreturns_{f,t}$ is 1 year abnormal returns beginning 5 days after the filings of the 10-K.

We also use the Fama-French Carhart four-factor model to further test the association between accruals estimation and future abnormal returns. More specifically, we construct 25 portfolios each month based on the amount of accruals of the firm and the amount of estimation conveyed in the notes to financial statements – 5 rankings of accrual x 5 rankings of estimation. For each portfolio we then estimate the four-factor alpha using the following model.

$$MonthyExr_t = \beta_0 + \beta_1 MktExr_t + \beta_2 HML_t + \beta_3 SMB_t + \beta_4 UMD_t + \epsilon_t \quad (7)$$

Where $MonthyExr_t$ is the monthly excess return of the value (equal) weighted portfolios. $MktExr_t$ is the monthly return of the value-weighted index minus the risk free rate. HML_t = monthly premium of the book-to-market factor. SMB_t is the monthly premium of the size factor; UMD_t is the monthly premium on winners minus losers.[13]

4.4.1 Estimation and Returns Findings

The second columns of Table 10 present the baseline results shown in Sloan [1996]. For our sample we find that accruals are negatively associated with the future abnormal returns of the firm. Column 3 of Table 10 shows the results of our test of P3 - the association between future long term abnormal returns and the amount of estimation information found in the footnotes to the financial statements. The coefficient on the interaction between estimation and total accruals is negative and statistically significant at the 1% level. Therefore, accruals which required more estimation are more negatively associated with future long term abnormal returns. This is consistent with the hypothesis that investors overvalue firms with less persistent positive accruals

[13] RF, HML, SMB, and UMD factors are from Ken French's website.

and undervalue those with less persistent negative accruals. Column 4 presents the results of the model when we decompose estimation into BAE and WAE. We find that the negative association between abnormal returns and the interaction of accruals and estimation is primarily driven by BAE.

Table 11 presents the results for our test of P3 using the four-factor model. Overall we find mixed results as to whether the accrual anomaly is most concentrated in those firms with the greatest amount of accruals estimation. Panel A of Table 11 presents the results for the top and bottom most portfolios in terms of estimation and accruals. We construct hedged returns by going long in those firms with the greatest accruals and short in those with the least amount of accruals. When using equal weighted hedged returns we find evidence that accruals are associated with future abnormal returns in our sample. However, we do not find any association between the future abnormal returns of the firm and accruals when using value weighted hedge returns.

We then examine whether the association between future abnormal returns and the current years accruals are greater for those firms with the greatest amount of accruals estimation. We do not find any evidence consistent with the hypothesis that the accrual anomaly is exacerbated in those firms with greater estimation. This result is consistent with the hypothesis that investors utilize the estimation information found in the notes to the financial statements when valuing the firm.

Lastly, Panel B disaggregates estimation into BAE and WAE. We find no discernible association between future abnormal returns and the amount of estimation and accruals when using value weighted returns. Using equal weighted returns we find little difference between the high and low estimation portfolios when using BAE. However, we do find some evidence that those firms with the least WAE are associated with the future abnormal returns of the firm.

5. Conclusion

The primary focus of this study was to examine the association between the estimation needed during the accrual generating process and the persistence of accruals. Sloan [1996] and Richardson et al. [2005] suggest that the estimation needed when recording accruals reduces the persistence of accruals. Their hypotheses are based on the idea that if accruals require a greater degree of estimation they are more likely to be recorded with error (i.e. accruals are less precise). If so, then accruals will be less predictive of future earnings. While this conjecture has been generally accepted in the accounting literature few have explicitly examined the association between the estimation involved during the accrual generating process and the persistence of accruals.

This study provides evidence consistent the conjectures that the estimation needed during the accrual generating process plays a key role in the persistence of accruals. Specifically, we find that when accruals have more estimation they are less predictive of future earnings. We also find that accruals map less into the past, current, or future cash flows of the firm when they require more estimation. Next, we find that both the between-accounts and within-accounts portions of estimations drive our results. Lastly, we find mixed evidence as to whether the markets do not quickly incorporate the estimation information found in firms' footnotes into their valuation of the firm.

In conclusion, the findings in this study provide insight into the accrual generating processing of the managers. More importantly, the findings in our study suggest that understanding the process which managers undergo when recording accruals plays an important role in understanding the persistence and quality of accruals.

References

Dechow, P., I. Dichev, 2002, The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors, *The Accounting Review* 77, 35-59.

Dechow, P., W. Ge, 2006, The Persistence of Earnings and Cash Flow and the Roles of Special Items: Implications for the Accrual Anomaly, *Review of Accounting Studies,* 253-296.

Dechow, P., W. Ge, C. Schrand, 2010, Understanding earnings quality: A review of the proxies, their determinants and their consequences. Journal of Accounting and Economics, 50(2-3), 344–401.

Dechow, P. R. Sloan, A. Sweeney, 1995, Detecting Earnings Management, *The Accounting Review* 70, 193-225.

De Franco, G., F. Wong, Y. Zhou, 2011, Accounting Adjustments and the Valuation of Financial Statement Note Information in 10-K Filings, *The Accounting Review,* Forthcoming.

Francis, J., R. LaFond, P. Olsson, K. Schipper, 2005, The Market Pricing of Accruals Quality, *Journal of Accounting and Economics* 39, 295-327.

Hanlon, M., 2005, The Persistence and Pricing of Earnings, Accruals, and Cash Flows When Firms Have Large Book-Tax Differences, *The Accounting Review* 80, 137-166.

Hribar, P., D. Collins, 2002, Errors in Estimating Accruals: Implications for Empirical Research, *Journal of Accounting Research* 40.

Jones, J., 1991, Earnings Management During Import Relief Investigations. *Journal of Accounting Research* 29, 193–228.

Klein, D. and C. Manning, 2003, "Accurate Unlexicalized Parsing" *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, 423-430.

Kraft, A., A. Leone, C. Wasley, 2006, An Analysis of the Theories and Explanations Offered for the Mispricing of Accruals and Accrual Components. Journal of Accounting Research, 44(2), 297–339.

Li, F., 2008, Annual Report Readability, Current Earnings, and Earnings Persistence, *Journal of Accounting and Economics* 45, 221-247.

Li, F., 2011, Textual Analysis of Corporate Disclosures: A Survey of the Literature, *Journal of Accounting Literature,* 29, 143-165.

Marneffe, M., B. MacCartney, C. Manning, 2006, Generating Typed Dependency Parses from Phrase Structure Parses. *LREC 2006*.

Merkley, K., 2010, More Than Numbers: R&D-related Disclosure and Firm Performance, Working Paper, Cornell University.

Mashruwala, C., S. Rajgopal, & T. Shevlin, 2006, Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. Journal of Accounting and Economics, 42(1-2), 3–33.

McNichols, M., 2002, Discussion of The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors, *The Accounting Review* 77, 61-69.

Radin, A., 2010, Have We Created Financial Statement Disclosure Overload?, The CPA Journal.

Richardson, S., R. Sloan, M. Soliman, I. Tuna, 2005, Accrual Reliability, Earnings Persistence and Stock Prices, *Journal of Accounting and Economics* 39, 437-485.

Riedl, E., S. Srinivasan, 2010, Signaling Firm Performance Through Financial Statement Presentation: An Analysis Using Special Items, *Contemporary Accounting Research* 27, 289-332.

Sloan, R. 1996, Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings?, *The Accounting Review* 71, 289-315.

United States Securities and Exchange Commission, SEC Form 10-K General Instructions, http://www.sec.gov/about/forms/form10-k.pdf.

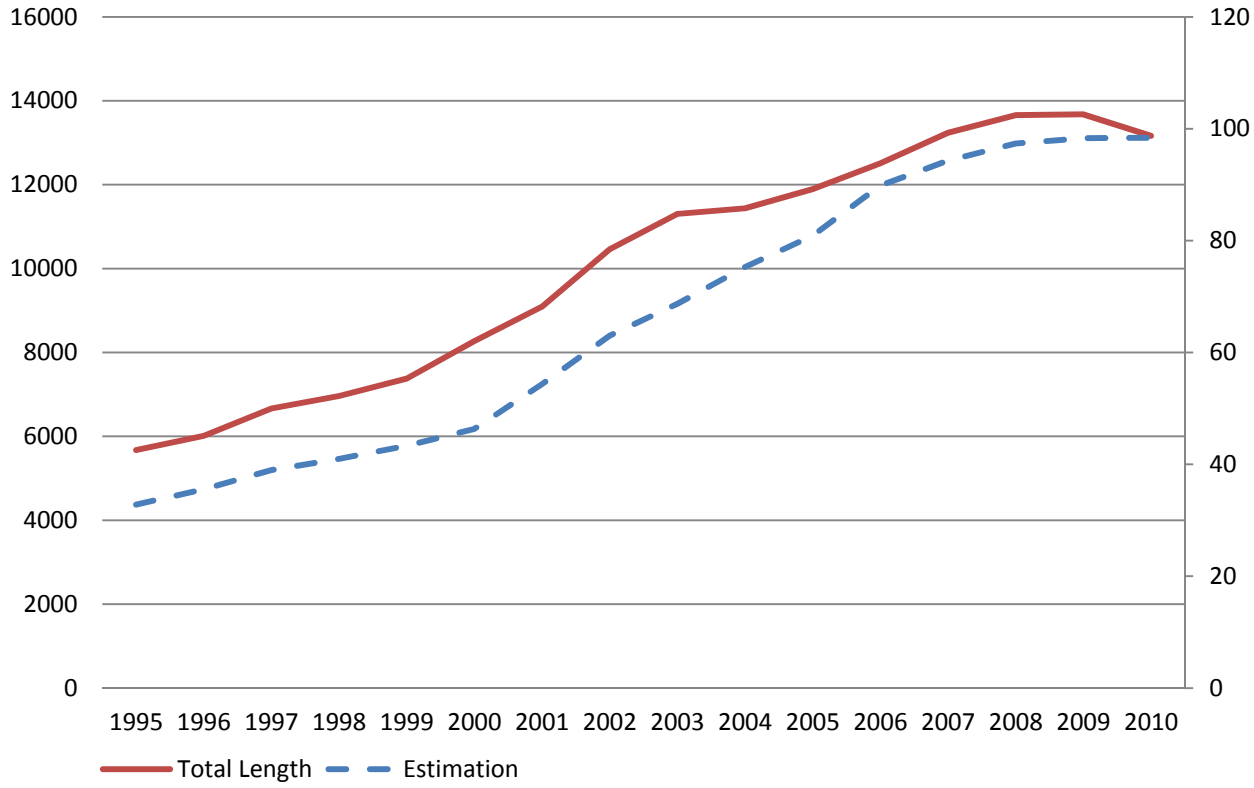
United States Securities and Exchange Commission, SEC Act of 1934, http://www.sec.gov/about/laws/sea34.pdf.

White, H., 2010, An Accrual Balance Approach to Assessing Accrual Quality. *Working Paper.*

Xie, H., 2001, The Mispricing of Abnormal Accruals, (3), 357–373.

Figure 1

Length of the Footnotes and Estimation Trend



Notes: Figure 1 shows the trend in the in the length of the notes to the financial statements (blue) and the estimation count found in the footnotes (red).

Table 1

Estimation Trend

Year	Estimation	Length	N
1995	33	5,673	2,783
1996	36	6,015	4,818
1997	39	6,667	4,968
1998	41	6,963	4,830
1999	43	7,380	4,948
2000	46	8,270	4,789
2001	54	9,092	4,491
2002	63	10,461	4,257
2003	69	11,302	4,031
2004	75	11,436	3,978
2005	81	11,897	3,683
2006	90	12,508	3,680
2007	94	13,236	3,703
2008	97	13,657	3,950
2009	98	13,678	3,788
2010	98	13,164	3,564
Average	66	10,087	4,141

Notes: Table 1 presents the average number of estimation related linguistic cues and the total number of words in the footnotes found in the notes to the financial statements.

Table 2

Estimation by Industry

Industry	Two Digit SIC	Estimation	Total Length	N
Agricultural Production Crops	1	31	5651	16
Metal Mining	10	72	10561	378
Coal Mining	12	98	14020	54
Oil And Gas Extraction	13	75	10420	2446
Mining And Quarrying Of Nonmetallic Minerals, Except Fuels	14	45	6332	77
Building Construction General Contractors And Operative Builders	15	57	8767	422
Heavy Construction Other Than Building Construction Contractors	16	67	9002	155
Construction Special Trade Contractors	17	41	5930	99
Food And Kindred Products	20	56	8432	1508
Textile Mill Products	22	48	6986	308
Apparel And Other Finished Products Made From Fabrics And Similar Materials	23	54	8449	684
Lumber And Wood Products, Except Furniture	24	54	7311	316
Furniture And Fixtures	25	52	7138	452
Paper And Allied Products	26	74	9334	661
Printing, Publishing, And Allied Industries	27	59	8141	897
Chemicals And Allied Products	28	65	10274	6264
Petroleum Refining And Related Industries	29	77	10577	408
Rubber And Miscellaneous Plastics Products	30	57	8197	835
Leather And Leather Products	31	40	7084	203
Stone, Clay, Glass, And Concrete Products	32	58	7838	402
Primary Metal Industries	33	63	8863	1027
Fabricated Metal Products, Except Machinery And Transportation Equipment	34	56	7289	1071
Industrial And Commercial Machinery And Computer Equipment	35	62	8246	4314
Electronic And Other Electrical Equipment And Components	36	66	8798	5733
Transportation Equipment	37	69	8711	1455
Measuring, Analyzing, And Controlling Instruments	38	58	8232	4357
Miscellaneous Manufacturing Industries	39	53	7552	736

Table 2

Estimation by Industry (continued)

Industry	Two Digit SIC	Estimation	Total Length	N
Railroad Transportation	40	51	7848	49
Motor Freight Transportation And Warehousing	42	49	6770	540
Water Transportation	44	66	9986	266
Transportation By Air	45	71	9598	479
Transportation Services	47	61	9619	276
Communications	48	77	11765	2471
Electric, Gas, And Sanitary Services	49	92	13172	3303
Wholesale Trade-durable Goods	50	50	7791	1851
Wholesale Trade-non-durable Goods	51	57	9234	1039
Building Materials, Hardware, Garden Supply, And Mobile Home Dealers	52	35	5905	36
General Merchandise Stores	53	55	7337	427
Food Stores	54	54	7531	438
Automotive Dealers And Gasoline Service Stations	55	66	10125	343
Apparel And Accessory Stores	56	51	7749	753
Home Furniture, Furnishings, And Equipment Stores	57	39	5997	288
Eating And Drinking Places	58	55	8062	1236
Miscellaneous Retail	59	52	8608	1499
Hotels, Rooming Houses, Camps, And Other Lodging Places	70	54	7589	265
Personal Services	72	48	7788	143
Business Services	73	64	9398	8689
Automotive Repair, Services, And Parking	75	34	5020	21
Motion Pictures	78	65	10251	402
Amusement And Recreation Services	79	62	10081	858
Health Services	80	67	9972	1376
Educational Services	82	64	9165	292
Social Services	83	41	6917	65
Engineering, Accounting, Research, Management, And Related Services	87	63	9798	1529
Nonclassifiable Establishments	99	55	9020	298

Notes: Table 2 shows average number of estimation related linguistic cues and the average number of words in the notes to the financial statements by industry for fiscal periods between 1995-2010.

Table 3

Summary Statistics

Variable	N	Mean	Minimum	P1	P25	Median	P75	P99	Maximum	Std. Dev.
Total Earnings	64510	-0.067	-1.603	-1.603	-0.082	0.024	0.071	0.320	0.320	0.294
Accruals	64510	-0.085	-0.952	-0.952	-0.118	-0.058	-0.014	0.298	0.298	0.166
Operating Cash Flows	64510	0.018	-1.016	-1.016	-0.015	0.067	0.128	0.376	0.376	0.217
Estimation	64510	64	10	10	33	52	84	217	217	42
BAE	61408	64	13	13	39	55	84	153	153	32
WAE	61408	1	-59	-59	-16	-3	13	108	108	29
Length	64510	9247	1743	1743	4579	7164	11308	44229	44229	7300
Operating Cycle	62585	4.591	1.827	1.827	4.183	4.660	5.094	6.654	6.654	0.817
log(Market Value)	60204	5.298	0.031	0.031	3.794	5.297	6.779	10.599	10.599	2.182
NEGEARN	53625	1.719	0	0	0	1	3	5	5	1.820
BTM	60043	0.694	0.069	0.069	0.424	0.670	0.914	1.919	1.919	0.363
Stdev(DD Residual)	45412	0.088	0.008	0.008	0.032	0.057	0.105	0.581	0.581	0.095
Beta	52869	0.840	-5.081	-0.386	0.374	0.777	1.227	2.633	5.809	0.650

Notes: Table 3 shows the summary statistics for the sample used in this study. *Total Earnings* is the firms income before extraordinary items scaled by average total assets. *Accruals* are total accruals scaled by average total assets. *Operating Cash Flows* are operating cash flows scaled by average total assets. *Estimate* is the number of estimation related linguistic cues found in the footnotes section of the firm's 10-K. *BAE* and *WAE* are the fitted value and residuals from the regression of *Estimate* on the footnote item fixed effects in equation (1) and footnote 5. *Length* is measured as the total number of words in the footnotes section of the firm's 10-K. *Length* is the total number of words in the notes to the financial statements. *Operating Cycle* is the log of a operating cycle of the firm. *Log(Market Value)* the market value of the firm's equity is calculated as the share price of the firm's stock at the filing date multiplied by the number of shared outstanding. *NEGEARN* is the number of years over the past 5 years in which the company had negative earnings. *BTM* is the book to market ratio. This ratio is calculated as the book value of assets divided by the market value of equity plus liabilities. *ETP* is the earnings to price ratio calculated as the firms operating income after depreciation divided by price. *Beta* is the firm annual beta.

Table 4

Correlations

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Total Earnings	-	0.367	0.664	-0.113	-0.020	-0.136	-0.152	-0.013	0.383	-0.642	-0.298	-0.278	0.006
2	Accruals	0.574	-	-0.308	-0.099	-0.065	-0.059	-0.105	0.231	0.065	-0.211	-0.013	-0.142	-0.030
3	Operating Cash Earnings	0.782	-0.042	-	-0.021	0.047	-0.089	-0.064	-0.200	0.351	-0.503	-0.208	-0.229	0.012
4	Estimation	-0.032	-0.071	0.017	-	0.760	0.481	0.863	-0.099	0.396	0.096	-0.001	0.059	0.357
5	BAE	0.031	-0.040	0.070	0.730	-	-0.132	0.705	-0.098	0.367	0.009	-0.015	0.034	0.352
6	WAE	-0.080	-0.059	-0.055	0.637	-0.055	-	0.355	-0.015	0.127	0.122	0.020	0.038	0.076
7	Length	-0.083	-0.067	-0.052	0.697	0.508	0.450	-	-0.113	0.353	0.145	-0.010	0.094	0.331
8	Operating Cycle	-0.025	0.140	-0.135	-0.063	-0.064	-0.020	-0.061	-	-0.078	0.013	-0.022	0.072	0.012
9	log(Market Value)	0.310	0.109	0.304	0.406	0.370	0.178	0.265	-0.043	-	-0.407	-0.390	-0.320	0.398
10	Negative Earnings	-0.597	-0.243	-0.556	0.059	-0.012	0.099	0.110	-0.001	-0.383	-	0.017	0.519	0.066
11	BTM	0.011	-0.002	0.012	0.000	-0.016	0.017	-0.003	-0.004	-0.419	-0.028	-	-0.138	-0.205
12	Stdev(DD Residual)	-0.433	-0.243	-0.352	0.030	0.007	0.036	0.072	0.017	-0.251	0.462	-0.140	-	0.071
13	Beta	-0.024	-0.024	-0.010	0.320	0.328	0.098	0.213	0.016	0.347	0.079	-0.188	0.066	-

Notes: Table 4 presents the Spearman (above diagonal) and Pearson (below diagonal) correlation for the main variables used in this study. *Accruals* are total accruals scaled by average total assets. *Operating Cash Flows* are operating cash flows scaled by average total assets. *Estimate* is the number of estimation related linguistic cues found in the footnotes section of the firm's 10-K. *BAE* and *WAE* are the fitted value and residuals from the regression of *Estimate* on the footnote item fixed effects in equation (1) and footnote 5 *Length* is measured as the total number of words in the footnotes section of the firm's 10-K. *Length* is the total number of words in the notes to the financial statements. *Operating Cycle* is the log of a operating cycle of the firm. *Log(Market Value)* the market value of the firm's equity is calculated as the share price of the firm's stock at the filing date multiplied by the number of shared outstanding. *NEGEARN* is the number of years over the past 5 years in which the company had negative earnings. *BTM* is the book to market ratio. This ratio is calculated as the book value of assets divided by the market value of equity plus liabilities. *ETP* is the earnings to price ratio calculated as the firms operating income after depreciation divided by price. *Beta* is the firm annual beta.

Table 5

Determinants of Estimation

Dependent Variable	**Estimation Count**
	(p-value)
log(Market Value)	6.8628***
	(0.000)
Operating Cycle	0.7047*
	(0.092)
Stdev(Sales)	5.7875***
	(0.000)
Stdev(Operating Cash Flows)	-43.1774***
	(0.000)
Negative Earnings	5.3469***
	(0.000)
Constant	-17.5249***
	(0.000)
Observations	49,822
Pseudo R-squared	0.0607

Notes: Table 5 shows the regression of determines of estimation on the estimation. P-values are reported in parenthesis below their respective coefficients. Please refer to section 4 of the study for detailed descriptions of each of the variables. All continuous variables are winsorized at 1% and 99% of their respective sample distributions. ***, **, and * denote two-tailed statistical significance at 1%, 5%, and 10% significance levels respectively.

Table 6

Earnings Persistence

Dependent Variable	**Earnings t+1** *(p-value)*	**Earnings t+1** *(p-value)*
Earnings	0.3071***	0.3119**
	(0.010)	(0.013)
Estimation	-0.0001*	
	(0.062)	
Estimation x Earnings	-0.0013***	
	(0.000)	
BAE		-0.0001
		(0.262)
BAE x Earnings		-0.0013***
		(0.002)
WAE		-0.0001***
		(0.007)
WAE x Earnings		-0.0014***
		(0.000)
Constant	Yes	Yes
Controls	Yes	Yes
Controls Interact w/ Earnings	Yes	Yes
Fixed Effects	Industry, Year, Auditor	Industry, Year, Auditor
Cluster	Industry, Year	Industry, Year
Observations	45,718	43,155
Adjusted R-squared	0.538	0.538

Notes: Table 6 presents the results for our tests of the amount of estimation on earnings persistence. The controls include size, operating cycle, standard deviation of sales, standard deviation of operating cash flows, the number of years over the past 5 years in which the firm had negative earnings, and the total length of the footnotes. Please refer to section 4 of the study for detailed descriptions of each of the variables. P-values are reported in parenthesis below their respective coefficients. All continuous variables are winsorized at 1% and 99% of their respective sample distributions. ***, **, and * denote two-tailed statistical significance at 1%, 5%, and 10% significance levels respectively.

Table 7

Accruals and Operating Cash Flows Persistence

Dependent Variable	**Earnings t+1** *(p-value)*	**Earnings t+1** *(p-value)*	**Earnings t+1** *(p-value)*	**Earnings t+1** *(p-value)*
Accruals	0.5340***	0.5457***	0.2420***	0.2563***
	(0.000)	(0.000)	(0.006)	(0.004)
Operating Cash Flows	0.8828***	0.8904***	0.5959***	0.5969***
	(0.000)	(0.000)	(0.000)	(0.000)
Estimation	-0.0001		-0.0001**	
	(0.169)		(0.010)	
Estimation x Accruals	-0.0012***		-0.0011***	
	(0.003)		(0.000)	
Estimation x Operating Cash Flows	0.0002		-0.0001	
	(0.305)		(0.771)	
BAE		0.0001		-0.0002**
		(0.355)		(0.036)
BAE x Accruals		-0.0015**		-0.0012***
		(0.012)		(0.000)
BAE x Operating Cash Flows		0.0001		-0.0001
		(0.832)		(0.857)
WAE		-0.0001**		-0.0001**
		(0.013)		(0.029)
WAE x Accruals		-0.0011***		-0.0008**
		(0.003)		(0.014)
WAE x Operating Cash Flows		0.0003		-0.0002
		(0.269)		(0.562)
Constant	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Controls Interact w/ Cash earnings	-	-	Yes	Yes
Controls Interact w/ Accrual earnings	-	-	Yes	Yes
Fixed Effects	Industry, Year, Auditor	Industry, Year, Auditor	Industry, Year, Auditor	Industry, Year, Auditor
Cluster	Industry, Year	Industry, Year	Industry, Year	Industry, Year
Observations	58,692	55,797	45,718	43,155
Adjusted R-squared	0.603	0.602	0.569	0.568

Notes: Table 7 presents the results for our tests of estimation on the persistence of cash and accruals. The controls include size, operating cycle, standard deviation of sales, standard deviation of operating cash flows, the number of years over the past 5 years in which the firm had negative earnings, and the total length of the footnotes. Please, refer to section 4 of the study for detailed descriptions of each of the variables. P-values are reported in parenthesis below their respective coefficients. All continuous variables are winsorized at 1% and 99% of their respective sample distributions. ***, **, and * denote two-tailed statistical significance at 1%, 5%, and 10% significance levels respectively. An F-test of BAE x Accruals = WAE x Accruals yields a p-value of 0.3949.

Table 8

Panel A: Estimation on Persistence with other Measures of Accruals Quality

Dependent Variable	**Earnings t+1**	**Earnings t+1**	**Earnings t+1**	**Earnings t+1**
	(p-value)	*(p-value)*	*(p-value)*	*(p-value)*
Accruals	0.2672***	0.2876***	0.3568***	0.3707***
	(0.003)	(0.002)	(0.001)	(0.000)
Operating Cash Flows	0.6315***	0.6453***	0.7091***	0.7636***
	(0.000)	(0.000)	(0.000)	(0.000)
Estimation	-0.0001*	-0.0001**	-0.0002***	-0.0001**
	(0.092)	(0.028)	(0.003)	(0.031)
Estimation x Accruals	-0.0009***	-0.0008***	-0.0007**	-0.0006***
	(0.001)	(0.000)	(0.015)	(0.007)
Estimation x Operating Cash Flows	-0.0001	-0.0000	0.0001	0.0001
	(0.697)	(0.940)	(0.742)	(0.836)
Stdev(DD Residual)	-0.0956***			-0.0955***
	(0.000)			(0.001)
Stdev(DD Residual) x Accruals	-0.6143***			-0.3338***
	(0.000)			(0.002)
Stdev(DD Residual) x Operating Cash Flows	0.1903*			0.2280*
	(0.096)			(0.063)
Abs(Accruals)		-0.0552*		-0.0693*
		(0.094)		(0.052)
Abs(Accruals) x Accruals		-0.3035***		-0.2212***
		(0.000)		(0.002)
Abs(Accruals) x Operating Cash Flows		-0.1431***		-0.0974
		(0.003)		(0.128)
Special Items			0.0070	0.0134**
			(0.170)	(0.016)
Special Items x Accruals			-0.1870***	-0.1643***
			(0.000)	(0.000)
Special Items x Operating Cash Flows			-0.1117***	-0.1110***
			(0.000)	(0.000)
Constant	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Controls Interact w/ Cash earnings	-	-	Yes	Yes
Controls Interact w/ Accrual earnings	-	-	Yes	Yes
Fixed Effects	Industry, Year, Auditor	Industry, Year, Auditor	Industry, Year, Auditor	Industry, Year, Auditor
Cluster	Industry, Year	Industry, Year	Industry, Year	Industry, Year
Observations	39,943	43,155	43,155	39,943
Adjusted R-squared	0.572	0.571	0.574	0.578

Notes: Table 8 Panel B presents the results for our tests of estimation on the persistence of cash flows and accruals including other measure of accruals quality found in the accounting literature. The controls include size, operating cycle, standard deviation of sales, standard deviation of operating cash flows, the number of years over the past 5 years in which the firm had negative earnings, and the total length of the footnotes. Please, refer to section 4 of the study for detailed descriptions of each of the variables. P-values are reported in parenthesis below their respective coefficients. All continuous variables are winsorized at 1% and 99% of their respective sample distributions. ***, **, and * denote two-tailed statistical significance at 1%, 5%, and 10% significance levels respectively.

Table 8

Panel B: BAE and WAE on Persistence with other Measures of Accruals Quality

Dependent Variable	**Earnings t+1**	**Earnings t+1**	**Earnings t+1**	**Earnings t+1**
	(p-value)	*(p-value)*	*(p-value)*	*(p-value)*
Accruals	0.2637***	0.2964***	0.3667***	0.3565***
	(0.002)	(0.002)	(0.001)	(0.000)
Operating Cash Flows	0.6186***	0.6448***	0.7035***	0.7447***
	(0.000)	(0.000)	(0.000)	(0.000)
BAE	-0.0001*	-0.0002*	-0.0002***	-0.0002**
	(0.083)	(0.057)	(0.007)	(0.012)
BAE x Accruals	-0.0008**	-0.0011***	-0.0008**	-0.0006*
	(0.015)	(0.001)	(0.031)	(0.073)
BAE x Operating Cash Flows	0.0001	0.0000	0.0002	0.0003
	(0.889)	(0.983)	(0.698)	(0.528)
WAE	-0.0001	-0.0001*	-0.0001**	-0.0001
	(0.263)	(0.079)	(0.015)	(0.170)
WAE x Accruals	-0.0007*	-0.0006*	-0.0005	-0.0005
	(0.050)	(0.092)	(0.153)	(0.172)
WAE x Operating Cash Flows	-0.0004	-0.0001	0.0000	-0.0002
	(0.123)	(0.723)	(0.938)	(0.491)
Stdev(DD Residual)	-0.0954***			-0.0942***
	(0.000)			(0.001)
Stdev(DD Residual) x Accruals	-0.6190***			-0.3566***
	(0.000)			(0.001)
Stdev(DD Residual) x Operating Cash Flows	0.1888			0.2302*
	(0.105)			(0.065)
Abs(Accruals)		-0.0583*		-0.0722*
		(0.077)		(0.050)
Abs(Accruals) x Accruals		-0.2914***		-0.2031***
		(0.000)		(0.007)
Abs(Accruals) x Operating Cash Flows		-0.1519***		-0.1050*
		(0.001)		(0.089)
Special Items			0.0069	0.0139**
			(0.172)	(0.013)
Special Items x Accruals			-0.1872***	-0.1670***
			(0.000)	(0.000)
Special Items x Operating Cash Flows			-0.1118***	-0.1134***
			(0.000)	(0.000)
Constant	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Controls Interact w/ Cash earnings	-	-	Yes	Yes
Controls Interact w/ Accrual earnings	-	-	Yes	Yes
Fixed Effects	Industry, Year, Auditor	Industry, Year, Auditor	Industry, Year, Auditor	Industry, Year, Auditor
Cluster	Industry, Year	Industry, Year	Industry, Year	Industry, Year
Observations	39,943	43,155	43,155	39,943
Adjusted R-squared	0.572	0.571	0.574	0.578

Notes: Table 8 Panel B presents the results for our tests of estimation on the persistence of cash flows and accruals including other measure of accruals quality found in the accounting literature. The controls include size, operating cycle, standard deviation of sales, standard deviation of operating cash flows, the number of years over the past 5 years in which the firm had negative earnings, and the total length of the footnotes. Please, refer to section 4 of the study for detailed descriptions of each of the variables. P-values are reported in parenthesis below their respective coefficients. All continuous variables are winsorized at 1% and 99% of their respective sample distributions. ***, **, and * denote two-tailed statistical significance at 1%, 5%, and 10% significance levels respectively.

Table 9

Estimation and the Mapping of Accruals into Cash Flows

Dependent Variable	**Stdev(DD Residual)** *(p-value)*	**Stdev(DD Residual)** *(p-value)*	**Stdev(DD Residual)** *(p-value)*
Estimation		0.1050***	
		(0.002)	
BAE			0.1203***
			(0.004)
WAE			0.1077***
			(0.001)
Length		0.0330***	0.0590***
		(0.001)	(0.000)
log(Market Value)	-0.0016**	-0.0025***	-0.0024***
	(0.015)	(0.000)	(0.000)
sum(NEGEARN)	0.0124***	0.0115***	0.0117***
	(0.000)	(0.000)	(0.000)
Stdev(Sales)	0.0493***	0.0476***	0.0472***
	(0.000)	(0.000)	(0.000)
Stdev(Operating Cash Flows)	0.4552***	0.4617***	0.4602***
	(0.000)	(0.000)	(0.000)
Avg(Operating Cycle)	0.0036***	0.0033***	0.0036***
	(0.001)	(0.003)	(0.001)
Constant	0.0129	0.0154	0.0124
	(0.781)	(0.732)	(0.788)
Fixed Effects	Industry, Year, Auditor	Industry, Year, Auditor	Industry, Year, Auditor
Cluster	Industry, Year	Industry, Year	Industry, Year
Observations	41,648	41,648	39,198
Adjusted R-squared	0.418	0.421	0.418

Notes: Table 9 shows the association between estimation and accruals quality, as measure by Dechow and Dichev [2002]. P-values are reported in parenthesis below their respective Please refer to section 4 of the study for detailed descriptions of each of the variables. Length is scaled by 100,000. All continuous variables are winsorized at 1% and 99% of their respective sample distributions. ***, **, and * denote two-tailed statistical significance at 1%, 5%, and 10% significance levels respectively. An F-test of BAE = WAE yields a p-value of 0.6768.

Table 10

Estimation and Abnormal Returns

Dependent Variable	**Abnormal Returns t+1** *(p-value)*	**Abnormal Returns t+1** *(p-value)*	**Abnormal Returns t+1** *(p-value)*
Accruals	-0.0946***	0.0198***	0.0037
	(0.004)	(0.004)	(0.960)
Estimation		-0.0000	
		(0.641)	
Estimation x Accruals		-0.0004***	
		(0.001)	
BAE			0.0002
			(0.374)
BAE x Accruals			-0.0050***
			(0.000)
WAE			0.0002
			(0.253)
WAE x Accruals			0.0013
			(0.304)
Length	-0.0163	0.1506*	0.1040
	(0.105)	(0.050)	(0.204)
Length x Accruals	0.0000	0.0000***	0.0000***
	(0.179)	(0.000)	(0.005)
log(Market Value)	0.0009***	-0.0159***	-0.0138***
	(0.002)	(0.000)	(0.000)
BTM	0.0105***	0.1770***	0.1833***
	(0.000)	(0.000)	(0.000)
Beta	-0.0054***	0.0321***	0.0333***
	(0.000)	(0.000)	(0.000)
ETP	0.0042***	-0.0647***	-0.0688***
	(0.000)	(0.000)	(0.000)
Constant	-0.1656***	-0.0032	-0.1640***
	(0.000)	(0.309)	(0.000)
Fixed Effects	Year	Year	Year
Observations	51,931	51,841	49,249
Adjusted R-squared	0.066	0.006	0.068

Notes: Table 10 presents the association between 1 year abnormal earnings beginning 5 days following the filing date and accrual earnings. All regressions were estimated using ordinary least squares. P-values are reported in parenthesis below their respective coefficients. Refer to section 4 of the study for detailed descriptions of each of the variables. Length is scaled by 100,000. All continuous variables are winsorized at 1% and 99% of their respective sample distributions. ***, **, and * denote two-tailed statistical significance at 1%, 5%, and 10% significance levels respectively.

Table 11

Fama-French Carhart Four-Factor Alpha

Panel A: Accruals Estimation

	Equal Weighted			Value Weighted		
	Accruals			Accruals		
	1	**5**	**Hedge**	**1**	**5**	**Hedge**
All Observations	0.7583***	0.0817	0.5816**	0.0745	-0.3200**	0.3945
	(0.005)	(0.478)	(0.029)	(0.752)	(0.027)	(0.176)
Estimation	**1**	**5**	**Hedge**	**1**	**5**	**Hedge**
1	1.0450***	-0.0150	1.0599***	0.7996***	-0.3029	1.1025***
	(0.000)	(0.946)	(0.000)	(0.009)	(0.253)	(0.005)
5	0.7104**	-0.0329	0.7433**	-0.0693	-0.0769	0.0077
	(0.040)	(0.855)	(0.025)	(0.855)	(0.688)	(0.986)

Panel B: BAE and WAE

	Equal Weighted			Value Weighted		
	Accruals			Accruals		
	1	**5**	**Hedge**	**1**	**5**	**Hedge**
All Observations	0.7218**	0.1402	0.5816**	-0.0456	-0.2483	0.2027
	(0.011)	(0.478)	(0.029)	(0.847)	(0.151)	(0.533)
BAE	**1**	**5**	**Hedge**	**1**	**5**	**Hedge**
1	0.8726**	0.1731	0.6995***	0.1059	-0.2903	0.3962
	(0.011)	(0.472)	(0.009)	(0.705)	(0.251)	(0.268)
5	0.8072***	0.0559	0.7513***	0.0848	-0.1673	0.2521
	(0.006)	(0.739)	(0.010)	(0.801)	(0.397)	(0.515)
WAE	**1**	**5**	**Hedge**	**1**	**5**	**Hedge**
1	-0.3726	0.4440***	0.7165***	0.2027	0.2573	0.4982
	(0.239)	(0.007)	(0.002)	(0.533)	(0.250)	(0.226)
5	0.7684**	0.0228	0.5219	0.2705	-0.2028	-0.0018
	(0.023)	(0.915)	(0.119)	(0.403)	(0.478)	(0.996)

Notes: Table 11 presents the results for the Carhart Four-Factor Alpha. All regressions were estimated using ordinary least squares. P-values are reported in parenthesis below their respective coefficients. ***, **, and * denote two-tailed statistical significance at 1%, 5%, and 10% significance levels respectively.

Appendix 1: Overview of Grammatical Relationships

In this study we are interested in determining when estimation was needed by management during the accrual generating process. Since estimation is a function of the actions taken during the accrual generating process and not necessarily of the accruals themselves it is difficult to determine the amount of estimation by simply examining the magnitude of a company's accruals. For public companies the notes to the financial statements proves a wealth of information about the accrual generating process and, more importantly, information about the estimations needed by management.

Unlike the accruals which are denoted quantitatively, the notes to the financial statements, and hence the information pertaining to the estimation involved, is qualitative in nature. More importantly to assess when an estimate was needed by management we need to infer from the words and the placement of the words in the sentence (i.e. the grammatical relationships) when an estimate was needed during the accrual generating process.

The words and the grammatical relationships that are used in written language are not random and are in fact highly structured. For example in the sentence "I like football." The object "football" is the target of "like". Since the word "like" is conveying the enjoyment of something and "like" is targeting "football" this linguistic cue is conveying that football is enjoyed. More over the association between "I" and "like" denote that the person performing the action is "I". Even though this example may be simple it illustrates a powerful idea and provides us with a structure to help us infer meaning from the qualitative footnotes.

To infer meaning from the footnotes we therefore need the grammatical relationships in each sentence in the notes to the financial statements and several dictionaries of terms associated with estimation (Appendix 2). To get the grammatical relationships in each sentence we use a technique pioneered in the field of Natural Language Processing called Statistical Parsing to map the structure of the sentences in the notes to the financial statements. We use a specific implementation of statistical parsing from the Stanford Natural Language Processing Group – see http://nlp.stanford.edu/software/index.shtml

for details (Marneffe et al [2006]). Essentially, this implementation finds the most likely map of the sentence by matching the sentence to a tree bank of manually parsed sentences to find the layout of the sentence which is most likely.

To illustrate this technique, the following sentence was parsed using the Stanford parser.

"We estimated receivables and purchased inventory."

[nsubj(estimated-2, We-1), nsubj(purchased-5, We-1), dobj(estimated-2, receivables-3), conj_and(estimated-2, purchased-5), dobj(purchased-5, inventory-6)]

We see that the object "receivables" is the direct object of the action (verb) "estimated". This linguistic cue indicates that the sentence is conveying that receivables were estimated. As illustrated, using the grammatical relationships removes any ambiguity about the meaning of the sentence.

Appendix 2: Estimation Dictionaries and Grammatical Relationships

Estimation Dictionaries

We construct four dictionaries to help measure the estimation conveyed in each companies notes to the financial statements:

Estimation Actions

Estimate, Estimating, Estimated, Anticipate, Anticipates, Anticipating, Anticipated, Approximate, Approximates, Approximated, Approximating, Assess, Assesses, Assessed, Assessing, Believe, Believed, Believes, Believing, Determine, Determined, Determining, Determines, Evaluate, Evaluated, Evaluating, Evaluates, Expect, Expects, Expected, Expecting, Forecast, Forecasts, Forecasted, Forecasting

Estimation Objects

Estimate, Estimates, Estimation, Estimations, Approximation, Approximations, Assumption, Assumptions, Belief, Beliefs, Forecast, Forecasts

Estimation Adjectives

Estimated, Anticipated, Approximately, Expected, Forecasted, Likely, Probable

Use Words

Make, Makes, Made, Making, Use, Uses, Used, Using, Include, Includes, Included, Including

Estimation Actions are verbs which convey that an estimation action was performed (e.g. "we estimated accruals"). Estimation Objects are estimation related objects/nouns (e.g. "we used estimates"). Estimation Adjectives modify an object to convey that the object was estimated (e.g. "estimated accruals"). Lastly, the Use Words dictionary contains action words which convey that something was used or done by management.

Grammatical Relationships

We use the following grammatical relationships in conjunction with the Estimation Dictionaries above to find the linguistic cues that infer that estimation was needed.

Direct Object

The accusative object of an action (i.e. "estimate receivables" or "used estimates"). For this grammatical relationship we look for when a word from the Estimation Action dictionary targets some object or when a word from the Use Word dictionary targets a word from the Estimation Object dictionary. An example of the first case is "estimate receivables" here the action "estimate" targets the object "receivables" thereby implying that receivables were estimated. An example of the second scenario is "used estimates". In this example one of the words from the Use Words dictionary, "used", targets a word from the Estimation Object dictionary, "estimates", which implies that they used an estimate.

(Passive) Nominal Subject

This grammatical relationship is similar to the direct object in that it relates information about an object. This grammatical association is similar to the Direct Object and the relationships that we look for are the same as those for the direct object.

Adjective Modifier

Adjective Modifies modify the meaning of an object (i.e. "likely receivable" or "anticipated value"). For this grammatical relationship we find when a word from the Estimation Adjective dictionary targets some object – this implies that the object was estimated.

Quantifier Phrase Modifier

This grammatical relationship is a modifier to a number (i.e. "approximately $100"). This grammatical association is similar to the adjective modifier but specific to numbers. We look for when a word from the Estimation Adjectives dictionary targets a number.

Noun Compound Subjects

A noun used to modify another noun (i.e. “value estimates”). This grammatical relationship is similar to the adjective modifier except that a word from the Estimation Object dictionary is modifying another object.