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

Preliminary – January 6th, 2012

Abstract

This paper directly examines the link between the amount of estimation needed during the accrual generating process and the persistence of earnings. We measure the amount of estimation during the accrual generating process by counting the number of linguistic cues in the notes to the financial statements that convey that estimation was needed. Consistent with the conjectures in Sloan [1996], we find that accrual earnings which needed more estimation are less predictive of future earnings. We also find that greater estimation is associated with a lower mapping of accrual earnings into the past, current, or future cash flows. Lastly, we find some evidence that the market reacts as if it does not incorporate the amount of estimation in accruals earnings into its valuation of the firm in a timely manner. Overall, our results suggest that the estimation needed during the accrual generating process plays an important in understanding the persistence of accrual earnings.

We thank members of the University of Michigan Accounting Department and Michigan State 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

This study examines whether accrual earnings which required more estimation are less predictive of future earnings (i.e. less persistent). Sloan [1996] and Richardson et al. [2005] argue that greater estimation needed during the accrual generating process[1] explains why accrual earnings are less persistent than cash earnings. Unlike cash earnings, accrual earnings incorporate estimates of future cash flows, cash flow deferrals, depreciation and amortization, and fair value estimates; the amount of estimation in accrual earnings is greater than that needed when estimating cash earnings. Richardson et al [2005] argues that greater estimation of accrual earnings suggests that accrual earnings are recognized with lower precision and therefore will be less predictive of the future earnings of the firm.

Few studies have directly examined the estimation involved in the accruals generating process. Rather, prior research has primarily examined the different components that comprise accrual earnings to explain accruals persistence. For example, Dechow et al. [2006] find that low accrual earnings exhibit lower persistence when they contain special items. Richardson et al. [2005] partitions accrual earnings into three distinct components: working capital accruals, non-current operating accruals, and financing accruals. They find evidence that accrual earnings due to financing are more persistent than working capital and non-current operating accrual earnings. While these and other studies provide valuable insight into accrual earnings persistence they do not directly test the implications of accruals estimation on persistence.[2] Moreover, these studies have not directly examined Sloan's conjecture that the estimation of accruals explains the lower persistence of accruals earnings.

We measure the amount of estimation in accrual earnings using the amount of estimation involved in the accruals recognition process, calculated as the number of estimation related

[1] The accrual generating process denotes the accounting process used by the firm to calculate accruals.
[2] Conclusions about accruals estimation inferred from these studies are confounded by the economics of the firms since accruals are correlated with the economic fundamentals (e.g., growth status) of the firm.

linguistic cues in the notes to the financial statements. To do this we first construct a dictionary of estimation words by examining dozens of footnotes and compiling a list of common estimation related words. We then count the number of times these estimation words target an object (e.g. "we *estimated* receivables"), are the target of an action (e.g. "we used *estimates*"), or are an adjective to an object in a sentence (e.g. "*estimated* costs"). (See. Appendix 2)

This textual analysis approach is used because the amount of estimation needed when recording accruals is difficult to determine by simply examining the amount of accruals. The amount of estimation in the recorded accruals is determined by the accrual generating process used to arrive at the final number rather than the magnitude of the final number itself. Hence, the accrual earnings of two firms could be identical in magnitude but the amount of estimation in one firm's accrual earnings could differ drastically from the others. Since this is the case, we naturally look towards the qualitative portion of a firm's financial disclosures for information pertaining to the accrual generating process.

We focus on the notes to the financial statements of the 10-K because they provide 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 et al. [2010]). More importantly, this section provides information pertaining to the estimations made and the assumptions needed during the accrual generating process. While other sections of the 10-K, such as the Management Discussion and Analysis (MD&A), may also provide some information about the accruals process, these other sections do not typically include details about the accrual generating process. Moreover, much of the accounting related information in these other sections is reiterated in the notes to the financial statements.

We hypothesize that accrual earnings which required more estimation are less predictive of future earnings. Our findings are consistent with this hypothesis. We find that when there is more estimation, accrual earnings are significantly less persistent but that there is no difference

in the persistence of cash earnings. These findings lend support to Sloan's (1996) conjecture that the amount of estimation in accrual earnings partially explains their lower persistence. Moreover, this finding indirectly reinforces his hypothesis that the difference in the amount of estimation needed for accrual earnings in comparison to cash earnings partially explains the difference in their persistence.

Next, we find that accruals that involve more estimation 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 accrual earnings are less precise, and therefore have greater error, and provides further evidence that the estimation of accruals drives the lower persistence.

Lastly, we find some evidence that the market reacts as if they do not incorporate the amount of estimation in accruals into their valuation of the firm in a timely manner. Specifically, we find that the accrual anomaly documented in Sloan [1996] is more significant when more estimation is needed during the accruals generating process.[3] This is consistent with the hypothesis that the market more greatly overvalues more estimated accrual earnings in the short term and vice versa for negative accrual earnings.

By providing direct evidence of the association between the estimation in accrual earnings and accruals persistence, this paper makes several contributions to the literature. First, this study directly incorporates manager's estimation into our understanding of the persistence of accruals. Many prior studies have ignored this important component of the accrual process and have simply focused on the actual accrual numbers. As such, they have ignored the differences in the estimation component in accrual generating process and hence the difference in the persistence of the accruals due to estimation.[4] Our study addresses this issue and directly

[3] Accumulation of abnormal returns begins 5 days after the 10-K filing.
[4] Richardson et al. [2005] indirectly test the association between accruals estimation and persistence.

incorporates the accruals estimation information found in the notes to the financial statements. In doing so, our finding suggests that we can't simply take accrual numbers at face value but rather we need to understand the process used when deriving the accrual to understand the complete picture of accruals.

Second, our findings strengthen Sloan's argument that the estimation of accruals explains differences in the persistence of cash earnings and accrual earnings (Sloan [1996], Richardson et al. [2005]). Until now the argument that the difference in the persistence of accrual earnings and cash earnings is driven by the difference in the amount of estimation of the two components of earnings has not been directly tested. Our findings suggest that estimation does partially explain the lower persistence of accrual earnings. Additionally, while this finding does not directly examine the differences in the estimation between the cash and accrual components of earnings, our findings provide some evidence consistent with this hypothesis.[5]

Finally, this study contributes to the textual analysis accounting literature by using statistical parsing to extract meaning from qualitative financial information. This method provides structure to how we examine the qualitative information. Additionally, this study adds to a growing field of textual analysis studies which suggest that qualitative information can help provide a richer understanding of the firm than that conveyed from simply examining the quantitative information (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 our sample of 10-K footnotes and financial information was prepared. In Section 4 we present the research design and main results and in Section 5 we present some robustness tests. Section 6 concludes the paper.

[5] Some suggest that no estimation is required for cash earnings.

2. Motivation and Prior Literature

Sloan [1996] finds that accrual earnings are less predictive of future earnings than cash earnings (i.e. accruals earnings are less persistent than cash earnings). He argues that the difference between the persistence of accrual earnings and cash earnings is due to the greater estimation in accrual earnings. Accrual earnings incorporate estimates of future cash flows, depreciation and allocations, deferrals, and valuations which are not needed when recording cash earnings (Richardson et al. [2005]). Since accrual earnings are more greatly estimated than cash earnings, they are less precise and therefore are less persistent than cash earnings.

Richardson et al. [2005] expands upon the hypothesis in Sloan [1996] and formally models accruals estimation as an error-in-variables problem, the errors in variables problem is also known as a measurement error in the independent variable problem. Their model suggests the existence of accrual earnings which truly reflect perfect foresight of the wealth generated by the firm. These perfect foresight accruals are measured without error and therefore the firm will not need to correct accrual errors in future periods. However, actual recorded accruals are measured with error and are estimated as the perfect foresight accrual plus some error. This measurement error thereby reduces the association between accrual earnings and future earnings (i.e. lowers their persistence).

The estimation of accrual earnings is a function of the process and estimations needed when recording the accrual and not necessarily a reflection of the amount of accruals. Even if two companies have the same total amount of accruals, the amount estimation in the accruals in these two companies may be vastly different. Simply examining the total amount of accrual earnings will not provide information about the firm's accrual generating process nor will it provide information about the estimation of accrual earnings. 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 less estimated financial accruals. Even if the total amount of accrual earnings is

the same, the degree of estimation between the two company's accrual earnings may be vastly different.

Prior studies have examined how specific components that comprise accruals earnings affect how well accruals predict future earnings. However, none of them have directly explored how the estimation of accrual earnings influences their persistence. For instance, Dechow and Ge [2005] examines the persistence of low accrual earnings when the firm has special items. Consistent with their hypothesis, they find that accrual earnings are less persistent when the firm has special items. Richardson et al. [2005] disaggregates accrual earnings 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 accrual earnings. The greater estimation in these accruals implies that these accruals are less likely to be realized in the cash flows of the firm 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 and directly measure and examine the implications of accruals estimation on accruals persistence. Specifically, we use a textual analysis approach to examine the notes to the financial statements and measure the amount of estimation needed during the accruals generation process. Even though prior studies have provided interesting insight into the persistence of accrual earnings, none of them 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: *Accrual earnings that involve more estimation are less persistent.*

Dechow et al. 2002 find that firms which exhibit a lower mapping of accrual earnings 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, accrual earnings will be less predictive of future earnings.

Following a similar thought process, if there is greater estimation in accrual earnings then accrual earnings are likely to be recorded with lower precision (i.e. accrual earnings map less into realized cash flows). Greater estimation in accrual earnings implies that managers needed to make more estimates during the accruals 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, and/or subsequent periods. Following this reasoning, more estimation in recorded accruals will be associated with a lower mapping of cash flows into accrual earnings.

Prediction 2: *Accrual earnings 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 the accruals portion of earnings is not quickly incorporated by investors into their valuation of the firm (Sloan [1996], Hanlon [2005], Richardson et al. [2005]). One explanation is that investors fixate on total earnings thereby disregarding the lower persistence of accruals earnings' affect on how predictive current earnings are of future earnings (Sloan [1996]). Accordingly, Sloan [1996] finds that future abnormal returns of the firms are negatively associated with the firms' accrual earnings. This

finding is consistent with his hypothesis that investors misinterpret the persistence of accrual earnings. If investors fixate on earnings and ignore the accruals portion of earnings, which could be calculated directly from the statement of cash flows, 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 accrual earnings) and over value firms with more estimated income increasing accruals (positive accrual earnings). Hence, we would find that the estimation information in the footnotes is informative of future long term abnormal returns.

On the other hand, investors may quickly incorporate the amount of estimation in accrual earnings into their valuation of the firm since this information is available in the firm's disclosures. 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, and hence the lower persistence, of accrual earnings can be found in the footnotes then investors may become informed of the lower persistence of these earnings. If so, then investors will quickly incorporate this information and the estimation information in the footnotes will not be informative of future long term abnormal returns.

Prediction 3: *The market reacts as if it does not incorporate the amount of estimation in accrual earnings into 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 1994 and 2010 from the SEC Edgar Website. We then extract the footnotes to the financial statements from each of the 10-K filings using Perl. The extracted footnotes were stripped of all html tags

and tables. Footnotes were truncated at the 1% and 99% level of length. Since there isn't strict standardization in the headers of each section of the 10-K the script wasn't always able to identify the notes sections. Some section (item) headers in the 10-K filing differed significantly from the suggested section header defined by the SEC (SEC Form 10-K General Instructions). To mitigate the impact of these non-standard filings we include many of the non-standard section headers into our Perl script. We verified the accuracy of our script by hand collecting the footnotes to the financial statements from approximately 50 randomly chosen 10-K filings and compared the extracted notes to our hand collect sample. Overall, our parser was able to correctly extract the footnotes section in well over 90% of the randomly selected test sample. In total we extracted 85,691 footnotes from the 10-K filings.

3.2 Measuring Estimation

We use a NLP statistical parsing technique to determine the amount of estimation used during the accrual generating process. More specifically, we use this technique to extract estimation information from the notes to the financial statements. Many prior studies have used simple word counts to extract information from qualitative disclosures (Loughran et al [2011], Rogers et al [2011]). While word counts perform reasonably well for certain questions this technique's ability to infer meaning from sentences in qualitative disclosures is limited.

Our NLP approach uses an implementation of statistical parsing by the Stanford NLP group to map the structure of each sentence in the notes to the financial statements. The map of each sentences' structure identifies its noun modifiers, direct object modifiers, adjective modifier, etc. Deconstructing sentences in this manner allows us to identify specific linguistic cues which indicate that estimation was needed and thus improves our ability to measure the amount of estimation needed during the accruals generating process (Klein et al [2003], see Appendix 1).

We begin this procedure by first constructing a dictionary of estimation related words. The dictionary of estimation terms is limited to the common estimation words and therefore may not encompass all the words that firm *could* use to convey estimation. The alternative approach would have been to use a larger, more expansive, dictionary of estimation term. However, using a larger dictionary would induced noise in our measure because some firms may use certain terms to denote estimation while other firms may use the same term in other ways. While our dictionary of common estimation terms is small, the use and interpretation of each of the words is specific. Another issue with the dictionary approach is that the meaning of the sentence is not used. That is, the estimation word in the sentence may not be used to denote that an accrual was estimated. This issue would positively inflate the estimation count and adds noise to our measure. The small dictionary mitigates this problem because of the limited scope of how each of the words in the dictionary can be used.

Next, each sentence of the notes to the financial statements was mapped using the Stanford open source statistical parser (Marneffe et al [2006]). We then examine the mapping of each sentence using the dictionary of estimation words for linguistic cues which convey that estimation was needed or used (see Appendix 2).[6] The count of the number of linguistic cues (hereafter "estimation count") and the count of the number of linguistic cues scaled by the number of sentences (hereafter "scaled estimation count") are used as our measures of the amount of estimation.

Table 1 presents the average estimation count, the average total number of sentences and the scaled estimation count found in the footnotes which are estimation words. Consistent with prior studies, we find that the average length of the footnotes to the 10-K has steadily increased over time (Li [2008]). On average, the length of the notes to the financial statements has doubled

[6] 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

in size over our sample period. 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 the estimation count and the scaled estimation count have both increased monotonically over the years.

Figure 1 plots the trend in the estimation count and in the length of the footnote (as measured by the total number of sentences in the footnote). Overall, the length of the footnotes has been growing over the years in our sample. Moreover, the estimation count has been growing as well. Prior to the 2001 fiscal period the growth of the length and the number of estimation words appears identical. However, after 2001 the growth of the number of sentences and the estimation count has dramatically increased.

Table 2 shows the average estimation count, number of sentence in the footnotes, and the scaled estimation count by industry. Overall, industry appears to play an important role in the amount of estimation and the length of the notes to the financial statements. The bottom five industries by footnote length include fishing, forestry, miscellaneous repair services, home improvement, and furniture and fixtures average approximately 200 sentences in length. On the other hand the top five industries by footnote length average approximately 400 sentence and include tobacco products, utilities, coal mining, depository institution, and insurance carries.

3.3 Sample Preparation

We merge the estimation count with annual financial information from the Wharton Research Data Services (WRDS) Compustat database. Financial information was not available for 34,270 of the extracted footnotes. The primary reason for many of the non-matches stemmed from not being able to find an appropriate gvkey for the cik specified in the header of the 10-K filing. After merging the footnotes with Compustat we are left with 52,377 firm year observations.

The future long window abnormal returns of the firm were 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 following the filing of the 10-K beginning five days after the filing of the 10-K. We also calculate 1 year value weighted compounded market beginning five days after the 10-K filing for each of the firms in our sample. We then calculate the firms' compounded abnormal returns for the firm by subtracting the 1 year value weight compounded market returns from the 1 year compounded returns of the firm.

4. Research Design and Results

4.1 Determinates of Estimation

We examine some of the possible determinates of the estimation count and scaled estimation count in the notes to the financial statements: 1. Sentences– longer footnotes are more likely to have greater estimation count (lower scaled estimation count) 2. Size – larger firms are more likely to need more estimates to properly measure the cash flows of their firm and how they map into accrual earnings 3. Operating Cycle – longer operating cycles imply that more estimates of future cash receipts are needed. 4. Standard Deviation of Cash Earnings – less stable cash flows suggests that a greater number of estimates are needed to properly estimate the cash flows associated with an accrual 5. Standard Deviation of Sales – less stable sales suggest more volatile operating and therefore more estimates and approximations are needed 6. Big Four Auditor – the auditor is intimately tied to the accruals generating process; higher quality auditors are less likely to need estimation when recording accruals; 7. Book-to-Market – firms which are growing need more accounting estimation to convey their growth; 8. Negative Accruals and Negative Earnings – negative accruals and/or earnings are unstable times in the firm's life and therefore may require more accounting estimation to reflect this; 9. Number of Accounting Items – the number of non-empty annual financial items in Compustat indicates the complexity of the firms accounting; more complexity accounting requires more estimation; 10 Special Items –

special items suggest that some event outside of normal operation may have taken place and its impact may need to be estimated.

$$est_{f,t} = \beta_0 + \beta_1 * Total\_Words_{f,t} + \beta_2 * \text{size}_{f,t} + \beta_3 * btm_{f,t} + \beta_4 \quad (1)$$
$$* Operating\_Cycle_{f,t} + \beta_5 * stdev\_sales_{f,t} + \beta_6$$
$$* stdev\_cashearnings_{f,t} + \beta_7 * BIGFOUR_{f,t} + \beta_8$$
$$* NEGEARNINGS_{f,t} + \beta_9 * NITEMS_{f,t} + \beta_{10} * SpecialItems_{f,t} + \epsilon_{f,t}$$

Where $est_{f,t}$ = the amount of estimation in the footnotes measured as either the number of estimation related linguistic cues (estimation count) or the estimation count scaled by the number of sentences; $size_{f,t}$ = log of the market value of the firm = log(prcc*csho);$btm_{f,t}$= book to market ratio calculated as total assets / (market value of equity + liabilities); $Operating\_Cycle_{f,t}$= operating cycle of the firm calculated as 360/(sales/average receivables) + 360/(cost of goods sold/average assets); $stdev\_sales_{f,t}$= standard deviation of sales for a 10 year rolling window (5 firm year observations minimum per window); $stdev\_cashearnings_{f,t}$= standard deviation of cash earnings for a 10 year rolling window (5 firm year observations minimum per window); $BIGFOUR_{f,t}$= 1 if the firm has a big four auditor; $NEGEARNINGS_{f,t}$= flag indicating that total earnings are negative; $NITEMS_{f,t}$= number of non-empty annual financial items listed for the firm in Compustat; $SpecialItems_{f,t}$= 1 when special items are less than two percent of average total assets.

4.1.1 Determinates of Estimation Findings

Table 5 presents results for the determinates of the amount of estimation. We use a simply linear regression model to test the relationship between determines and the estimation count and the scaled estimation count of the notes to the financial statements. Several of the

characteristics including the length of the footnotes, size of the firm, number of non-empty annual items in Compustat, and whether the firm has negative accruals or negative earnings are consistent with our hypotheses. Two of the three determines of the accruals quality identified in Dechow and Dichev 2005, operating cycle and standard deviation of cash flows, are statistically significant but the coefficient is opposite of what we predicted. Additionally, the result for book-to-market is not consistent with our predictions.

4.2 Estimation and the Persistence of Accrual Earnings

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

First, we examine the marginal effect of estimation on the persistence of total earning. 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 they are for a given level of estimation. If our hypothesis is correct then we should find a negative coefficient on this interaction term.

$$Earnings_{f,t+1} = \beta_0 + \beta_1 * Earnings_{f,t} + \beta_2 * est_{f,t} + \beta_3 * Earnings_f * est_{f,t} \quad (2)$$
$$+ \Sigma\beta_i * controls_{f,t} + Earnings_{f,t} * \Sigma\beta_j * controls_{f,t} + \Sigma\beta_k$$
$$* fyear + \Sigma\beta_k * sic\_two + \epsilon_{f,t}$$

Next, we disaggregate earnings into cash earnings and accrual earnings and interact each component with the estimation measure.[7] If increased estimation lowers the ability of current year's accruals earnings to predict following years earnings then the interaction between estimation and accrual earnings will be negative. Ideally, if the number of estimation cues in the footnotes does not capture the uncertainty in the cash flows the interaction between cash earnings and estimation should be insignificant.

$$
\begin{aligned}
Earnings_{f,t+1} \quad & (3)\\
&= \beta_0 + \beta_1 * cash_{f,t} + \beta_2 * accruals_{f,t} + \beta_3 * est_{f,t}\\
&+ \beta_4 * cash_{f,t} * est_{f,t} + \beta_5 * accruals_{f,t} * est_{f,t} + \Sigma\beta_i * controls_{f,t}\\
&+ accruals_{f,t} * \Sigma\beta_j * controls_{f,t} + cash_{f,t} * \Sigma\beta_j * controls_{f,t}\\
&+ \Sigma\beta_k * fyear + \Sigma\beta_k * sic\_two + \epsilon_{f,t}
\end{aligned}
$$

Where $Earnings_{f,t+1}$= firm f's operating income after depreciation in the following year = ibc_f_t+1; $Earnings_{f,t}$ = firm f's operating income after depreciation in the current year =oiadp_f_t; $cash_{f,t}$ = earnings from cash flows = oiancf_f_t; $accruals_{f,t}$= total accruals equals operating income after cash flows minus operating cash flows = ibc – oancf; $est_{f,t}$ = the amount of estimation in the footnotes measured as either the number of estimation related linguistic cues (estimation count) or the estimation count scaled by the number of sentences; $btm_{f,t}$= book to market = at_f_t divided by prcc*csho; $size_{f,t}$= log of the market value of the firm = log(prcc*csho); $negearn_{f,t}$= 1 if current earnings are less than 0 and 0 otherwise = oiadp < 0; sic_two = fixed effect for each industry. We also include interactions between all control

[7] The measure of estimation may also capture overall uncertainty about the firm. If so then the interaction between cash flows and the measure of estimation would also be significant.

variables and earnings and year fixed effects.[8] All continuous variables are scaled by average assets.

4.2.1 Estimation and the Persistence of Accrual Earnings Findings

Table 6 presents the results of interacting estimation with total earnings. This table presents the results when both the estimation count and the scaled estimation count are used as proxies for estimation. The coefficient on the interaction between the estimation count and earnings in the current year is negative and statistically significant. The coefficient on the interaction between the scaled estimation count and total earnings is negative and statistically significant, consistent with our hypothesis. Overall, these findings are consistent with our prediction that the amount of estimation in accrual earnings is associated with lower persistence of earnings.

As discussed before, accrual earnings are but 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 portion. Table 7 shows the results when we disaggregate total earnings into cash earnings and accrual earnings. In untabluated results, consistent with Sloan [1996] we find that the persistence of accrual earnings is less than that of cash earning. Moreover, the magnitudes of the coefficients are similar to those found in Dechow et al [2006].[9] Panel A shows the results when the estimation count is used as the measure of the estimation of accrual earnings. As predicted the interaction between accrual earnings and estimation is negative and statistically significant, which is consistent with our prediction that accrual earnings which are more subjective exhibit lower persistence. The difference in the persistence of accrual earnings between the 25th percentile of estimation count to the 75th percentile is approximately -0.0672

[8] 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.

[9] We compare the magnitudes to Dechow et al [2005] and not Sloan [1996] because we use the statement of cash to estimate accruals, as in Dechow et al [2005].

((72-40)*(-0.0016)). This is equivalent to approximately a 10% difference in persistence when going from the 25th percentile to the 75th percentile of estimation. Next, the results show that the coefficient on the interaction between cash earnings and the amount of estimation is statistically insignificant. This shows that the measure of estimation captures some aspect of the accrual earnings but not cash earnings. Panel B of Table 7 shows the results when the scaled estimation count is used as the measure of the estimation of accrual earnings. Similar to Panel A the coefficient between the interaction of percent and accrual earnings is negative. We also see that the interaction between the scaled estimation count and cash earnings is statistically insignificant. Economically, the difference in the persistence when going from the 25th percentile to 75th percentile is similar to that when using the estimation count.

4.3 Estimation and the Mapping of Cash Flows into Accruals

Prediction 2 (P2) suggests that more subjective accrual earnings are less likely to be realized as cash by the firm. The primary measure of how well accrual map into cash flows used in this study was developed by Dechow and Dichev [2002] (hereafter DD). This model captures accruals quality by measuring how well working capital accruals map to realized operating cash flows. This model is based on the idea that accruals are a way to shift the recognition of cash flows.[10] If the realized cash flows of a firm maps well into the accruals of the firm then the firms 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 accruals quality for the firm.

Our second measure is the absolute value of the residual accruals model developed in Jones [1991] (hereafter Jones Model), as modified in Dechow et al. [1995] (hereafter Modified

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

Jones Model) (Francis et al. [2005]). This model first estimates the level of normal accruals for a firm in a given year then uses estimate to calculate the abnormal (residual) portion of accruals. The normal amount of total accruals is estimated by regressing the total accruals of the firm on the difference in the change in revenues minus the change in receivables and on total property, plant and equipment. The coefficients from this regression are then use to estimate the normal amount of accruals. The abnormal (residual) portion of accruals is then the firms total accruals minus the estimated normal accruals.

Operationally the DD model is limited to working capital accruals (Dechow et al. [2002], White [2010]). Specifically, the model does not include non-current accruals because of the long-lag times between the recognition of the accrual and the eventual realization of the cash flow. Unfortunately, our measure of the amount of accruals estimation does not make a distinction between current and non-current accruals. Therefore, this may limit the ability of our measure to explain any variation in the DD measure. On the other hand the Jones Model accounts for current and non-current accruals and therefore does not suffer from the same limitation as the DD model. However, unlike the DD model, the Jones Model doesn't directly measure the quality of accruals. Rather, the Jones Model estimates the abnormal accruals of a firm. While an abnormal amount of accruals is likely correlated with the quality of the accruals it is not a direct measure of accruals quality. Even with each models limitations we believe that together they provide insight into accruals quality and therefore we include both measures in our tests.

The specification of the DD model is shown in equation (1). We include change in revenues and Property, Plant and Equipment (PPE) as proposed in McNichols [2002].

$$TCACC_{f,t} = \beta_0 + \beta_1 * CFO_{f,t-1} + \beta_2 * CFO_{f,t} + \beta_3 * CFO_{f,t+1} \quad (4)$$
$$+\beta_4 * \Delta Rev_{f,t} + \beta_5 * PPE_{f,t} + \epsilon_{f,t}$$

Where $CFO_{f,t}$ = firm f's operating cash flows for fiscal year t = oancf_f,t; $TCACC_{f,t}$ = total working capital accruals of firm f during fiscal period t = ibc_f,t – oancf_f,t + dp_f,t; $\Delta Rev_{f,t}$= change in sales from the prior year = sale_f,t – sale_f,t-1; $PPE_{f,t}$= firm f's total property plant and equipment for the current fiscal period = ppegt_f,t. All variables are scaled by average total assets; average total assets is the current period assets (at_f,t) plus the firms prior period assets (at_f,t-1) divided by 2.

We follow prior research and make two adjustments to the DD model to attain a yearly measure of accruals quality (Ogneva [2008], Francis et al. [2005]). First, we estimate equation (1) by industry rather than over the life of the firm - We define a firms industry as its two digit SIC code. Second, the absolute value of the residual from equation (1) is used as the estimate of accruals quality rather than the standard deviation of the residual from (1) across the life of the firms. These modifications allows for a firm-year specific measure of current accruals quality.

The Modified Jones Model is specified in equation (2). We estimate this model by fiscal period for each industry. The residuals are calculated using the estimated coefficients from (2) for each firm-year observation in the sample. The absolute value of the residual is used as the second proxy for accruals quality.

$$TACC_{f,t} = \beta_0 + \beta_1 * (\Delta Rev_{f,t} - \Delta Rec_{f,t}) + \beta_2 * PPE_{f,t} + \epsilon_{f,t} \quad (5)$$

Where $TACC_{f,t}$ = total accruals for firm f during fiscal year t = ibc – oancf; $\Delta Rev_{f,t} - \Delta Rec_{f,t}$ = change in revenues minus the change in total receivables = (sale_f,t – sale_f,t-1) – (rect_f,t – rect_f,t-1); $PPE_{f,t}$ = property, plant, and equipment for firm f during fiscal period t.

4.3.1 Estimation and the Mapping of Cash Flows into Accruals Findings

Table 8 presents the results for how estimation affects how well cash flows map into accrual earnings (P2). We include controls for inherent determinates of accruals quality as identified in Dechow et al [2002]. All but one of the determinates of accruals quality load as in the same direction as found in prior studies (Francis et al [2005] and Dechow et al [2002]). The one determine which does not load is the standard deviation of sales which is statistically insignificant. This is likely due to the firm-year modification of the DD model used in this study. However when the specification of the DD model used in Francis et al [2005] is used all determines of accrual quality are statistically significant and load in the same direction as prior studies (see section 5).

Next, we see that the estimation count is positivity associated with both the absolute value of the Dechow and Dichev residual and the absolute value of the residual from the modified Jones model. Both of these findings are consistent with the hypothesis that the degree of estimation in accrual earnings is associated with greater accrual errors and therefore associated with a lower mapping of cash flows into accrual earnings. The fourth and fifth columns of Table 8 present the results when the scaled estimation count is used as the measure of estimation. Once again we see that the coefficients are positively associated with a lower mapping of cash flows into accrual earnings. Overall this provides evidence that more subjective accrual earnings are associated with a lower mapping of accrual earnings 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 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. If accrual earnings are positive this is associated with negative future abnormal returns. On the other hand, negative accrual earnings are associated with positive future abnormal returns. These findings are consistent with his hypothesis that the market over value the persistence of accrual earnings.

We make several small but important modifications to their research design to better suit this study. Since we are interested in the incremental effect of the amount of estimation on the persistence of accrual earnings 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 over values accrual earnings.

Next, we make two small changes to the specification of their measurements 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 firms' 10-K. The information about the estimation of accrual earnings used in this study is found in the firms' 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. 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 accrual earnings 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 accrual earnings, we use the raw accruals amount rather than the decile rank of the amount of accruals.

$$Abnreturns_{f,t} = \beta_0 + \beta_1 * accruals_{f,t} + \beta_2 * est_{f,t} + \beta_3 * est_{f,t} * accruals_{f,t} + \beta_4 * size_{f,t} + \beta_5 * BTM_{f,t} + \beta_6 * ETP_{f,t} + \beta_7 * beta_{f,t} + \epsilon_{f,t} \quad (6)$$

Where $accruals_{f,t}$= value of total accruals equals operating income after cash flows minus operating cash flows = ibc – oancf; $size_{f,t}$ = log of the market value of the firm = log(prcc*csho); $btm_{f,t}$= book to market = at_f_t divided by prcc*csho; $ETP_{f,t}$= earnings to price ratio = oiadp / prcc; $beta_{f,t}$= firms beta for fiscal period t = betav; $Abnreturns_{f,t}$= 1 year abnormal returns beginning 5 days after the filings of the 10-K; $est_{f,t}$ = is the amount of estimation in the footnotes, this is measured as either the estimation count or the estimation count scaled by the number of sentences:

4.4.1 Estimation and Returns Findings

The second and third columns of Table 9 present 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 the estimation count and the total accrual earnings is negative and highly statistically significant. Therefore, more subjective accruals (those with more estimation) are negatively associated with future long term abnormal returns. This is consistent with the hypothesis that investors overvalue firms with positive accrual earnings and undervalue those with negative accruals earnings.

Columns four and five of Table 9 show the results using the scaled estimation count. Similar to the findings when the estimation count is used, we see that the interaction between accrual earnings and the scaled estimation count is negative and statistically significant. This provides further evidence strengthening our hypothesis that that the market reacts as if in the short then they overvalue more subjective positive accrual earnings and vice-versa for negative accrual earnings. [11]

5. Sensitivity Tests

5.1 Alternative Measure of Estimation

We use two alternative proxies for our measure of the amount of accruals estimation. First, we use the number of estimation words found in the footnotes. This is a direct measure of total amount of estimation found in the footnotes. However, this measure does not directly control for the possibility that longer footnotes disclosures are simply more likely to mention more estimation words than shorter ones. Next, we use the percent of estimation words found in the footnote. The percent of estimation words is calculated as the number of estimation words divided by the total number of words in the footnotes multiplied one hundred. This measure controls for some of the shortcomings of the raw estimation word count. Our results are not sensitive to this alternative specification.

5.2 Dummy Estimation Word List

One criticism of our estimation words dictionary is that the words may be "random" and not informative. To address this issue we construct a dummy list of words and rerun our tests of accrual earnings and cash earnings persistence. To construct the dummy list of words, we first sort all words used in 10,000 randomly selected footnotes by their frequency. For each of the

[11] Similar results are found when the decile rank of the accrual earnings is used rather than the raw number.

words in our estimation words dictionary we then pick the word just below it in our sorted frequency list of all words. [12]

Our results suggest that the dummy list of words is, in some cases, marginally informative of greater persistent of accrual earnings. This provides evidence that the estimation words list provides some information about the accruals generating process and isn't simply random. Moreover, the dummy list of words is not informative of lower persistence of cash earnings.

5.3 Alternative Specification of Dechow and Dichev Measure

We used a modified version of the Dechow and Dichev measure to attain a firm year measure how well accrual earnings map into cash flows since we measure the amount of estimation needed during the accrual generating process on a firm year basis. A cross-section measure of accruals quality was never the intent of the original Dechow and Dichev measure and therefore the modified measure may no longer measure accruals quality. We follow Francis et al [2005] and use the standard deviation of the cross-section residual for the past five years to better conform to the original DD measure. We then use the average amount of estimation using a rolling 10 year window (requiring a minimum of 5 observations for each windows) to measure the inherent amount of estimation needed during the accruals generating process for the firm. Our results are not sensitive to this alternative specification.

6. Conclusion

This study directly examines whether more subjective accruals are less predictive of future earnings (i.e. less persistent). Sloan [1996] suggests that the estimation needed when recording accruals reduces the persistence of accrual earnings. Their hypothesis is based on the

[12] The dummy list includes the following words and any of their permutations: understanding, mine, negotiated, ratio, positive, inherent, advance, independence, sufficient, reduced, Arkansas, liquids, certificate, primary

idea that since accrual earnings require a greater degree of estimation they are more likely to be recorded with error (i.e. accrual earnings are less precise). If so, then accrual earnings will be less predictive of future earnings. Richardson et al. [2005] expands upon this hypothesis and suggests that accruals which require more estimation, and thus are more subjective, will be less persistent than accrual earnings which require less estimation. While subsequent research has examined the effects of different components which comprise accruals earnings on accruals persistence none have directly examined the hypothesis of Sloan [1996].

This study is the first to provide direct evidence consistent with the hypothesis that the estimation of accrual earnings reduces their persistence. We find that when accrual earnings have estimation they are less predictive of future earnings. Moreover, we find that the firms' 1 year past, current, and 1 year future cash flows map less into more subjective accrual earnings. We also find some evidence that the markets do not quickly incorporate the estimation of accrual earnings into their valuation of the firm.

In conclusion, the findings in this study provide insight into the accruals generating processing of the firm. More importantly, our findings suggest that understanding the process behind the accrual numbers themselves is important to understanding the persistence of accrual earnings.

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

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

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.

Loughran, T., B. McDonald, 2011, When is a Liability Not a Liability? Textual Analysis, Dictionary, and 10-Ks, *Journal of Finance* 66, 35-65

Marie-Catherine de Marneffe, Bill MacCartney and Christopher D. 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

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

Ogneva, M., 2008, Accrual Quality and Expected Returns: The Importance of Controlling for Cash Flow Shocks, *Working Paper*

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

Rogers, J., A. Van Buskrik, S. Zechman, 2011, Disclosure Tone and Shareholder Litigation, *The Accounting Review*, Forthcoming

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

Tetlock, P., 2007, Giving Content to Investor Sentiment: The Role of Media in the Stock Market, *Journal of Finance* 62,

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

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

Table 1

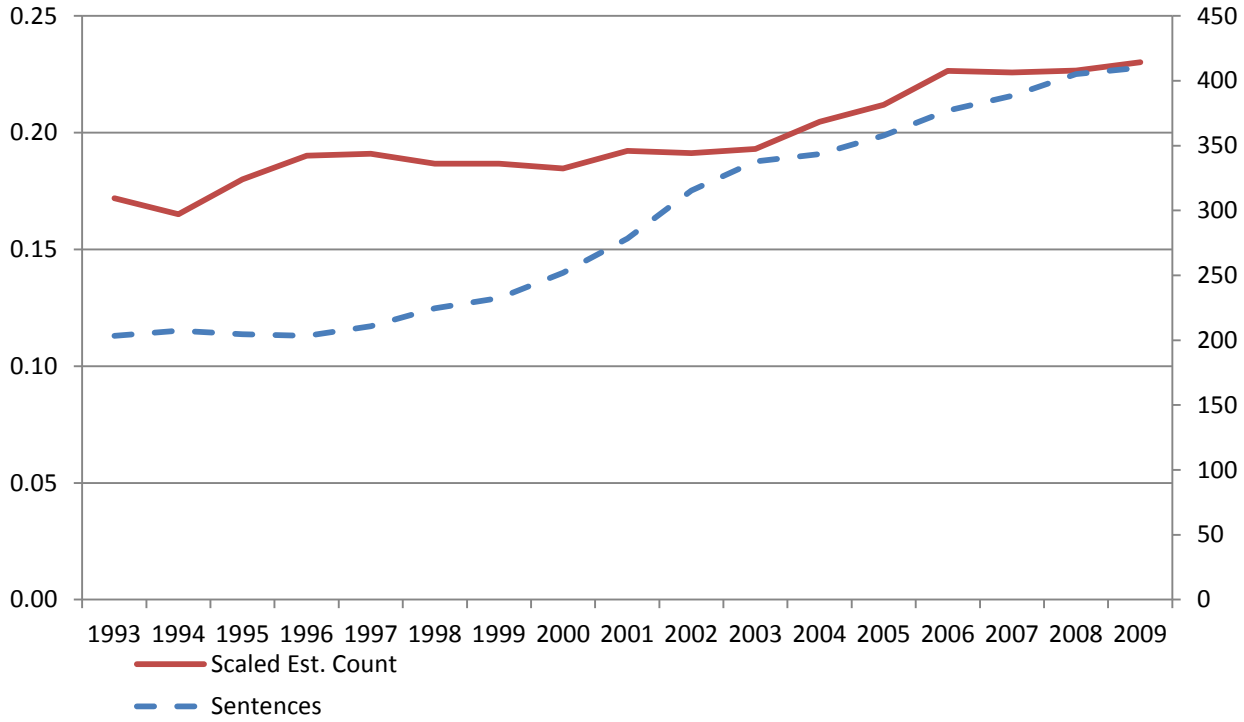
Estimation Trend

Year	Estimation Count	Sentences	Scaled Est. Count	N
1993	34	203	0.17	722
1994	32	207	0.17	820
1995	34	205	0.18	2023
1996	35	203	0.19	3693
1997	37	211	0.19	4185
1998	39	225	0.19	4286
1999	40	233	0.19	4425
2000	43	252	0.18	4490
2001	50	278	0.19	4134
2002	58	315	0.19	3664
2003	63	338	0.19	3294
2004	68	344	0.20	3145
2005	73	358	0.21	3009
2006	82	377	0.23	2879
2007	85	389	0.23	2719
2008	88	405	0.23	2907
2009	91	410	0.23	2786
Average	56	291	0.20	3128

Notes: Table 1 presents the average number estimation count, total number of sentences, and the scaled estimation count found in the notes to the financial statements.

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 2

Estimation by Industry

Industry	Two Digit SIC	Count	Sentences	Scaled	N
Agricultural Production Crops	1	38.68	249.41	0.16	139
Agriculture production livestock and animal specialties	2	44.27	252.97	0.18	30
Agricultural Services	7	46.84	252.35	0.19	49
Forestry	8	28.36	184.14	0.14	14
Fishing, hunting, and trapping	9	22.00	148.00	0.15	2
Metal Mining	10	57.85	321.42	0.18	293
Coal Mining	12	70.48	400.40	0.18	93
Oil And Gas Extraction	13	60.76	321.33	0.20	1900
Mining And Quarrying Of Nonmetallic Minerals, Except Fuels	14	50.54	269.92	0.19	99
Building Construction General Contractors And Operative Builders	15	47.31	282.30	0.18	380
Heavy Construction Other Than Building Construction Contractors	16	64.96	343.43	0.20	178
Construction Special Trade Contractors	17	51.01	244.57	0.20	183
Food And Kindred Products	20	45.14	276.13	0.17	1184
Tobacco Products	21	63.43	414.20	0.16	65
Textile Mill Products	22	42.72	251.16	0.18	317
Apparel And Other Finished Products Made From Fabrics And Similar Materials	23	44.26	277.08	0.17	550
Lumber And Wood Products, Except Furniture	24	49.98	265.71	0.20	260
Furniture And Fixtures	25	42.99	234.79	0.18	353
Paper And Allied Products	26	55.90	297.02	0.19	552
Printing, Publishing, And Allied Industries	27	46.40	271.41	0.18	724
Chemicals And Allied Products	28	51.55	313.81	0.17	4930
Petroleum Refining And Related Industries	29	60.30	349.21	0.18	359
Rubber And Miscellaneous Plastics Products	30	44.87	255.73	0.18	678
Leather And Leather Products	31	37.17	243.13	0.17	251
Stone, Clay, Glass, And Concrete Products	32	48.68	263.48	0.19	310
Primary Metal Industries	33	52.93	288.65	0.19	864
Fabricated Metal Products, Except Machinery And Transportation Equipment	34	47.57	260.57	0.19	868
Industrial And Commercial Machinery And Computer Equipment	35	49.43	267.12	0.19	3635
Electronic And Other Electrical Equipment And Components	36	52.46	279.64	0.19	4446
Transportation Equipment	37	55.16	287.20	0.19	1232
Measuring, Analyzing, And Controlling Instruments	38	46.12	265.49	0.18	3308
Miscellaneous Manufacturing Industries	39	42.81	247.22	0.18	571
Railroad Transportation	40	55.10	292.43	0.20	157
Local And Suburban Transit And Interurban Highway Passenger Transportation	41	54.48	353.00	0.16	23
Motor Freight Transportation And Warehousing	42	44.17	255.19	0.18	409
Water Transportation	44	52.84	314.24	0.17	209
Transportation By Air	45	60.72	315.99	0.19	416

Table 2

Estimation by Industry

Industry	Two Digit SIC	Count	Sentences	Scaled	N
Pipelines, Except Natural Gas	46	58.03	347.01	0.17	79
Transportation Services	47	51.87	311.90	0.18	215
Communications	48	57.05	338.38	0.17	1993
Electric, Gas, And Sanitary Services	49	74.30	410.29	0.19	2389
Wholesale Trade-durable Goods	50	39.91	253.79	0.17	1465
Wholesale Trade-non-durable Goods	51	45.57	287.56	0.17	860
Building Materials, Hardware, Garden Supply, And Mobile Home Dealers	52	37.31	221.67	0.18	123
General Merchandise Stores	53	44.04	251.99	0.17	351
Food Stores	54	43.85	248.69	0.18	378
Automotive Dealers And Gasoline Service Stations	55	53.18	298.26	0.19	269
Apparel And Accessory Stores	56	42.64	266.70	0.17	583
Home Furniture, Furnishings, And Equipment Stores	57	36.56	241.26	0.17	271
Eating And Drinking Places	58	42.81	257.80	0.18	980
Miscellaneous Retail	59	41.55	269.52	0.17	1198
Depository Institutions	60	75.21	376.42	0.21	2363
Non-depository Credit Institutions	61	62.46	343.31	0.19	991
Security And Commodity Brokers, Dealers, Exchanges, And Services	62	44.70	310.68	0.15	787
Insurance Carriers	63	64.55	355.48	0.19	2104
Insurance Agents, Brokers, And Service	64	43.77	285.38	0.16	292
Real Estate	65	47.20	287.22	0.17	534
Holding And Other Investment Offices	67	51.83	318.01	0.17	2445
Hotels, Rooming Houses, Camps, And Other Lodging Places	70	52.11	281.39	0.18	244
Personal Services	72	48.17	296.03	0.17	206
Business Services	73	50.35	292.26	0.18	6775
Automotive Repair, Services, And Parking	75	55.98	328.85	0.17	130
Miscellaneous Repair Services	76	35.03	205.13	0.19	32
Motion Pictures	78	50.85	308.19	0.18	341
Amusement And Recreation Services	79	47.62	307.25	0.16	672
Health Services	80	54.31	304.47	0.19	1130
Legal Services	81	66.17	342.92	0.21	12
Educational Services	82	49.69	281.32	0.18	218
Social Services	83	54.16	323.67	0.19	149
Membership Organizations	86	54.75	331.25	0.17	4
Engineering, Accounting, Research, Management, And Related Services	87	49.66	292.99	0.18	1235
Miscellaneous Services	89	65.50	353.00	0.18	2
Nonclassifiable Establishments	99	56.15	314.76	0.18	197

Notes: Table 2 shows the amount of estimation, number of sentences and the scaled estimation (estimation / sentences) by industry for the period 1994-2010.

Table 3

Summary Statistics

Variable	N	Mean	Minimum	P1	P25	Median	P75	P99	Maximum	Std. Dev.
Total Earnings	52377	-0.06	-1.47	-1.47	-0.08	0.02	0.07	0.30	0.30	0.28
Accrual Earnings	52377	-0.08	-0.87	-0.87	-0.12	-0.06	-0.02	0.29	0.29	0.16
Cash Earnings	52377	0.02	-0.96	-0.96	-0.01	0.07	0.13	0.37	0.37	0.21
Estimation Count	53181	56.00	11.00	11.00	30.00	45.00	72.00	192.00	192.00	36.08
Sentences	53181	291.96	104.00	104.00	166.00	239.00	357.00	1016.00	1016.00	179.13
Scaled Est. Count	53181	0.27	0.10	0.10	0.20	0.25	0.31	0.64	0.64	0.10
Operating Cycle	41066	138.54	9.20	9.20	74.32	115.33	171.09	702.64	702.64	105.09
NITEMS	53181	254.46	143.00	193.00	223.00	250.00	288.00	326.00	362.00	36.00
Special Items	53181	0.25	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.43
log(Market Value)	49506	5.32	0.02	0.02	3.82	5.32	6.80	10.67	10.67	2.18
NEGEARN	53181	0.40	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.49
BTM	49377	0.70	0.07	0.07	0.43	0.68	0.92	1.94	1.94	0.36
|DD Residual|	44557	0.08	0.00	0.00	0.02	0.04	0.10	0.67	0.67	0.11
|Jones Residual|	48771	0.10	0.00	0.00	0.02	0.05	0.11	0.82	0.82	0.13
Beta	43950	0.83	-5.08	-0.41	0.36	0.76	1.21	2.65	5.81	0.65

Notes: Table 3 shows the summary statistics for the sample used in this study. *Total Earnings* is the firms operating income after depreciation scaled by average total assets. *Accrual Earnings* are total accrual earnings scaled by average total assets. *Cash Earnings* are cash earnings scaled by average total assets. *Estimate Count* is the number of estimation related linguistic cues found in the footnotes section of the firm's 10-K. *Total Length* is measured as the total number of words in the footnotes section of the firm's 10-K. *Size* is measured as the log of the market value of the firm's equity. 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. *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	14	15
1	Total Earnings	-	0.37	0.65	-0.10	-0.12	0.01	-0.11	0.06	-0.37	0.41	-0.81	-0.32	-0.21	-0.17	0.04
2	Accrual Earnings	0.58	-	-0.31	-0.10	-0.10	-0.03	0.14	-0.06	-0.31	0.06	-0.36	-0.04	-0.04	0.03	-0.02
3	Cash Earnings	0.78	-0.04	-	0.00	-0.03	0.04	-0.23	0.12	-0.15	0.38	-0.53	-0.20	-0.24	-0.26	0.03
4	Estimation Count	-0.02	-0.08	0.04	-	0.80	0.55	-0.07	0.58	0.09	0.33	0.07	0.02	0.01	-0.04	0.28
5	Sentences	-0.06	-0.07	-0.01	0.74	-	0.01	-0.08	0.51	0.09	0.31	0.09	0.01	0.03	-0.03	0.25
6	Scaled Est. Count	0.04	-0.02	0.07	0.49	-0.11	-	0.00	0.25	0.03	0.12	-0.01	0.02	-0.01	-0.02	0.12
7	Operating Cycle	-0.21	0.02	-0.27	-0.05	-0.04	-0.01	-	0.01	0.06	-0.13	0.12	-0.03	0.13	0.12	0.03
8	NITEMS	0.04	-0.05	0.09	0.57	0.42	0.26	-0.01	-	0.02	0.29	-0.02	-0.08	0.01	-0.03	0.31
9	Special Items	-0.31	-0.34	-0.12	0.09	0.08	0.03	0.06	0.02	-	-0.05	0.38	0.04	0.17	0.15	0.06
10	log(Market Value)	0.33	0.11	0.33	0.35	0.28	0.11	-0.12	0.27	-0.05	-	-0.37	-0.41	-0.23	-0.26	0.39
11	NEGEARN	-0.64	-0.39	-0.50	0.05	0.07	-0.01	0.16	-0.02	0.38	-0.37	-	0.12	0.28	0.23	0.02
12	BTM	-0.01	-0.03	0.00	0.02	0.01	0.01	-0.01	-0.07	0.05	-0.43	0.14	-	-0.09	-0.04	-0.21
13	abs(DD Residual)	-0.47	-0.34	-0.33	0.01	0.03	-0.01	0.13	0.01	0.19	-0.22	0.29	-0.09	-	0.53	0.04
14	abs(Mjones Residual)	-0.42	-0.33	-0.28	-0.02	0.01	-0.02	0.10	-0.02	0.18	-0.24	0.24	-0.05	0.72	-	0.01
15	Beta	-0.01	-0.03	0.01	0.27	0.20	0.11	0.01	0.31	0.06	0.34	0.03	-0.20	0.04	0.02	-

Notes: Table 4 presents the Spearman (above diagonal) and Pearson (below diagonal) correlation for the main variables used in this study.

Table 5

Determinates of Estimation

Dependent Variable	**Estimation Count**		**Scaled Est. Count**	
	(p-value)		*(p-value)*	
Sentences	121.0724	***	-0.1090	***
	(0.000)		(0.000)	
log(Market Value)	2.9012	***	0.0048	***
	(0.000)		(0.000)	
Operating Cycle	-0.0040	***	-0.0000	
	(0.002)		(0.396)	
sdev(Sales)	1.3168	**	0.0047	***
	(0.038)		(0.007)	
sdev(Cash Earnings)	-8.6233	***	-0.0253	***
	(0.000)		(0.000)	
NEGEARNINGS	5.4856	***	0.0087	***
	(0.000)		(0.000)	
NITEMS	0.3012	***	0.0006	***
	(0.000)		(0.000)	
BTM	9.0969	***	0.0175	***
	(0.000)		(0.000)	
BIGFOUR	1.5788	***	0.0035	***
	(0.000)		(0.001)	
Special items	1.6103	***	0.0035	***
	(0.000)		(0.000)	
Constant	-81.3168	***	0.0202	***
	(0.000)		(0.000)	
Observations	32,151		32,151	
R-squared	0.651		0.139	

Notes: Table 5 shows the regression of determines of estimation on the estimation count and scaled estimation count. 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. Sentences is scaled by 1,000. 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 6

Earnings Persistence

Dependent Variable	**Earnings_t+1**		**Earnings_t+1**		**Earnings_t+1**		**Earnings_t+1**	
	(p-value)		*(p-value)*		*(p-value)*		*(p-value)*	
Earnings_t	0.1855	**	0.1402		0.3355	***	0.2982	***
	(0.033)		(0.107)		(0.000)		(0.000)	
Est. Count	-0.0001		-0.0000					
	(0.111)		(0.187)					
Est. Count x Earnings_t	-0.0023	***	-0.0025	***				
	(0.000)		(0.000)					
Scaled Est. Count					0.0261	**	0.0313	***
					(0.039)		(0.011)	
Scaled Est. Count x Earnings_t					-0.4912	***	-0.5256	***
					(0.001)		(0.000)	
Sentences	-0.0339	***	-0.0365	***	-0.0396	***	-0.0409	***
	(0.000)		(0.000)		(0.000)		(0.000)	
Sentences x Earnings_t	0.0435		0.0395		-0.2606	***	-0.2848	***
	(0.537)		(0.576)		(0.000)		(0.000)	
Special Items	0.0383	***	0.0384	***	0.0386	***	0.0388	***
	(0.000)		(0.000)		(0.000)		(0.000)	
log(Market Value)	0.0063	***	0.0067***	***	0.0060	***	0.0064	***
	(0.000)		(0.000)		(0.000)		(0.000)	
NEGEARN	-0.0358	***	-0.0394	***	-0.0361	***	-0.0396	***
	(0.000)		(0.000)		(0.000)		(0.000)	
BTM	-0.0304	***	-0.0172	***	-0.0314	***	-0.0180	***
	(0.000)		(0.000)		(0.000)		(0.000)	
BIGFOUR	0.0061	**	0.0071	**	0.0055*	*	0.0066	**
	(0.041)		(0.015)		(0.064)		(0.024)	
NITEMS	0.0003	***	0.0001	*	0.0003	***	0.0001	
	(0.000)		(0.092)		(0.000)		(0.102)	
sdev(Sales)	-0.0111	**	-0.0080	*	-0.0114	**	-0.0079	*
	(0.020)		(0.075)		(0.017)		(0.077)	
Special items x Earnings_t	0.4970	***	0.5082	***	0.4990	***	0.5106	***
	(0.000)		(0.000)		(0.000)		(0.000)	
log(Market Value) x Earnings_t	0.0382	***	0.0420	***	0.0355	***	0.0392	***
	(0.000)		(0.000)		(0.000)		(0.000)	
NEGEARN x Earnings_t	0.4413	***	0.4649	***	0.4317	***	0.4551	***
	(0.000)		(0.000)		(0.000)		(0.000)	
BTM x Earnings_t	-0.1357	***	-0.1124	***	-0.1457	***	-0.1226	***
	(0.000)		(0.001)		(0.000)		(0.000)	
BIGFOUR x Earnings_t	-0.0962	***	-0.0890	***	-0.0959	***	-0.0884	***
	(0.000)		(0.000)		(0.000)		(0.000)	
NITEMS x Earnings_t	0.0009	***	0.0010	***	0.0007	**	0.0007	**
	(0.002)		(0.001)		(0.022)		(0.011)	
sdv(Sales) x Earnings_t	-0.0447		-0.0584		-0.0464		-0.0600	
	(0.327)		(0.202)		(0.311)		(0.192)	
Constant	-0.0482	**	0.0005		-0.0529	***	-0.0061	
	(0.018)		(0.979)		(0.010)		(0.750)	
Fixed Effects	Industry & Year		Year		Industry & Year		Year	
Cluster	Firm		Firm		Firm		Firm	
Observations	35,909		36,015		35,909		36,015	
R-squared	0.563		0.558		0.562		0.557	

Notes: Table 6 presents the results for our tests of the amount of estimation on earnings persistence. 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. Sentences is scaled by 1,000. 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 7

Panel A: Accruals Persistence (Estimation Count)

Dependent Variable	**Earnings_t+1**		**Earnings_t+1**		**Earnings_t+1**	
	(p-value)		*(p-value)*		*(p-value)*	
Cash Earnings_t	0.8920	***	0.5652	***	0.5974	***
	(0.000)		(0.000)		(0.000)	
Accrual Earnings_t	0.5548	***	0.3052	***	0.3017	***
	(0.000)		(0.005)		(0.005)	
Est. Count	-0.0000		-0.0002	***	-0.0002	**
	(0.737)		(0.008)		(0.013)	
Est. Count x Accrual Earnings_t	-0.0016	***	-0.0015	***	-0.0014	***
	(0.000)		(0.006)		(0.007)	
Est. Count x Cash Earnings_t	0.0004		-0.0005		-0.0005	
	(0.151)		(0.358)		(0.384)	
Sentences			-0.0310	***	-0.0311	***
			(0.006)		(0.005)	
Sentences x Accrual Earnings_t			-0.0738		-0.0831	
			(0.465)		(0.408)	
Sentences x Cash Earnings_t			0.2047	**	0.2113	**
			(0.018)		(0.014)	
Special Items			0.0060		0.0062	
			(0.115)		(0.109)	
log(Market Value)			0.0064	***	0.0060	***
			(0.000)		(0.000)	
NEGEARN			-0.0319	***	-0.0282	***
			(0.000)		(0.000)	
BTM			-0.0038		-0.0145	***
			(0.483)		(0.009)	
BIGFOUR			0.0039		0.0035	
			(0.299)		(0.345)	
NITEMS			0.0001	**	0.0002	***
			(0.024)		(0.000)	
sdev(Sales)			0.0044		0.0024	
			(0.440)		(0.683)	
Constant	-0.0341	***	-0.0489	**	-0.0725	***
	(0.000)		(0.017)		(0.001)	
Controls Interact w/ Cash earnings	NA		Yes		Yes	
Controls Interact w/ Accrual earnings	NA		Yes		Yes	
Fixed Effects	Industry & Year		Year		Industry & Year	
Cluster	Firm		Firm		Firm	
Observations	47,656		36,587		36,479	
R-squared	0.594		0.593		0.596	

Panel B: Accruals Persistence (Scaled Estimation Count)

Dependent Variable	**Earnings_t+1**		**Earnings_t+1**		**Earnings_t+1**	
	(p-value)		*(p-value)*		*(p-value)*	
Cash Earnings_t	0.9162	***	0.5952	***	0.6265	***
	(0.000)		(0.000)		(0.000)	
Accrual Earnings_t	0.5229	***	0.4046	***	0.3978	***
	(0.000)		(0.000)		(0.000)	
Scaled Est. Count	0.0456	***	-0.0242		-0.0221	
	(0.005)		(0.199)		(0.243)	
Scaled Est. Count x Accrual Earnings_t	-0.2769	*	-0.431	**	-0.4299	**
	(0.075)		(0.016)		(0.016)	
Scaled Est. Count x Cash Earnings_t	-0.0319		0.0787		0.0942	
	(0.799)		(0.643)		(0.582)	
Sentences			-0.053	***	-0.0513	***
			(0.000)		(0.000)	
Sentences x Accrual Earnings_t			-0.2895	***	-0.2929	***
			(0.001)		(0.001)	
Sentences x Cash Earnings_t			0.1632	**	0.1763	**
			(0.040)		(0.027)	
Special Items			0.0059		0.0060	
			(0.122)		(0.116)	
log(Market Value)			0.006	***	0.0057	***
			(0.000)		(0.000)	
NEGEARN			-0.0324	***	-0.0287	***
			(0.000)		(0.000)	
BTM			-0.0045		-0.0153	***
			(0.402)		(0.006)	
BIGFOUR			0.0036		0.0032	
			(0.336)		(0.390)	
NITEMS			0.0001	**	0.0002	***
			(0.024)		(0.000)	
sdev(Sales)			0.0043		0.0021	
			(0.447)		(0.716)	
Constant	-0.0392	***	-0.0457	**	-0.0685	***
	(0.000)		(0.024)		(0.001)	
Controls Interact w/ Cash earnings	NA		Yes		Yes	
Controls Interact w/ Accrual earnings	NA		Yes		Yes	
Fixed Effects	Industry & Year		Year		Industry & Year	
Cluster	Firm		Firm		Firm	
Observations	47,656		36,587		36,479	
R-squared	0.594		0.593		0.596	

Notes: Table 7 presents the results for our tests of the amount of estimation on the persistence of cash and accrual earnings. 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. Sentences is scaled by 1,000. 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 8

Accruals Quality

Dependent Variable	**abs(DD Residual)**		**abs(Mjones Residual)**		**abs(DD Residual)**		**abs(Mjones Residual)**	
	(p-value)		*(p-value)*		*(p-value)*		*(p-value)*	
Est. Count	0.0001	***	0.0001	***				
	(0.000)		(0.000)					
Scaled Est. Count					0.0278	***	0.0252	***
					(0.001)		(0.009)	
Sentences	0.0195	***	0.0291	***	0.0379	***	0.0459	***
	(0.000)		(0.000)		(0.000)		(0.000)	
log(Market Value)	-0.0084	***	-0.0122	***	-0.0081	***	-0.0119	***
	(0.000)		(0.000)		(0.000)		(0.000)	
sum(NEGEARN)	0.0032	***	0.0024	***	0.0033	***	0.0025	***
	(0.000)		(0.000)		(0.000)		(0.000)	
sdev(Sales)	-0.0025		0.0127	***	-0.0024		0.0127	***
	(0.331)		(0.000)		(0.345)		(0.000)	
sdev(Cash Earnings)	0.2370	***	0.2664	***	0.2358	***	0.2653	***
	(0.000)		(0.000)		(0.000)		(0.000)	
avg(Operating Cycle)	0.0000	***	0.0000	**	0.0000	***	0.0000	**
	(0.000)		(0.015)		(0.000)		(0.017)	
Special Items	0.0318	***	0.0375	***	0.0319	***	0.0376	***
	(0.000)		(0.000)		(0.000)		(0.000)	
NEGEARN	0.0122	***	0.0041	**	0.0124	***	0.0043	**
	(0.000)		(0.021)		(0.000)		(0.015)	
NITEMS	-0.0000		-0.0001	***	0.0000		-0.0000	**
	(0.716)		(0.005)		(0.439)		(0.049)	
BTM	-0.0450	***	-0.0473	***	-0.0442	***	-0.0466	***
	(0.000)		(0.000)		(0.000)		(0.000)	
BIGFOUR	-0.0077	***	-0.0173	***	-0.0075	***	-0.0172	***
	(0.000)		(0.000)		(0.000)		(0.000)	
Constant	0.1064	***	0.1591	***	0.0949	***	0.1488	***
	(0.000)		(0.000)		(0.000)		(0.000)	
Observations	28,621		30,893		28,621		30,893	
R-squared	0.215		0.203		0.214		0.203	

Notes: Table 8 shows the association between accrual quality, as measure by the Dechow and Dichev 2001 model and the Modified Jones model. 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. Sentences is scaled by 1,000. 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 9

Panel A: Abnormal Returns (Estimation Count)

Dependent Variable	**Abnormal Returns_t+1**		**Abnormal Returns_t+1**		**Abnormal Returns_t+1**		**Abnormal Returns_t+1**	
	(p-value)		*(p-value)*		*(p-value)*		*(p-value)*	
Accrual Earnings_t	0.0202		-0.1373		0.0091		-0.1568	*
	(0.794)		(0.107)		(0.905)		(0.061)	
Est. Count	0.0005	***	0.0000		0.0001		-0.0002	
	(0.003)		(0.975)		(0.672)		(0.356)	
Est. Count x Accrual Earnings_t	-0.0026	**	-0.0078	***	-0.0014		-0.0067	***
	(0.041)		(0.000)		(0.284)		(0.000)	
Sentences			0.1447	***			0.088	*
			(0.002)				(0.057)	
Sentences x Accrual Earnings_t			1.5043	***			1.5561	***
			(0.000)				(0.000)	
log(Market Value)	-0.0194	***	-0.0196	***	-0.0172	***	-0.0172	***
	(0.000)		(0.000)		(0.000)		(0.000)	
BTM	0.2634	***	0.2608	***	0.1737	***	0.171	***
	(0.000)		(0.000)		(0.000)		(0.000)	
Beta	0.0389	***	0.039	***	0.0337	***	0.0337	***
	(0.000)		(0.000)		(0.000)		(0.000)	
ETP	-0.0819	***	-0.0833	***	-0.0699	***	-0.0718	***
	(0.000)		(0.000)		(0.000)		(0.000)	
Constant	-0.0625	***	-0.0728	***	-0.0263		-0.0331	
	(0.008)		(0.002)		(0.423)		(0.315)	
Fixed Effects	None		None		Year		Year	
Observations	42,899		42,899		42,899		42,899	
R-squared	0.022		0.022		0.061		0.062	

Table 9

Panel B: Abnormal Returns (Scaled Estimation Count)

Dependent Variable	**Abnormal Returns_t+1**		**Abnormal Returns_t+1**		**Abnormal Returns_t+1**		**Abnormal Returns_t+1**	
	(p-value)		*(p-value)*		*(p-value)*		*(p-value)*	
Accrual Earnings_t	0.286	**	0.1845		0.3176	***	0.1208	**
	(0.014)		(0.213)		(0.005)		(0.407)	
Scaled Est. Count	-0.1302		-0.0653		-0.147	*	-0.1302	
	(0.121)		(0.444)		(0.082)		(0.140)	
Scaled Est. Count x Accrual Earnings_t	-2.1881	***	-2.0807	***	-2.0037	***	-1.7833	***
	(0.000)		(0.000)		(0.000)		(0.002)	
Sentences			0.1427	***			0.0415	
			(0.000)				(0.266)	
Sentences x Accrual Earnings_t			0.3247				0.5477**	
			(0.205)				(0.029)	
log(Market Value)	-0.0143	***	-0.0181	***	-0.0161	***	-0.0163	***
	(0.000)		(0.000)		(0.000)		(0.000)	
BTM	0.2772	***	0.2657	***	0.1756	***	0.1739	***
	(0.000)		(0.000)		(0.000)		(0.000)	
Beta	0.0451	***	0.0415	***	0.034	***	0.0345	***
	(0.000)		(0.000)		(0.000)		(0.000)	
ETP	-0.0863	***	-0.0849	***	-0.0712	***	-0.073	***
	(0.000)		(0.000)		(0.000)		(0.000)	
Constant	-0.0519	*	-0.0723		0.0036		-0.0074	**
	(0.054)		(0.009)		(0.919)		(0.845)	
Fixed Effects	None		None		Year		Year	
Observations	42,899		42,899		42,899		42,899	
R-squared	0.021		0.022		0.062		0.062	

Notes: Table 9 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. Sentences is scaled by 1,000. 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.

Appendix 1

Statistical parsing is a technique pioneered in the field of Natural Language Processing used to diagram the structure of a sentence. For this study, 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. Essentially, this implementation finds the most likely map of the sentence by matching the sentence to a tree bank of manually parsed sentence to find the layout of the sentence which is most likely.

To illustrate, 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. Using the sentences map removes any ambiguity that the sentence is conveying that something was estimated.

Appendix 2

The following linguistic cues which contain an estimation word are counted: 1. Direct Object – the accusative object of a verb (i.e. "estimate receivables" or "used estimates") - the verb or object can be an estimation word 2. (Passive) Nominal Subject – similar to a direct object, in that it relates information about an object; the verb or object may be an estimation word 3. Adjective Modifier – Modifies the meaning of a noun (i.e. "likely receivable" or "anticipated value") 4. Noun Compound Subjects – A noun used to modify another noun (i.e.

“estimates value”) 5. Quantifier Phrase Modifier – Modifier to a number (i.e. “approximately $100”). A Perl script is used to count the number of times a linguistic cue conveys that estimation was used.