Estimating the Amount of Estimation in Accruals

Jason V. Chen
University of Illinois at Chicago
jchen19@uic.edu

Feng Li
Shanghai Advanced Institute of Finance
Shanghai Jiao Tong University
fli@saif.sjtu.edu.cn

February 24, 2016

Abstract

This paper examines the relation between the amount of estimation needed during the accrual generating process and the quality of accruals. Specifically, we identify estimation related linguistic cues contained in the notes and critical accounting policies sections of 10-K filings to measure the amount of estimation needed during the accruals generating process. Consistent with the conjecture in Sloan (1996), we find that accruals requiring more estimation have a lower relation to future earnings. We also find that these accruals are less likely to map into past, current, or future cash flows in the spirit of Dechow and Dichev (2002). Furthermore, we find that our results are driven by specific accrual accounts as well as within account variation in estimation. Lastly, we find no relation between the amount of estimation in accruals and the magnitude of the relation between accruals and future abnormal returns. Overall, our findings suggest that the amount of estimation needed during the accrual generating process plays an important role in the persistence of accruals.

We thank Matt DeAngelis, Ilia Dichev, Scott Richardson, Lakshmanan Shivakumar, Irem Tuna, and the workshop participants at University of Buffalo, Columbia University, Florida Atlantic University, Florida State University, London Business School, London School of Economics, University of Michigan, Michigan State University, Purdue University, Rice University, University of Illinois at Chicago, and Temple University for their comments. We gratefully acknowledge the support of the Ross School of Business, the financial support from the Paton Accounting Fellowship, the Harry Jones Endowment, and the China Academy of Financial Research, Shanghai Jiao Tong University.

1. Introduction

The accounting literature has long observed that accruals are less persistent than cash flows. One reason proposed for this lower persistence is the degree of estimation required during the accrual generating process (Sloan 1996; Richardson et al. 2005). In our study, we examine this conjecture by calculating the amount of estimation embedded in accruals, measured as the number of estimation-related linguistic cues in the notes and critical accounting policy (CAP) disclosures in the management discussion and analysis section of firm 10-K filings. We then use the amount of estimation as a measure of accruals quality to empirically investigate the association between accruals quality and the persistence of accruals.

Unlike cash flows, accruals incorporate estimates of future cash flows, cash flow deferrals, depreciation and amortization, and fair value estimates. However, prior research on accruals persistence has focused on the role of accrual components rather than an examination of how estimates impact persistence (Dechow et al. 2010). For example, Dechow and Ge (2006) find that accruals exhibit lower persistence when they contain special items. In another study, Richardson et al. (2005) find that accruals persistence is impacted by a firm's working capital, non-current operating, and financial asset accruals. Our study complements this body of research by examining how the characteristics of the accruals generating process impact accruals persistence.

To do so, we focus on the notes and CAP disclosures in the management discussion and analysis section of firm 10-K filings. This focus serves a dual purpose. First, these sections contain detail on the nature and generation of accruals for us to gauge the amount of estimation in accruals. Second, it allows us to examine the usefulness of these sections in assessing accruals quality, an assertion which has been questioned by both auditors and investors (Radin 2007).

To conduct our study, we construct a dictionary of estimation-related words which we use to search for linguistic cues in the notes and CAP disclosure sections (see Appendix 2 for details). These cues fall into one of three linguistic relations which we identify from a preliminary reading of a large sample of financial statements: (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"), and (3) an

estimation word is an adjective to an object (e.g. "estimated costs"). We use statistical parsing techniques to automate the search for these linguistic cues in the notes to the financial statements and CAP section of firms' 10-K filings (see Appendix 1).

Our first hypothesis is that accruals that require more estimation have a lower association with future earnings. Consistent with this, we find that accruals are significantly less persistent when a firm's footnotes and CAP disclosures indicate a greater amount of accrual estimation. Moreover, we find no significant relation between the amount of estimation and the persistence of cash flows. Together, these findings support Sloan's conjecture that accruals persistence is at least partially explained by the amount of estimation.

Our findings also show that a greater amount of accrual estimation is associated with lower accruals quality, as measured by the relation between accruals and a firm's past, current, or future cash flows in the sense of Dechow and Dichev (2002). This is consistent with the hypothesis that a greater amount of estimation leads to lower accruals quality and provides further evidence that the estimation needed during the accrual generating process leads to lower persistence.

To gain further insight into our findings, we decompose the number of estimation cues in the notes and CAP disclosures into two components: those due to specific accruals accounts and those due to within-account variation.[1] We do so by identifying 49 common accounting items (see footnote 5) and estimate an expected number of linguistic estimation cues for each account. This expected number of estimation cues captures the number of cues due to different accruals accounts while the residual captures within-accounts variations. We find that both the persistence and quality of accruals are driven by a combination of specific accruals accounts and within-account variations.

Lastly, we find no relation between the amount of estimation in accruals and the accrual anomaly (Sloan 1996; Xie 2001; Mashruwala et al. 2006).[2] This result is potentially due to the accrual anomaly

[1] We thank Irem Tuna for this observation.
[2] This relation is found for cross-sectional OLS regressions of accruals on future abnormal returns. Our study uses one-year abnormal returns which begin 5 days after the 10-K filing.

being weaker in more recent years (Green et al. 2011). Robustness tests of these findings using a Carhart four-factor Alpha model yield similar results.

To check the robustness of our empirical results, we conduct a number of additional tests. First, we build pseudo word counts by randomly selecting a dictionary of words with similar frequencies as our estimation cue words in the 10-K samples. Our bootstrapping test shows a low probability of obtaining our results using a pseudo random dictionary. Second, we re-run our analyses including the non-CAP section of the management discussion and analysis section. The results from this analysis show no relation between the amount of estimation in this section and accruals or earnings quality. This suggests that our measure of accruals estimation does not simply capture generic business uncertainty. Lastly, we control for numerous other textual disclosure characteristics of the financial statements documented in prior studies, including the Fog index (Li 2008) as well as a textual measure of competition (Li et al. 2012); our results are robust to these controls.

Our study makes several contributions to the literature on accruals. First, this study proposes a new measure of accrual quality based upon the qualitative information in a firms notes to the financial statements and CAP disclosures. This complements the findings in recent studies which suggest that qualitative disclosures can provide relevant information about a firms accruals (Frankel et al. 2015).[3] Prior studies have mostly overlooked this important source of information about accruals when examining accruals quality.

Second, our findings strengthen Sloan's argument that the amount of estimation involved in generating accruals explains the difference in the persistence of the cash portion of earnings and the accruals portion of earnings. Some studies argue that this difference is driven by omitted fundamental differences such as growth. By contrast, our findings lend support to the explanation that accruals persistence is impacted by the amount of accruals estimation.

[3] Frankel et al. 2015 uses support-vector regressions on the qualitative information in the MD&A to predict accruals. They find that these accruals exhibit lower persistence.

Finally, our study contributes to the textual analysis accounting literature by using grammatical relations to extract meaning from qualitative financial information. These relations 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 the combination of qualitative and quantitative accounting information provides 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 the motivation for our hypotheses. In Section 3, we discuss our sample of 10-K footnotes and CAP disclosures and financial information. In Section 4, we present our research design and main results. Section 5 presents the results of our robustness checks and Section 6 concludes the study.

2. Prior Literature and Hypotheses

One area of research our study contributes to is the literature on accrual quality. In an early study in this area, Sloan (1996) finds that the accruals portion of earnings has a lower association with future earnings compared to the cash portion earnings. He attributes this lower persistence to the greater amount of estimation needed to incorporate estimates of future cash flows, depreciation and allocations, deferrals, and valuations in deriving accruals. Richardson et al. (2005) formally models this accruals estimation process as an error-in-variables problem. In Richardson's model, recorded accruals are assumed to be measured with error since managerial estimation is needed during the accrual generating process. This error subsequently reduces the relation between a firm's accruals and its future earnings.

Following Richardson's model, a number of studies have examined how well specific accruals components predict future earnings. For instance, Dechow and Ge (2006) find that accruals are less persistent when a firms earnings contain special items. In another study, Richardson et al. (2005) find that financing accruals, working capital accruals, and non-current operating accruals exhibit different degrees of estimation. Specifically, financial accruals are typically contractually defined and thus require less estimation than working capital or non-current operating accruals. Conversely, the greater estimation

required to determine working capital and non-current operating accruals implies that these accruals are less likely to be realized as cash flows and therefore will be less informative of future earnings. Accordingly, they find that financing accruals are more persistent than either working capital or non-current operating accruals.

While on average accruals require more estimation than cash flows, all accruals are likely not created equal. Specifically, differences in business fundamentals, accounting policies, and earnings management incentives may lead to differences in accrual estimates across firms, *ceteris paribus*. This means that even if two companies have the same total dollar amount of accruals, the amount of estimation needed in their respective accruals may be vastly different.

Focusing on the implications of these cross-sectional differences in the amount of estimation in accruals on accruals persistence, we can now state our first hypothesis:

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

In another study, Dechow and Dichev (2002) find lower earnings persistence for firms which exhibit a lower mapping of accruals into past, current, and future cash flows. They posit that the lower level of mapping implies greater estimation errors. Consistent with this hypothesis, we argue that greater accruals estimation means that accruals are likely to be recorded with lower precision (i.e., accruals will map less into realized cash flows). This leads to our second hypothesis:

Prediction 2: *Accruals that involve more estimation are less strongly associated with a firm's past, current, or future cash flows.*

In another stream of research, a number of studies have found that the lower persistence of accruals is not quickly incorporated by investors into their valuations of firms (Sloan 1996; Hanlon 2005; Richardson et al. 2005). One explanation for this finding is that a focus on total earnings leads investors

to disregard how the predictive ability of current earnings is affected by the persistence of accruals (Sloan 1996; Kraft et al. 2006). Accordingly, Sloan (1996) finds that future abnormal returns are negatively associated with the magnitude of a firm's accruals.

If investors ignore the difference between accruals and cash flows, then they may not fully incorporate this information in a timely manner. In this case, a greater amount of estimation would exacerbate the accrual anomaly as accruals that need more estimation would be more likely to be mispriced and negatively associated with future stock returns.

On the other hand, since accrual estimation information is readily available in a firm's 10-K filings, it is possible that investors can easily incorporate this information quickly. Indeed, prior research suggests that investors and analysts utilize information provided in a firm's footnote disclosures (De Franco et al. 2011). If investors are aware of the lower persistence of earnings upon the filing of the 10-K, then this information may not be informative about the relation between accruals and future long term abnormal returns of the firm. This leads to our final prediction, stated in the null hypothesis format:

Prediction 3: *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.*

3. Data Preparation

3.1 Extracting the Footnotes to the Financial Statements

Our data preparation consists of two components. We begin with a sample of all 10-K documents filed with the SEC between 1993 and 2012 from the SEC EDGAR website.[4] From this sample, we remove all filings not explicitly labeled as either "10-K" or "10-K405." Using Perl, we then extract the notes to the financial statements and the critical accounting policies (hereafter CAP) disclosure in the management discussion and analysis section (hereafter MD&A) of each of the 10-K filings. To aid in our analysis, we strip the footnotes and CAP disclosures of all HTML tags and tables. To mitigate issues related to

[4] Mandatory filing through the EDGAR online system for all publically traded companies within the U.S. was enacted around 1995.

identifying each section, we eliminate notes with fewer than 1,000 words and CAP disclosures with fewer than 400 words.

3.2 Measuring Accruals Estimation

After obtaining our sample, we next use textual analysis of the footnotes and CAP disclosures to capture the amount of accruals estimation. These sections provide a wealth of information not found in other sections of the 10-K filing, including information about a firm's accounting process and the assumptions made during the accrual generating process (Merkeley 2011; Riedl and Srinivasan 2010).

In particular, all 10-K documents filed after May 2002 are required to provide CAP disclosures that include a discussion of any accounting policies that entail highly uncertain assumptions for which differing estimates would have a material influence on the firm's financial statements (Billings 2011). Research on the impact of this 2002 SEC ruling has found an impact of the new disclosure requirements on investor valuation decisions (Levine and Smith 2011). Since our sample begins in 1993, we have CAP disclosures available for only the post-2002 portion of our sample.

The basis of our textual analysis is the identification of three distinct linguistic cues that signal estimation in the text of the 10-K footnotes and CAP disclosures. These linguistic cues are derived from a study of numerous notes and CAP disclosures. The first linguistic cue is the indication that 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 receivables were estimated. The second linguistic cue is the indication that 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." Our third linguistic cue is the indication that 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 modified by the adjective "estimated" thereby conveying that the costs are estimates.

We use the Stanford open source statistical parser to parse each sentence into components including noun modifiers, direct object modifiers, adjective modifiers, and others (Marneffe et al. 2006).

We then use these grammatical relations between words to identify the linguistic cues that convey that an estimate was used by management in the accruals generating process (see Appendix 1).

After parsing each sentence to determine the grammatical relations between words, we create fours dictionaries to help us extract meaning from the parsed sentences. 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 contains Estimation Objects (Nouns). This dictionary includes estimation related objects and contains words such as "Belief," "Estimates," and "Approximations." We use this dictionary in conjunction with our third dictionary, a Use Words dictionary. This dictionary includes words that denote that management uses or needs some object and includes words such as "Make," "Use," and "Include." Our fourth dictionary contains Estimation Adjectives and includes estimation words that are used to modify some object such as "Likely," "Estimated," and "Anticipated".

With our parsed sample and word dictionaries, we are able to examine each sentence in our sample of firms' footnotes and CAP disclosures for linguistic cues that convey that an estimate was used by management (see Appendix 2).[5] The tally of these estimation-related linguistic cues is used as our measure of the amount of estimation needed by management of a firm during their accrual generating process.

3.2.1 Measuring Between-Accounts Estimations and Within-Account Estimations

As mentioned, we distinguish between two estimation components in our study: between-accounts estimations (BAE) and within-account estimations (WAE). A firm's notes and CAP sections could have a relatively greater amount of estimation cues because the firm has specific accrual accounts. For instance, a defined benefits pension plan accrual tends to involve more estimated components than other accruals. We refer to this amount of estimation as BAE. It is also possible for there to be different levels of

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

estimations within the same accrual account or transaction across companies. For example, the calculation of uncollectible receivables may require more parameters to be estimated for one company than another. We refer to this estimation as WAE. In our study, we distinguish between these types of estimation by examining whether BAE and WAE have different implications for accruals persistence and quality. We posit that both BAE and WAE partially explain the lower persistence of accruals in comparison to that of cash flows.

To measure BAE and WAE, we first extract all footnote headers from our sample of 10-K filings and sort them based on their frequency. Using a pool of the most frequent footnote headers, we next manually categorize approximately 1,000 unique footnote headers into 49 distinct footnote items.[6][7] We then search for these 49 items in the notes to the financial statements for each company in a given year. We use the following regression to calculate the expected number of estimation-related linguistic cues given the items identified in the notes to the financial statements:

$$Estimation_{f,t} = \beta_0 + \Sigma\beta_j Footnote\_item_{j,f,t} + \epsilon \quad (1)$$

where $Estimation_{f,t}$ is the number of estimation-related linguistic cues in firm f's notes to the financial statement and CAP disclosures in year t and $Footnote\_item_{j,f,t}$ is an indicator which equals 1 if the company's notes to the financial statements contain the specific footnote item j and 0 otherwise The predicted (residual) value from this model captures BAE (WAE), or the number of estimation cues explained by the existence of different accruals accounts (within-account variations). Appendix 3 presents the top 10 and bottom 10 account items in terms of the amount of estimation as indicated by the β_j

[6] The footnote account 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.

[7] Our set of 1,000 unique manually categorized footnote headers accounts for approximately 70% of all footnote headers in our sample.

coefficients estimated in model (1). The items associated with the greatest number of estimation-related linguistic cues include the following: the recording of Fair Value, Regulatory, Intangibles, Derivatives, Reinsurance, Restatement, Restructuring, Stock Compensation, Contingencies, and Segments. By contrast, the items associated with the lowest number of estimation-related linguistic cues include the following: Other Expenses, Long-Term Debt, Credit Arrangements, Significant Customers, Leases, Inventory, Accounting Policies, Stock Options, Taxes, and Home Loans.

3.3 Sample Preparation

To obtain our final sample for our analysis of the effect of the amount of accruals estimation on accruals persistence, we merge our firm sample from our textual analysis with firm financial information obtained from Compustat and equity market information from the Center for Research in Security Prices database (CRSP).[8] Any firm years with data missing from either Compustat or CRSP are eliminated, as are financial institutions, as these firms may have idiosyncratic accruals and disclosures.[9] This process gives us a final sample of 60,389 firm year observations. We calculate future one-year abnormal returns as the buy-and-hold returns of the firm minus the buy-and-hold returns of a value weighted market portfolio over the same window. We calculate one-year abnormal buy-and-hold returns beginning five days after the filing of the 10-K.

3.3.1 Summary Statistics

Table 1 presents the average estimation-related linguistic cues and the average total number of words found in the notes to the financial statements and CAP disclosures for the firms in our sample. Consistent with prior studies, we find that the average length of the footnotes and CAP disclosures has steadily increased over time (Li 2008; Radin 2009). On average, the length of the notes to the financial statements and CAP disclosures has tripled in size over our sample period, from an average of 3,725 words in fiscal

[8]We were unable to find corresponding financial data or market information for a handful of firms. The main reason for many of these cases stemmed from not being able to find an appropriate GVKEY for the CIK specified in the header of the 10-K filing.
[9] We identify financial firms as those firms having SIC codes between 6000 and 6999.

period 1993 to 11,971 words in 2012.[10] We also find that the number of estimation-related linguistic cues in the footnotes and CAP disclosures increased monotonically during our sample period, from an average of 36 in the fiscal period 1993 to an average of 155 in 2012. However, as shown in Figure 1, we see that while the number of estimation-related linguistic cues increased during the sample period, there is a slight leveling off in the growth of the number of footnotes in the later periods of our sample. To control for the time-trend in section length, we include year fixed effects and the total number of words in our empirical analysis.

In Table 2, we present the average number of estimation-related linguistic cues and the average number of words in the notes to the financial statements and CAP disclosures by industry. These statistics indicate that industry appears to play an important role in the amount of estimation required. Specifically, Coal Mining, Electric Services, Metal Mining, Communications, and Non-classifiable Establishments have the greatest amount of accruals estimation while Agricultural Production, Automotive Repair, Building Materials, Construction Contractors, and Social Services have the least.

Next, examine how our measure of estimation varies from year to year to gain a better understanding of the variation in our measure. Table X presents the yearly quintile ranking of estimation count by the following years quintile rank. We find that approximately 56% of our firm year observations remain in the same quintile ranking of estimation in the following year across our sample. This suggests that the average firm does not drastically change its accounting policy from year to year. Next, we see that firms which do not remain in the same quintile ranking of estimation are ranked in an adjacent quintile in approximately 42% of our observations. For instance, 46% of firm year observations in quintile rank 2 are ranked in either estimation quintile 1 or 3 in the following year. This percentage is drastically reduced for non-adjacent quintile estimation ranks.

4. Research Design and Results

4.1 Determinants of Estimation

[10] The length of the notes and CAP sections are calculated as the total number of non-stop words in each section.

We begin our analysis by exploring the relation between the number of estimation-related linguistic cues and several determinants of accruals estimation identified in prior research, namely firm size, negative earnings, operating cycle, and volatility (Dechow and Dichev 2002; Francis et al. 2005). This examination serves two purposes. First, it provides some intuitive validation to our measure. Second, it indicates whether these determinants should be included as control variables in our subsequent analyses.

Size - Larger firms typically have more operational complexity than smaller firms. On the one hand, this suggests that greater estimation is needed to convey the activities of the firm through accruals. However, the transactions of these firms may have diversification effects, which may make the estimation of accruals more precise. For instance, firms with a diverse set of receivables may be able to estimate their bad debt ratio more precisely if their different sources of receivables offer some diversification. Therefore, *ex ante*, we do not have a clear prediction on the association between firm size and the amount of estimation in the accruals generation process.

Negative Earnings - Accounting conservatism suggests that investors require more verification of good news than bad news (Basu 1997). If so, then we would expect that investors would require greater precision for positive than negative earnings. Therefore, we hypothesize that firms with negative earnings are likely to have a greater amount of estimation in their accruals.

Operating Cycle - Longer operating cycles imply a longer horizon for accruals to be realized as cash flows. This longer horizon suggests a greater amount of estimation to calculate and recognize accruals. Therefore, we predict a positive association between the operating cycle of a firm and the amount of estimation in its accruals.

Volatility of the Business - Managers who operate in more volatile business environments are less likely to know the future of their business. This uncertainty is likely to be reflected in their calculation of accruals. In this study, we examine two potential proxies for the volatility of a firm's business environment: the standard deviation of cash flows and the standard deviation of sales (both scaled by the book value of assets). We include both proxies as sales could be affected by the amount of accruals booked (e.g., receivables).

Including the above firm characteristics, we examine the relation between the number of estimation-related linguistic cues in the notes and CAP disclosures using the following Tobit model, left censored at 0:[11]

$$Estimation_{f,t} = \beta_0 + \beta_1 \text{Size}_{f,t} + \beta_2 OperatingCycle_{f,t} \quad (2)$$
$$+\beta_3 \sigma(Sales)_{f,t} + \beta_4 \sigma(Operating\ Cash\ Flows)_{f,t} + \beta_5 Negearn_{f,t} + \epsilon$$

where $Estimation_{f,t}$ is as defined earlier; $\text{S}\,ize_{f,t}$ is the log of the market value of the firm's equity; $OperatingCycle_{f,t}$ is the length of the operating cycle of the firm calculated as $\log(\frac{invt}{cogs} * 360 + \frac{rect}{sales} * 360)$, where *invt* is the average inventory balance and *cogs* is the cost of goods sold; $\sigma(Sales)_{f,t}$ is the standard deviation of the firm's sales (scaled by the book value of assets) over the past five years; $\sigma(Operating\ Cash\ Flows)_{f,t}$ is the standard deviation of the firm's cash flows (scaled by the book value of assets) over the past five years; and $Negearn_{f,t}$ is the number of years that the firm had negative earnings over the past five years.

4.1.1 Determinants of Accruals Estimation

Table 5 presents the results for the relation between the determinants of accruals estimation and our measure of the amount of accruals estimation. Consistent with our predictions, the standard deviation of sales, operating cycle, and negative earnings are all positively associated with *Estimation*. In addition, the positive and significant coefficient on *size* suggests that larger firms tend to report accruals that need more estimation.

Continuing with the results in Table 5, we see that the standard deviation of cash flows is statistically significant but that its coefficient loads in the opposite direction of our prediction. One

[11] Untabulated findings using Ordinary Least Squares (OLS) estimation provide similar results.

possible explanation for this finding is that managers facing an uncertain environment may not book their accruals, as they are uncertain of their future cash flows.

4.2 Accruals Estimation and Persistence

To test our first prediction (P1), we examine how the amount of estimation in accruals is related to both the persistence of earnings and the persistence of accruals relative to cash flows. Following prior literature, we regress the following year's earnings on the current year's earnings (accruals and cash flows) to estimate the relation between current and future earnings (Sloan 1996; Li 2008). If the estimated coefficient on current earnings is high, then we would conclude that current earnings are highly persistent.

To measure the impact of estimation on the persistence of earnings, we include the interaction between the current year's earnings and the amount of estimation in the following model:

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

where $Estimation_{f,t}$ is as defined earlier and $Earnings$ is income before extraordinary items scaled by average total assets. We include the following control variables in the above estimation: size, operating cycle, standard deviation of sales, standard deviation of operating cash flows, the number of years over the past five years in which the firm had negative earnings, and the total length of its footnotes. Note that the first five control variables represent the determinants of the amount of estimation in accruals examined in Table 5. We include them in the above regression to ensure that *Estimation* does not simply capture these common firm characteristics. We include the total length of a firm's footnotes to ensure that *Estimation* does not simply proxy for the length of a firm's disclosure documents. We also include

interactions between all control variables and earnings as well as auditor, year, and industry fixed effects.[12]

To disaggregate earnings into a cash flows component and an accruals component, we follow Hribar and Collins (2002) and calculate accruals using the statement of cash flows. If a greater amount of estimation in the accruals generation process lowers the association between the current year's accruals and the following year's earnings, then the interaction between our estimation measure and the accruals portion of earnings should be negative. Additionally, if the number of estimation-related linguistic cues does not capture the precision of the cash portion of earnings, the interaction between cash flows and estimation should be statistically insignificant in the following model:[13]

$$Earnings_{f,t+1} = \beta_0 + \beta_1 Cash_{f,t} + \beta_2 Accruals_{f,t} + \beta_3 Estimation_{f,t} \quad (4)$$
$$+ \beta_4 Cash_{f,t} * Estimation_{f,t} + \beta_5 Accruals_{f,t} * Estimation_{f,t} + \Sigma\beta_i Controls_{f,t}$$
$$+ \Sigma\beta_i Accruals_{f,t} * Controls_{f,t} + \Sigma\beta_i Cash_{f,t} * Controls_{f,t}$$
$$+ AuditorFE_{f,t} + YearFE_t + IndustryFE_{f,t} + \epsilon$$

where $Estimation_{f,t}$, $Earnings$, and the control variables are as defined in equation (3); $Cash_{f,t}$ is the portion of total earnings due to operating cash flows and A$ccruals_{f,t}$is the portion of total earnings due to accruals scaled by average total assets. Like other continuous variables in the regression, both *Cash* and *Accruals* are scaled by the average book value of assets.

4.2.1 Estimation and the Persistence of Accruals Findings

Table 6 Column 2 presents the results for our regression of the following year's earnings on current year's earnings, estimation, the interaction between estimation and the current year's earnings, and our control

[12] 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 their impact on future performance.
[13] The measure of estimation may also capture business uncertainty about the firm. If so, then the coefficient on the interaction between cash flows and estimation should also be negative and statistically significant.

variables (4). These results show that the coefficient on the interaction between estimation and earnings in the current year is negative and statistically significant at 1%. This suggests that earnings that entail greater estimation are less persistent. Economically, the effect of estimation on the persistence of earnings is approximately 17% of the estimated coefficient on the relation between current earnings and future earnings.

In Table 6 Column 3, we present the results for our regression of the next year's earnings on a firm's current year's earnings interacted with our two pieces of accruals estimation, BAE and WAE. These results show that the respective interactions between BAE and current earnings and WAE and current earnings are both statistically significant at the 1% level. Furthermore, these results show that the change in earnings persistence for BAE is greater than that for WAE (-0.053 and – 0.031, respectively), indicating that a one standard deviation change in BAE has a more significant economic impact on the persistence of earnings than WAE. Since an F-test of the difference between the coefficient estimates yields a probability of 0.77, we cannot reject our null hypothesis that the two coefficients will be equal.

In Table 7 Column 4, we present the results for our regression of the next year's earnings on each component of earnings interacted with the number of estimated-related linguistic cues.[14] These results show that the interaction between accruals and estimation is negative and statistically significant, consistent with our prediction that a greater amount of estimation leads to less accruals persistence. On the other hand, we find that the interaction between cash flows and estimation is not statistically significant. Specifically, the difference in the persistence of accruals for a one standard deviation change in estimation is -0.050; the estimated coefficient on the main effect of accruals is 0.549.

In Table 7 Column 5, we present the results for our regression of the next year's earnings on the interaction of BAE and WAE and both accruals and cash flows. Our results show that the respective interactions between the BAE and earnings and WAE and earnings are both statistically significant at the 1% level. Furthermore, an F-test of the coefficients on the respective interactions between BAE and

[14] 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.

accruals and WAE and accruals yields a probability of 0.84, suggesting that these coefficients are not statistically different. Economically, we find that BAE has a more significant impact on the persistence of the accruals portion of earnings than WAE (-0.629 and -0.349, respectively, for a one standard deviation change).

Finally, in Table 8, we present the results for our regression of future earnings on the persistence of accruals and cash flows including our three accruals quality measures: the absolute value of the magnitude of accruals, the standard deviation of the Dechow and Dichev (2002) residual, and special items. For the Dechow and Dichev (2002) residual, we use the standard deviation of the residual from the model over the past five years for each firm (McNichols 2002).

The results in Table 8 Columns 2, 3, and 4 Panel A show that all three measures of accruals quality are statistically significant at the 1% level and negatively associated with the persistence of accruals. Furthermore, the results in Column 5 suggest that, even after including alternate measures of estimation, our measure of estimation is still associated with a lower persistence of accruals at the 1% level of significance. When we exclude these alternate measures, the magnitude of the estimated coefficient on our measure of estimation decreases to -0.695 from -0.989. This finding suggests that our measure is informative about some aspect of accruals persistence not found in these other measures.

The results in Panel B show that the estimated coefficient on the interaction between BAE and accruals is negative and statistically significant at the 1% level while the estimated coefficient on the interaction between WAE and accruals is negative and statistically significant the 5% level. The estimated coefficients for the interactions between BAE and WAE and cash flows are both statistically insignificant. Again, these findings suggest that both BAE and WAE provide incremental information about the persistence of accruals beyond measures examined in prior studies.

4.3 Accruals Estimation and the Association between Accruals and Cash Flows

Prediction 2 (P2) suggests that accruals that entail greater estimation should be less likely to be realized as cash. To test this prediction, we use the model developed by Dechow and Dichev (2002) (DD).

Specifically, the DD model estimates how well working capital accruals map into realized operating cash flows. The model is based on the premise that accruals are a way to shift the recognition of cash flows.[15] That is, if accruals map into realized cash flows, then a firm's accruals are deemed to be of high quality. The DD model regresses 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 indicates how well a firm's accruals map into cash flows.

The specification of the DD model is shown in equation (5).

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

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; and $PPE_{f,t}$ is the total property plant and equipment for the current fiscal period, following McNichols (2002). All continuous variables are scaled by average total assets.

To estimate how well a firm's accruals map into its cash flows, we estimate the above model by industry and year and use σ*(DD Residual)*, the standard deviation of the residual from the model over the past five years for each firm. We expect that accruals with greater estimation to map less accurately into cash flows.

4.3.1 Accruals Estimation and the Mapping of Accruals into Cash Flow

Table 9 presents the results for our test of Prediction 2 (P2), the relation between the amount of accruals estimation and the association between accruals and past, present, and future cash flows. In tests excluding firm and year clustering, unreported, we find that all of the determinants of accruals quality are

[15] One limitation of this model is that it does not distinguish between earnings management and unintentional errors or management uncertainty.

statistically significant and load in the same direction as that found in prior studies (Francis et al. 2005; Dechow and Dichev 2002).

The results in Table 9 Column 3 show that the amount of estimation is statistically significant at the 1% level and positively associated with the standard deviation of the Dechow and Dichev (2002) residual. This finding is consistent with our prediction that the amount of estimation in accruals is associated with a greater number of accrual errors and therefore exhibits a lower relation between accruals and cash flows.

The results in Table 9 Column 4 show that our coefficients for BAE and WAE are statistically significant at the 5% and 1% level, respectively, and are both positively associated with a lower relation between accruals and cash flows. Since our F-test of the equality of the coefficients on BAE and WAE yields a probability of 0.25, we do not reject the null hypothesis that BAE and WAE have the same implications for accruals quality. This result suggests that both components of estimation affect the mapping of accruals into cash flows. 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

To test Prediction 3 (P3), we follow the research design of Sloan (1996) and Richardson et al. (2005) to determine how quickly the market appears to react to estimation information. In particular, we are interested in whether the market reacts as if it quickly incorporates the estimation information found in the footnotes and CAP disclosures. Using this approach, Sloan (1996) finds that positive (negative) accruals are associated with negative (positive) future abnormal returns. These findings are consistent with his hypothesis that the market over-values accruals persistence.

For our study, since we are interested in the incremental effect of the amount of estimation on the persistence of accruals, we add an interaction term to the model to allow us to examine the interaction between the amount of estimation and total accruals. The interaction term captures the marginal effect of the amount of estimation on the association between current accruals and future abnormal returns.

Specifically, a negative interaction would indicate that the market over-(under-) values positive (negative) accruals.

We make two changes to the specification of the model to better coincide with our research question. Sloan (1996) calculates future abnormal returns beginning four months after the end of the firm's fiscal period. Our abnormal returns accumulation begins five days after the 10-K filing date. Next, we use the raw amount of accruals in our model rather than a decile ranking. One of the purposes of Sloan (1996) is to show that a trading strategy can be implemented by purchasing stock in firms with extreme low accruals and shorting those with extreme high accruals. However, since our aim is to examine whether the market quickly incorporates the estimation information found in the footnotes, we use raw accruals to preserve more of the information contained in the accruals. Our modified version of the model is as follows:

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

where $Estimation_{f,t}$ is as defined earlier; $Accruals_{f,t}$ is income before extraordinary items minus operating cash flows scaled by average total assets; $Size_{f,t}$ is the log of the market value of the firm's equity; $BTM_{f,t}$ is the book to market ratio of the firm; $ETP_{f,t}$ is the firm's earnings to price ratio; $Beta_{f,t}$ is the market beta of the firm for fiscal period *t*; and $Abnreturns_{f,t}$ is the one-year buy-and-hold market adjusted abnormal returns, beginning five days after the filing of the 10-K. We estimate the model using Fama-Macbeth to correct for any cross sectional correlation in the error term.

We use a four-factor model to further test the association between the amount of estimation in a firm's accruals and its 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 by 5 rankings of estimation. For each portfolio, we then estimate the four-factor alpha using the following model.

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

where $MonthyExr_t$ is the monthly excess return of the value or equal weighted portfolios, $MktExr_t$ is the monthly return of the value-weighted index minus the risk free rate, HML_t is the monthly premium of the book-to-market factor, SMB_t is the monthly premium of the size factor, and UMD_t is the monthly premium on winners minus losers.[16]

4.4.1 Accruals Estimation and Returns

The results in Table 10 Column 3 show that the relation between the interaction between accruals and estimation and future abnormal returns is not statistically significant. Similarly, the results in Column 4 show that neither BAE nor WAE is informative of the relation between accruals and future abnormal returns. Together, these findings suggest that the market incorporates the estimation information found in the notes and CAP disclosures into their valuation of a firm.

Furthermore, the results in Table 11 for our test of P3 using the Carhart four-factor model show no evidence that the accrual anomaly is more concentrated in firms with a greater amount of estimation in their accruals. The results in Panel A show that using equal weighted hedged returns yields a significant relation between accruals and future abnormal returns. However, using value weighted hedge returns yields no significant relation between the two. Furthermore, the results in Panel A show no evidence that the accrual anomaly is exacerbated in firms with greater accruals estimation. This result is consistent with the hypothesis that investors utilize the estimation information found in the notes to the financial statements and CAP disclosures when valuing a firm.

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

Lastly, the results in Panel B show no discernible association between future abnormal returns and the amount of estimation and accruals when using either value-weighted or equal weighted returns for firms with high or low BAE and WAE.

5. Robustness Tests

5.1 Placebo Tests Based on Bootstrapping

To provide further confidence in our results, we conduct a number of robustness tests. First, we examine whether the results using our linguistic cues approach are "random" by conducting a bootstrap test of our main findings, Prediction 1 (P1) and Prediction 2 (P2). To do so, we begin by ranking all of the words in the notes to the financial statements in our sample by their frequencies. We next select the five words above and the five words below a word for each of the 47 unique words in our Estimation Actions, Estimation Objects, and Estimation Adjectives dictionaries that appears at least once in a firm's notes.[17] This process yields 10 unique dummy words for each estimation word and a total of 470 unique placebo estimation words. We list these words in a placebo dictionary that contains words which have a similar frequency of use as those in our main dictionaries but whose use variation differs.

We next randomly select 1 word from the 10 dummy words chosen for each estimation word and count the number of times each placebo word is mentioned in a firm's footnotes or CAP disclosures.[18] We then estimate the models for the tests of (P1) and (P2) substituting this placebo estimation count in place of our original count of the number of estimation related linguistic cues, repeating this process 1000 times.

Table 12 presents our results for the joint tests of significance (insignificance). These results show that only 2% of the bootstrap placebo tests yield results similar to those of our main findings for the persistence of earnings and the quality of accruals, (P1) and (P2) respectively. Specifically, only 2% of the bootstraps yield a negative and statistically significant result on the coefficient of earnings interacted

[17] The words anticipating, approximation, approximations, beliefs, believing, estimations, and expecting are not mentioned in any of our sample firms' footnotes but are included in the original dictionaries for completeness.

[18] As an additional robustness check, we randomly select 47 words from the 470 placebo words instead of 1 word from the 10 placebo words chosen for each estimation word. This change in the selection process does not significantly affect the results of our simulation.

with the placebo estimation count (P1a), a negative and statistically significant result on the interaction between accruals and the placebo estimation count and insignificant result on the coefficient of the interaction between cash flows and the placebo estimation count (P1b), and a positive and statistically significant association between the placebo estimation count at the standard deviation of the Dechow and Dichev residual (P2). Together, these results lend support to our findings.

To further examine whether our *Estimation* measure captures accruals estimation, we re-run our analyses including the number of estimation-related linguistic cues in the non-CAP section of the management discussion and analysis. Since this estimation count is not directly related to accruals recognition, it is likely to capture only generic business uncertainty only. Therefore, we do not expect to find similar empirical results using this variable. In untabulated results, we find that this alternate measure is not statistically significantly associated with earnings persistence, accruals persistence, or the Dechow-Dichev accruals quality.

5.2 Controlling for Other Textual Characteristics of 10-Ks

In a final set of robustness tests, we examine whether our findings reflect other textual characteristics of 10-K filings. For example, prior research finds that the persistence of earnings is lower for firms whose notes to the financial statements are more difficult to read (Li 2008). Following this result, if notes regarding accruals estimation are more difficult to read due to the complexity of conveying the estimation process, it may be that our results reflect textual difficulty rather than the amount of estimation in the accruals generation process. To address this possibility, we examine the readability of the notes and CAP disclosures in our study using the Gunning-Fox index. We then include this measure (and its interaction with earnings, accruals, and cash flows) into each of our main tests. In untabluated results, we find that our results are not sensitive to this specification.

Lastly, Li, Lundholm, and Minnis (2012) find that the perceived competition intensity reflected in firms' 10-K filings is associated with the mean-reverting speed of abnormal profit. In another set of tests,

untabulated, we find that our results remain essentially the same if the competition measure and its interactions with earnings, accruals, and cash flows are included as additional control variables.

6. Conclusion

Motivated by the findings of Sloan (1996) and Richardson et al. (2005), this study examines whether the amount of estimation in the accruals generation process is associated with the persistence of accruals. In particular, previous studies propose that accruals generated with a greater degree of estimation are less precise as they are more likely to be recorded with error. By consequence, they are then less predictive of future earnings. Using the notes to the financial statements and CAP disclosures from a sample of 10-K filings, we examine the association between the estimation involved during the accrual generating process and the persistence of accruals.

Our results provide evidence consistent with the conjecture that the estimation needed during the accrual generating process plays a key role in the persistence of accruals. Specifically, we find that when accruals are based on a greater degree of estimation, they are less predictive of future earnings. We also find that these accruals map less into the past, current, or future cash flows of the firm. Decomposing estimation into between- and within-account components, we find that both types of estimation drive our results. Lastly, we find that the amount of accruals estimation is not informative of the relation between accruals and the future abnormal returns of the firm. Overall, the findings in this study provide insight into the accrual generating processing. More importantly, our findings suggest that understanding the process that recording accruals entails plays an important role in understanding both the persistence and quality of a firm's accruals.

References

Basu, S. 1997. The Conservatism Principle and the Asymmetric Timeliness of Earnings, *Journal of Accounting & Economics* 24, 3-37.

Billings, M.B., 2011. Discussion of "Critical Accounting Policy Disclosures." Journal of Accounting, Auditing & Finance 26, 77–80.

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* 11, 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., Franco Wong, M.H., Zhou, Y., 2011. Accounting adjustments and the valuation of financial statement note information in 10-K filings. *The Accounting Review* 86, 1577–1604.

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

Frankel, R.M., Jennings, J.N., Lee, J.A., 2015. Using Unstructured and Qualitative Disclosures to Explain Accruals. *Working Paper,* Washington University in St. Louis and Florida State University

Green, J., Hand, J. R. M., & Soliman, M. T. (2011). Going, Going, Gone? The Apparent Demise of the Accruals Anomaly. *Management Science 57*(5), 797–816

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., Hribar, P., Collins, D.W., Collins, D.W., 2002. Errors in Estimating Accruals: Implications for Empirical Research. *Journal of Accounting Research* 40, 105–134.

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.

Levine, C. B., and Michael J. S. 2011, Critical Accounting Policy Disclosures, *Journal of Auditing, Accounting and Finance* 26 (Winter): 39–75.

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.

Li, F., Lundholm, R., Minnis, M., 2013. A Measure of Competition Based on 10-K Filings. *Journal of Accounting Research* 51, 399–436.

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

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.

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

Radin, A., 2007, Have We Created Financial Statement Disclosure Overload?, *The CPA Journal*. http://www.nysscpa.org/cpajournal/2007/1107/perspectives/p6.htm

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,* Penn State

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

Table 1 - Number of Estimation Linguistic Cues by Fiscal Year

This table presents the average number of estimation related linguistic cues and the total number of words found in the notes to the financial statements and the critical accounting policies sections of the 10-K. Our sample of firms spans from fiscal year 1993 to fiscal year 2012.

Fiscal Year	Estimation	Length	N
1993	36	3,725	802
1994	32	3,486	943
1995	34	3,626	2,445
1996	38	3,727	3,764
1997	41	4,131	4,103
1998	44	4,407	3,946
1999	46	4,689	3,811
2000	49	5,172	3,821
2001	71	6,511	3,703
2002	92	7,922	3,472
2003	102	8,597	3,340
2004	111	8,640	3,233
2005	115	8,786	2,992
2006	134	9,482	2,930
2007	141	9,987	2,926
2008	147	10,369	3,077
2009	150	10,432	2,989
2010	149	9,884	2,781
2011	157	12,023	2,638
2012	155	11,971	2,673
Average	**92**	**7,378**	**3,019**

Figure 1 – Average Number of Estimation Linguistic Cues in the Footnotes and Critical Accounting Policies Section by Fiscal Year

Figure 1 presents the average number of estimation linguistic cues for each fiscal year and average number of words the notes to the financial statements and the critical accounting policies section between fiscal years 1993 and 2012. The average total number of words is represented by the solid red line and its scale is represented by the left y-axis. The average number of estimation linguistic cues is represented by the blue dashed line and its scale is reflected in the right y-axis.

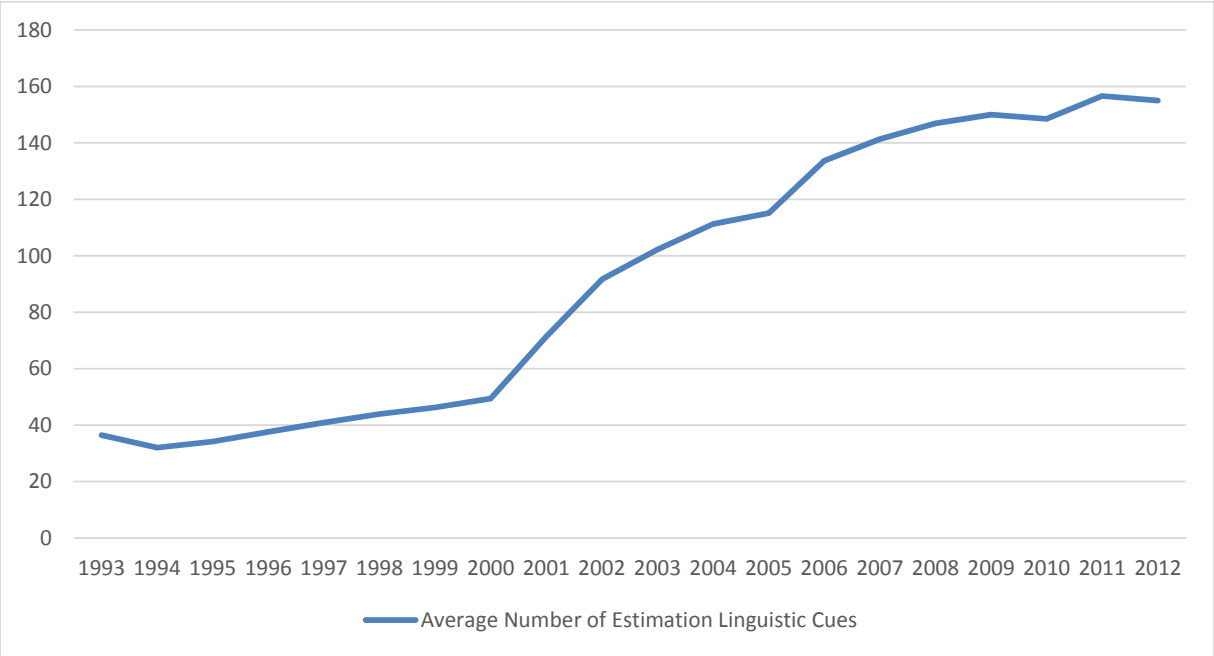

Table X - Change in Estimation Quintile

This table presents the quintile rank of estimation count in period t by quintile rank in period t+1.

	Estimation Quintile t + 1				
Estimation Quintile t	0	1	2	3	4
0	69%	20%	6%	3%	1%
1	20%	48%	22%	7%	2%
2	5%	23%	43%	23%	6%
3	2%	6%	23%	48%	21%
4	1%	1%	5%	20%	74%

Table 2 - Average Number of Words and Estimation Linguistic Cues in the Footnotes and Critical Accounting Policies Section by Industry

This table shows the average number of estimation related linguistic cues and the average number of words in the notes to the financial statements and critical accounting polices section by industry for the fiscal periods between 1993 and 2012.

Industry	SIC Code	Estimation	Length	N
Agricultural Production Crops	1	30	3,348	17
Metal Mining	10	123	8,892	261
Coal Mining	12	158	13,034	73
Oil And Gas Extraction	13	109	8,390	2,449
Mining And Quarrying Of Nonmetallic Minerals, Except Fuels	14	77	5,578	65
Building Construction General Contractors And Operative Builders	15	86	7,303	454
Heavy Construction Other Than Building Construction Contractors	16	111	8,878	182
Construction Special Trade Contractors	17	41	3,885	75
Food And Kindred Products	20	80	6,641	1,531
Textile Mill Products	22	58	5,263	284
Apparel And Other Finished Products Made From Fabrics	23	76	6,541	657
Lumber And Wood Products, Except Furniture	24	75	5,800	343
Furniture And Fixtures	25	74	5,596	453
Paper And Allied Products	26	98	7,001	639
Printing, Publishing, And Allied Industries	27	77	6,237	829
Chemicals And Allied Products	28	99	8,611	5,407
Petroleum Refining And Related Industries	29	101	8,292	400
Rubber And Miscellaneous Plastics Products	30	78	6,103	766
Leather And Leather Products	31	51	5,112	208
Stone, Clay, Glass, And Concrete Products	32	82	6,063	425
Primary Metal Industries	33	81	6,652	1,026
Fabricated Metal Products, Except Machinery And Transportation Equipment	34	79	5,822	1,038
Industrial And Commercial Machinery And Computer Equipment	35	90	6,621	4,274
Electronic And Other Electrical Equipment And Components	36	99	7,126	5,797
Transportation Equipment	37	102	7,156	1,522
Measuring, Analyzing, And Controlling Instruments	38	88	6,665	4,347
Miscellaneous Manufacturing Industries	39	79	6,002	737

Table 2 - Average Number of Words and Estimation Linguistic Cues in the Footnotes and Critical Accounting Policies Section by Industry (continued)

Industry	SIC Code	Estimation	Length	N
Railroad Transportation	40	48	5,023	46
Motor Freight Transportation And Warehousing	42	71	5,393	536
Water Transportation	44	98	8,595	261
Transportation By Air	45	94	7,405	456
Transportation Services	47	85	7,016	238
Communications	48	115	10,406	2,085
Electric, Gas, And Sanitary Services	49	115	9,537	2,437
Wholesale Trade-durable Goods	50	69	6,042	1,750
Wholesale Trade-non-durable Goods	51	78	7,124	984
Building Materials, Hardware, Garden Supply, And Mobile Home Dealers	52	32	3,573	43
General Merchandise Stores	53	81	6,323	470
Food Stores	54	76	5,573	407
Automotive Dealers And Gasoline Service Stations	55	95	8,964	341
Apparel And Accessory Stores	56	77	6,368	806
Home Furniture, Furnishings, And Equipment Stores	57	48	4,168	248
Eating And Drinking Places	58	77	6,376	1,146
Miscellaneous Retail	59	73	6,510	1,374
Hotels, Rooming Houses, Camps, And Other Lodging Places	70	87	7,357	207
Personal Services	72	60	6,341	133
Business Services	73	98	7,775	7,900
Automotive Repair, Services, And Parking	75	37	3,533	22
Motion Pictures	78	87	7,678	336
Amusement And Recreation Services	79	88	8,200	754
Health Services	80	97	8,421	1,340
Educational Services	82	96	7,409	280
Social Services	83	48	5,606	55
Engineering, Accounting, Research, Management, And Related Services	87	93	8,080	1,386
Nonclassifiable Establishments	99	168	12,947	89

Table 3 - Summary Statistics

Table 3 presents the summary statistics for the main sample. *Earnings* is the firm's 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. *Estimation* is the number of estimation related linguistic cues in the footnotes and the critical accounting policies section of the firm's 10-K scaled by 1,000. *Between Account Estimation* and *Within Account Estimation* is estimation broken down into the differences between estimation between accounts and differences in the amount of estimation within a specific account, estimated by industry and year. *Length* is measured as the total number of words in the footnotes and the critical accounting policies section of the 10-K scaled by 1,000. *Operating Cycle* is the log of the operating cycle of the firm. *Size* is measured as the log of the market value of the firm's equity calculated as the share price of the firm's stock at the filing date multiplied by the number of shared outstanding. *σ(Operating Cash Flows)* is the standard deviation of the operating cash flows over the past 5 years. *σ(Sales)* is the standard deviation of sales over the past 5 years. *Negearn* is the number of years in which the company had negative earnings over the past 5 years.

Variable	**N**	**Mean**	**P25**	**Median**	**P75**	**σ**
Earnings	60,389	-0.037	-0.052	0.031	0.077	0.242
Accruals	60,389	-0.077	-0.112	-0.056	-0.014	0.143
Operating Cash Flows	60,389	0.039	0.003	0.074	0.133	0.186
Estimation	60,389	0.092	0.039	0.072	0.127	0.069
Between Account Estimation	56,858	0.093	0.043	0.078	0.132	0.061
Within Account Estimation	56,858	0.000	-0.014	0.000	0.010	0.033
Length	60,389	7.402	3.233	5.622	9.337	6.379
Size	60,389	5.532	4.007	5.550	7.005	2.168
Operating Cycle	60,389	4.607	4.197	4.677	5.108	0.806
σ(Operating Cash Flows)	60,389	0.089	0.032	0.057	0.105	0.098
σ(Sales)	60,389	0.219	0.078	0.147	0.273	0.226
Negearn	60,389	1.631	0.000	1.000	3.000	1.744

Table 4 - Pearson and Spearman Univariate Correlations

This table presents the Spearman (above diagonal) and Pearson (below diagonal) correlation for the main variables used in this study. *Earnings* is the firm's 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. *Estimation* is the number of estimation related linguistic cues in the footnotes and the critical accounting policies section of the firm's 10-K. *Between Account Estimation* and *Within Account Estimation* is estimation broken down into the differences between estimation between accounts and differences in the amount of estimation within a specific account scaled by 1,000, estimated by industry and year. *Length* is measured as the total number of words in the footnotes and the critical accounting policies section of the 10-K scaled by 1,000. *Operating Cycle* is the log of the operating cycle of the firm. *Size* is measured as the log of the market value of the firm's equity calculated as the share price of the firm's stock at the filing date multiplied by the number of shared outstanding. *σ(Operating Cash Flows)* is the standard deviation of the operating cash flows over the past 5 years. σ(Sales) is the standard deviation of sales over the past 5 years. *Negearn* is the number of years in which the company had negative earnings over the past 5 years.

	Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	Earnings	1.00	0.41	0.67	-0.08	-0.05	-0.06	-0.12	0.39	-0.01	-0.25	-0.07	-0.67
2	Accruals	0.63	1.00	-0.24	-0.10	-0.09	-0.04	-0.11	0.09	0.23	-0.09	0.02	-0.25
3	Operating Cash Flows	0.79	0.05	1.00	0.01	0.03	-0.04	-0.04	0.37	-0.19	-0.27	-0.11	-0.53
4	Estimation	-0.01	-0.07	0.04	1.00	0.90	0.33	0.88	0.35	-0.09	-0.14	-0.13	0.10
5	Between Account Estimation	0.01	-0.05	0.05	0.88	1.00	-0.06	0.80	0.32	-0.09	-0.14	-0.13	0.08
6	Within Account Estimation	-0.05	-0.04	-0.02	0.48	0.00	1.00	0.27	0.10	0.00	-0.03	-0.02	0.05
7	Length	-0.06	-0.07	-0.03	0.70	0.60	0.35	1.00	0.31	-0.10	-0.10	-0.10	0.15
8	Size	0.33	0.14	0.32	0.36	0.35	0.12	0.24	1.00	-0.05	-0.40	-0.30	-0.42
9	Operating Cycle	-0.03	0.13	-0.14	-0.05	-0.06	0.01	-0.06	-0.01	1.00	0.10	-0.06	0.00
10	σ(Operating Cash Flows)	-0.46	-0.18	-0.46	-0.10	-0.11	-0.01	-0.04	-0.28	0.04	1.00	0.47	0.43
11	σ(Sales)	-0.08	-0.02	-0.10	-0.12	-0.13	-0.02	-0.06	-0.25	-0.11	0.34	1.00	0.18
12	Negearn	-0.59	-0.28	-0.55	0.06	0.04	0.06	0.10	-0.40	-0.01	0.41	0.12	1.00

Table 5 – Regression of the Number of Estimation Linguistic Cues on Determinants of Accruals Estimation

This table presents the tobit model left censored at 0 of the number of estimation linguistic cues found in the notes to the financial statements and the critical accounting policies section of the 10-K on determinants of accruals estimation scaled by 1,000. *Size* is measured as the log of 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. *Operating Cycle* is the log of the operating cycle of the firm. *σ(Sales)* is the standard deviation of sales over the past 5 years. *σ(Operating Cash Flows)* is the standard deviation of the firms operating cash flows over the past 5 years. *Negative Earnings* is the number of years in which the company had negative earnings over the past 5 years. 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. 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.

	Estimation
Size	0.0099***
	(0.000)
Operating Cycle	0.0011**
	(0.038)
σ(Sales)	0.0045***
	(0.000)
σ(Operating Cash Flows)	-0.0510***
	(0.000)
Negative Earnings	0.0071***
	(0.000)
Constant	-0.0483***
	(0.000)
Observations	60,389

Table 6 - Estimation and Earnings Persistence

This table presents the results for our tests of the amount of estimation on the persistence of earnings. *Earnings* is the firm's income before extraordinary items scaled by average total assets. *Estimation* is the number of estimation related linguistic cues in the footnotes and the critical accounting policies section of the firm's 10-K scaled by 1,000. *Between Account Estimation* and *Within Account Estimation* is estimation broken down into the differences between estimation between accounts and differences in the amount of estimation within a specific account; estimated by industry and year. Controls include *size*, *operating cycle*, *σ(sales)*, *σ(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 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.

	Earnings t+1	**Earnings t+1**
Earnings	0.3312***	0.3473***
	(0.001)	(0.000)
Estimation	0.0098	
	(0.639)	
Estimation x Earnings	-0.8410***	
	(0.008)	
Between Account Estimation		0.0339
		(0.295)
Between Account Estimation x Earnings (a)		-0.8730***
		(0.006)
Within Account Estimation		-0.0230
		(0.250)
Within Account Estimation x Earnings (b)		-0.9313***
		(0.001)
Equivalence of Coefficients (a) = (b)		Prob. > F = 0.77
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	60,360	56,832
Adjusted R-squared	0.558	0.559

Table 7 - Regression of Future Earnings on Current Accruals and Cash Flows Interacted with the Number of Estimation Linguistic Cues

This tables presents the regression of the next year's earnings on the current year's accruals and operating cash flows interacted with the number of estimation linguistic cues in the notes to the financial statements and critical accounting policies section of the 10-L. *Accruals* are total accruals scaled by average total assets. *Operating Cash Flows* are operating cash flows scaled by average total assets. *Estimation* is the number of estimation related linguistic cues in the footnotes and the critical accounting policies section of the firm's 10-K scaled by 1,000. *Between Account Estimation* and *Within Account Estimation* is estimation broken down into the differences between estimation between accounts and differences in the amount of estimation within a specific account, estimated by industry and year. 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.

	Earnings t+1	**Earnings t+1**	**Earnings t+1**	**Earnings t+1**
Accruals	0.5485***	0.5561***	0.3125***	0.3106***
	(0.000)	(0.000)	(0.000)	(0.000)
Operating Cash Flows	0.8798***	0.8816***	0.6404***	0.6425***
	(0.000)	(0.000)	(0.000)	(0.000)
Estimation	-0.0538		-0.0793**	
	(0.107)		(0.014)	
Estimation x Accruals	-0.9650***		-0.9890***	
	(0.000)		(0.000)	
Estimation x Operating Cash Flows	0.1603		-0.0236	
	(0.268)		(0.922)	
Between Account Estimation		-0.0112		-0.0684**
		(0.791)		(0.040)
Between Account Estimation x Accruals (a)		-1.0341***		-1.0324***
		(0.000)		(0.000)
Between Account Estimation x Operating Cash Flows		0.1351		-0.0426
		(0.513)		(0.890)
Within Account Estimation		-0.0982***		-0.1006***
		(0.010)		(0.006)
Within Account Estimation x Accruals (b)		-0.9530***		-1.0598***
		(0.000)		(0.000)
Within Account Estimation x Operating Cash Flows		0.1099		-0.0154
		(0.633)		(0.936)
Equivalence of Coefficients (a) = (b)		Prob. > F = 0.69		Prob. > F = 0.84
Constant	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Controls Interact w/ Operating Cash Flows	-	-	Yes	Yes
Controls Interact w/ Accruals	-	-	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	60,360	56,832	60,360	56,832
Adjusted R-squared	0.579	0.579	0.590	0.589

Table 8 - Estimation on Persistence with other Measures of Accruals Quality

Table 8 Panel A 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. 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.

Panel A: Estimation

	Earnings t+1	**Earnings t+1**	**Earnings t+1**	**Earnings t+1**
Accruals	0.3004***	0.3749***	0.4159***	0.3975***
	(0.000)	(0.000)	(0.000)	(0.000)
Operating Cash Flows	0.6070***	0.6857***	0.7149***	0.7239***
	(0.000)	(0.000)	(0.000)	(0.000)
Estimation	-0.0724**	-0.0682**	-0.0912***	-0.0847***
	(0.018)	(0.042)	(0.004)	(0.006)
Estimation x Accruals	-0.8397***	-0.8445***	-0.8445***	-0.6952***
	(0.000)	(0.000)	(0.000)	(0.001)
Estimation x Operating Cash Flows	-0.0057	0.0260	0.0957	0.1211
	(0.978)	(0.914)	(0.698)	(0.575)
σ(DD Residual)	-0.0747***			-0.0650***
	(0.000)			(0.000)
σ(DD Residual) x Accruals	-0.4334***			-0.2509**
	(0.000)			(0.014)
σ(DD Residual) x Operating Cash Flows	0.1263*			0.1817**
	(0.085)			(0.028)
Abs(Accruals)		-0.0424		-0.0424
		(0.137)		(0.109)
Abs(Accruals) x Accruals		-0.2658***		-0.1731***
		(0.000)		(0.007)
Abs(Accruals) x Operating Cash Flows		-0.1400***		-0.1724***
		(0.005)		(0.007)
Special Items			0.0027	0.0046**
			(0.204)	(0.013)
Special Items x Accruals			-0.1468***	-0.1352***
			(0.000)	(0.000)
Special Items x Operating Cash Flows			-0.0826***	-0.0834***
			(0.000)	(0.000)

Table 8 - Estimation on Persistence with other Measures of Accruals Quality

Panel B: Between Account and Within Account Estimation and Accruals Persistence with Other Measures of Accruals Quality

	Earnings t+1	**Earnings t+1**	**Earnings t+1**	**Earnings t+1**
Accruals	0.3033***	0.3693***	0.4113***	0.3964***
	(0.000)	(0.000)	(0.000)	(0.000)
Operating Cash Flows	0.6110***	0.6871***	0.7141***	0.7282***
	(0.000)	(0.000)	(0.000)	(0.000)
Between Account Estimation	-0.0627*	-0.0570	-0.0906***	-0.0857**
	(0.055)	(0.102)	(0.008)	(0.012)
Between Account Estimation x Accruals	-0.9014***	-0.8838***	-0.8817***	-0.7676***
	(0.000)	(0.000)	(0.000)	(0.002)
Between Account Estimation x Operating Cash Flows	-0.0241	0.0233	0.0917	0.1124
	(0.931)	(0.939)	(0.764)	(0.683)
Within Account Estimation	-0.0881**	-0.0908**	-0.1011***	-0.0903**
	(0.024)	(0.027)	(0.005)	(0.035)
Within Account Estimation x Accruals	-0.8322***	-0.9436***	-0.8932***	-0.6685**
	(0.000)	(0.001)	(0.000)	(0.021)
Within Account Estimation x Operating Cash Flows	-0.0042	0.0049	0.0997	0.1340
	(0.979)	(0.979)	(0.616)	(0.485)
σ(DD Residual)	-0.0731***			-0.0625***
	(0.000)			(0.001)
σ(DD Residual) x Accruals	-0.4244***			-0.2479**
	(0.000)			(0.028)
σ(DD Residual) x Operating Cash Flows	0.1244			0.1793**
	(0.110)			(0.039)
Abs(Accruals)		-0.0433		-0.0452*
		(0.130)		(0.090)
Abs(Accruals) x Accruals		-0.2595***		-0.1672**
		(0.000)		(0.012)
Abs(Accruals) x Operating Cash Flows		-0.1438***		-0.1782***
		(0.003)		(0.002)
Special Items			0.0031	0.0049***
			(0.136)	(0.002)
Special Items x Accruals			-0.1451***	-0.1358***
			(0.000)	(0.000)
Special Items x Operating Cash Flows			-0.0831***	-0.0854***
			(0.000)	(0.000)

Table 9 – Regression of the Standard Deviation of the Dechow-Dichev Residual on the Number of Estimation Linguistic Cues

This table shows the association between accrual quality, as measured by the Dechow and Dichev (2002) residual, and the number of estimation linguistic cues in the notes to the financial statements and the critical accounting policies section of the 10-K scaled by 1,000. *σ(DD Residual)* is the standard deviation of the Dechow and Dichev residual over the past 5 years, estimated by industry and year. *Beta* is the firm's annual beta. *Estimation* is the number of estimation related linguistic cues in the footnotes and the critical accounting policies section of the firm's 10-K. *Between Account Estimation* and *Within Account Estimation* is estimation broken down into the differences between estimation between accounts and differences in the amount of estimation within a specific account, estimated by industry and year. *Length* is measured as the total number of words in the footnotes and the critical accounting policies section of the 10-K. *Size* is measured as the log of the market value of the firm's equity 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 in which the company had negative earnings over the past 5 years. *σ(Sales)* is the standard deviation of sales over the past 5 years. *σ(Operating Cash Flow)* is the standard deviation of the operating cash flows over the past 5 years. *Operating Cycle* is the log of the operating cycle of the firm. 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. 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.

	σ(DD Residual)	**σ(DD Residual)**	**σ(DD Residual)**
Estimation		0.0570**	
		(0.020)	
Between Account Estimation (a)			0.0521**
			(0.046)
Within Account Estimation (b)			0.0819***
			(0.005)
Equivalence of Coefficients (a) = (b)			Prob. > F = 0.25
Length		0.0004***	0.0004***
		(0.001)	(0.002)
Size	-0.0023***	-0.0030***	-0.0032***
	(0.001)	(0.000)	(0.000)
Negearn	0.0130***	0.0123***	0.0122***
	(0.000)	(0.000)	(0.000)
σ(Sales)	0.0438***	0.0428***	0.0422***
	(0.000)	(0.000)	(0.000)
σ(Operating Cash Flows)	0.4042***	0.4079***	0.4076***
	(0.000)	(0.000)	(0.000)
Operating Cycle	0.0019	0.0016	0.0018
	(0.101)	(0.181)	(0.125)
Constant	Yes	Yes	Yes
Fixed Effects	Industry, Year, Auditor	Industry, Year, Auditor	Industry, Year, Auditor
Cluster	Industry, Year	Industry, Year	Industry, Year
Observations	56,271	56,271	52,860
Adjusted R-squared	0.394	0.396	0.394

Table 10 – Fama-Macbeth Regression of Future Abnormal Returns on Accruals Interacted with Estimation

This table presents the association between one-year abnormal returns beginning 5 days following the filing date and accruals. *Accruals* are total accruals scaled by average total assets. *Estimation* is the number of estimation related linguistic cues in the footnotes and the critical accounting policies section of the firm's 10-K scaled by 1,000. *Between Account Estimation* and *Within Account Estimation* is estimation broken down into the differences between estimation between accounts and differences in the amount of estimation within a specific account, estimated by industry and year. Controls include: *Length* is measured as the total number of words in the footnotes and the critical accounting policies section of the 10-K. *Size* is measured as the log of the market value of the firm's equity calculated as the share price of the firm's stock at the filing date multiplied by the number of shared outstanding. *Book-to-Market* is the book to market ratio calculated as the book value of assets divided by the market value of equity plus liabilities. *Earnings-to-Price* is the firm's earnings to price ratio. *Beta* is the firm annual beta. 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.

	Abnormal Returns t+1	**Abnormal Returns t+1**	**Abnormal Returns t+1**
Accruals	-0.2106***	-0.3475***	-0.3672***
	(0.001)	(0.002)	(0.003)
Estimation		-0.0261	
		(0.906)	
Estimation x Accruals		2.0918	
		(0.156)	
Between Account Estimation			-0.1294
			(0.671)
Between Account Estimation x Accruals			2.0862
			(0.240)
Within Account Estimation			0.1523
			(0.517)
Within Account Estimation x Accruals			1.7316
			(0.363)
Controls	Yes	Yes	Yes
Average R-squared	0.054	0.056	0.058

Table 11 - Estimation and Future Abnormal Returns Using Fama-French Carhart Four-Factor Alpha

This table presents the results for association between estimation and future abnormal returns using 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.

Panel A: Accruals Estimation

	Equal Weighted Returns				Value Weighted Returns		
	Accruals				Accruals		
	Low	High	Hedge		Low	High	Hedge
All Observations	0.640***	0.0429	0.597***		0.1560	-0.2236	0.3796
	(0.005)	(0.773)	(0.000)		(0.461)	(0.103)	(0.131)
Estimation	Low	High	Hedge		Low	High	Hedge
Low	0.733***	0.0100**	0.723***	Low	0.3809	-0.5376**	0.9185**
	(0.003)	(0.050)	(0.001)		(0.183)	(0.026)	(0.015)
High	0.4509	0.0052	0.4457	High	-0.1038	-0.1196	0.0158
	(0.130)	(0.974)	(0.115)		(0.767)	(0.479)	(0.968)
		Difference	0.2770			Difference	0.9027*
			(0.429)				(0.097)

Panel B: Between Accrual Estimation and Within Accrual Estimation

	Equal Weighted Returns				Value Weighted Returns		
	Accruals				Accruals		
	Low	High	Hedge		Low	High	Hedge
All Observations	0.657***	0.0883	0.568***		0.1153	-0.2515*	0.3668
	(0.006)	(0.577)	(0.001)		(0.589)	(0.086)	(0.147)
Between Accrual Estimation	Low	High	Hedge		Low	High	Hedge
Low	0.717***	-0.0862	0.804***	Low	0.3149	-0.2177	0.5326*
	(0.003)	(0.636)	(0.000)		(0.215)	(0.304)	(0.094)
High	0.3660	0.0585	0.3075	High	-0.3121	-0.1866	-0.1255
	(0.170)	(0.688)	(0.246)		(0.316)	(0.260)	(0.717)
		Difference	0.4962			Difference	0.6581
			(0.140)				(0.161)
Within Accrual Estimation	Low	High	Hedge		Low	High	Hedge
Low	0.6816**	0.1558	0.5258**	Low	0.0738	-0.3729	0.4467
	(0.012)	(0.415)	(0.028)		(0.832)	(0.688)	(0.277)
High	0.6794**	0.1985	0.4809*	High	0.0490	-0.3874*	0.4364
	(0.036)	(0.273)	(0.073)		(0.886)	(0.099)	(0.295)
		Difference	0.0449			Difference	0.0103
			(0.900)				(0.986)

Table 12 - Joint Tests of Prediction 1 and Prediction 2 Using Placebo Dictionaries

Table 12 presents the results of our joint tests of Predictions 1 and 2 using placebo estimation dictionaries. Placebo estimation dictionaries were created by randomly choosing a placebo word from the list of 10 nearest words in terms of frequency for each of the words in our Estimation Actions, Estimation Objects, and Estimation Adjectives dictionaries. We repeat this procedure 1,000 times and jointly test our main predictions: (P1A): Earnings x Placebo Estimation (Negative), P1B: Accruals x Placebo Estimation (Negative) and Operating Cash Flows x Placebo Estimation (Insignificant) P2: Placebo Estimation (Positive).

	Percentage of Statistically Significant Iterations		
	Statistical Significance Level		
Prediction(s)	**1%**	**5%**	**10%**
(P1A)	35%	48%	53%
(P1B)	15%	18%	15%
(P1A) and (P1B)	12%	16%	13%
(P2)	14%	21%	23%
(P1A) and (P1B) and (P2)	2%	6%	5%

Appendix 1 - Overview of Grammatical Relations

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 provide a wealth of information about the accrual generating process and, more importantly, information about the estimations needed by management.

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

The words and the grammatical relations that are used in written language are 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. Moreover, the association between "I" and "like" denotes 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 need to identify the grammatical relations in each sentence in the notes to the financial statements and several dictionaries of terms associated with estimation (Appendix 2). To identify the grammatical relations 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 relations removes any ambiguity about the meaning of the sentence.

Appendix 2 - Estimation Dictionaries and Grammatical Relations

Estimation Dictionaries

We construct four dictionaries to help measure the estimation conveyed in each firm's 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, Approximate, 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 Relations

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

Direct Object – This relation identifies the accusative object of an action (i.e. "estimate receivables" or "used estimates"). For this grammatical relation, 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 relation is similar to the direct object in that it relates information about an object. The relation that we look for is the same as for the direct object.

Adjective Modifier - Adjective Modifiers modify the meaning of an object (i.e. "likely receivable" or "anticipated value"). For this grammatical relation, we identify when a word from the Estimation Adjective dictionary targets some object – this implies that the object was estimated.

Quantifier Phrase Modifier - This grammatical relation is a modifier to a number (i.e. "approximately $100"). This grammatical association is similar to the adjective modifier but specific to numbers. Here, we look for when a word from the Estimation Adjectives dictionary targets a number.

Noun Compound Subjects – This indicates when a noun is used to modify another noun (i.e. "value estimates"). This grammatical relation is similar to the adjective modifier except that a word from the Estimation Object dictionary is modifying another object.

Appendix 3 - Top 10 and Bottom 10 Accrual Accounts by Estimation

This table presents the accrual accounts with the highest and lowest coefficients on the accrual account dummies from the regression of estimation on accrual account flags. The regression was performed using OLS with an intercept as described in equation (1).

Top 10	Bottom 10
Fair Value	Other Expenses
Regulatory	Long-Term Debt
Intangibles	Credit Arrangements
Derivatives	Significant Customers
Reinsurance	Leases
Restatement	Inventory
Restructuring	Accounting Policies
Stock Compensation	Stock Options
Contingencies	Taxes
Segments	Home Loan