Internet Appendix
(Not for Publication Purposes)

for

Taxes and Financial Constraints: Evidence from Linguistic Cues

March 2, 2015
1. **Banks Closure and Bank Loans**

We analyze the association of the number of failed banks in firms’ headquarter states and the associated loan spreads of carrying out bank loans in Internet Appendix Table 1. The data are obtained from the Loan Pricing Corporation’s DealScan database, a dataset contains detailed pricing information and contract terms of primarily syndicated loans provide by banks and various financial institutions. We follow the specification by Ivashina (2009), and construct several dependent variables to capture the costs of borrowing bank loans. *Bank Loan Spread* is the total annual spread in basis points paid over LIBOR for each dollar drawn from the loan net of upfront fee. *Commitment Fee* is the charge in basis points on the unused portion of the loan. *Upfront Fee* is a one-time charge in basis points collected at the loan initiation. *Standby Fee* is annual charge in basis points on the letter of credits issued. *Annual Fee* is the annual fee in basis points charged on the entire commitment loan amount. They are all in basis points for each loan in DealScan database. The main independent variable is *Number of Failed Banks*, which refers to the number of failed banks in a firm’s headquarter-state in a year based on data obtained from the Federal Deposit Insurance Corporation’s website. The sample period is from 2000 to 2011. We summarize a detailed discussion on the main findings in Section 3.3.

2. **Other Linguistic Cues, ETRs, and Tax Havens**

Our ETR and tax haven results are not sensitive to other linguistic cues. We re-estimate our baseline regressions in Tables V and VI by adding the other eleven linguistic cues and readability measures (19+11=30 variables in total). Internet Appendix Table 2 summarizes our results. Columns 1–6 summarize the results for *Cash ETRs* and *GAAP ETRs*, whereas columns 7–8 report the results *Use of Tax Havens* and *Number of Tax Havens*. Specifically, columns 2–3 and 5–6 report the long-run *Cash ETRs* and *GAAP ETRs* based on the subsequent 3 to 5 years. First, the strong explanatory power of negative words continues even after controlling for all other eleven linguistic cues. The direction, statistical significance, and economic magnitudes are all similar to those reported earlier. Second, there is an absence of any systematic and consistent pattern for any other linguistic cues. While some estimated
coefficients on *Number of Words* appear to be statistically significant in a few (4 out of 8) columns, there is no evidence consistently suggesting that complexity in annual reports is associated with firms’ aggressive tax planning activities. Third, *Use of Negative Words* on average has a much stronger explanatory power than other linguistic cues and complexity measures when we compare all estimated beta coefficients reported in Internet Appendix Table 2. In terms of economic significance (and ignoring the statistical insignificance of other linguistic cues), *Use of Negative Words* consistently ranks among the top five across all specifications when compared to other linguistic cues.

Overall, the above results provide strong support for our conjecture that firms’ tax planning activities are associated with their use of negative words in annual reports, even after controlling for other linguistic cues and readability measures.

3. **Individual UTB Components**

We examine individual components reconciling the beginning and ending UTB Balance. We re-estimate our regression with the full set of control variables identical to Table II, column 5. We interchange the dependent variables with the following three groups of variables. The first group of components measures the additions in *UTB Balance* (1) related to activities in current year and (2) related to activities in prior years. The second group examines the reduction in *UTB Balance* (3) related to activities in current year, (4) related to activities in prior years, (5) due to settlements with tax authorities, and (6) due to lapsed statute of limitations. The last set measures (7) how the *UTB Balance* would affect ETR if recognized, and (8) penalties and interest relating to *UTB Balance*.

Internet Appendix Table 3 summarizes our results for the components of UTB. All dependent variables in Panel A (Panel B) are measured contemporaneously in year $t$ (one-year-ahead in year $t+1$). In Panel A, financially constrained firms on average have higher addition in *UTB Balance* relating to firms’ activities in current and prior years. Given a one standard deviation increase in *Use of Negative Words* is associated with an increase in the *Addition in UTB Balance* by 3.3% ($t$-statistic: 1.76 in column 1) to 4.1% ($t$-statistic: 2.38 in column 2) of *Addition in UTB Balance*’s standard deviation, respectively. These firms
also have higher reductions in UTB due to settlements with tax authorities (coefficient: 0.030; \( t \)-statistic: 1.80 in column 5). However, \textit{Use of Negative Words} does not explain any other decreases, such as those for current and prior year positions or lapsed statute of limitations (in columns 3, 4, and 6). This non-result is no surprise because the “good news” of reductions arising from new information such as court cases affecting other firms in the industry, or lapsed statutes, are routine events unlikely to result from financial constraints. Last, we find that financially constrained firms have higher amounts of the \textit{UTB Balance} that would affect the book ETR if recognized (coefficient: 0.070; \( t \)-statistic: 3.72 in column 7) and higher accrued penalties and interest relating to \textit{UTB Balance} (coefficient: 0.066; \( t \)-statistic: 3.37 in column 8).

In Panel B, we further examine the predictability of firms’ use of negative words on future UTB individual components. We follow the same specification and additionally augment the corresponding contemporaneous, individual UTB components as control variables. Although firms’ use of negative words does not predict future UTB reductions and other UTB items, it predicts the addition to \textit{UTB Balance} in the next year. A one standard deviation increase in \textit{Use of Negative Words} is associated with an increase in \textit{Addition to UTB Balance Related to Current Year} in year \( t+1 \) in column 1 by 3.7\% (\( t \)-statistic: 3.24) of \textit{Addition to UTB Balance Related to Current Year}’s standard deviation, and with an increase in \textit{Addition to UTB Balance Related to Prior Years} in year \( t+1 \) in column 2 by 2.9\% (\( t \)-statistic: 2.00) of \textit{Addition to UTB Balance Related to Prior Years}’ standard deviation.

Overall, the evidence presented above shows that a firm’s use of negative words is strongly associated with—and even predicts—individual components reconciling the beginning and ending \textit{UTB Balance}.

4. \textbf{Alternative Explanation: The Second Test on Disclosure of Tax Risk}

Our second test on disclosure of tax risk exploits a common feature underlying tax-related disclosures: they are inherently uncertain. To illustrate this idea, consider the following 10-K filing of 2009 Altria Group, Inc.:

\textit{Altria Group, Inc.: Lease-in/Lease-out and Sale-in/Sale-out Transactions}
The Internal Revenue Service (IRS)… disallowed benefits pertaining to… leverage lease transactions… approximately $150 million of tax …. The IRS may in the future challenge and disallow more of PMCC’s leveraged lease benefits based on… subsequent case law addressing specific types of leveraged leases (lease-in/lease-out (“LILO”) and sale-in/lease-out (“SILO”) transactions)…. The total… would be approximately $1.0 billion.

This LILO/SILO (Lease-in-lease-out/Sale-in-lease-out) disclosure describes a tax planning activity that is highly uncertain, as the IRS would in future seek to disallow the taxpayers’ tax treatments.

The same uncertainty underlies Amazon’s online sales:

**Amazon.com, Inc.: Tax Collection from Online Sales**

We do not collect sales or other taxes on shipments of most of our goods into most states in the U.S…. The imposition by state and local governments of various taxes upon Internet commerce could create administrative burdens for us, put us at a competitive disadvantage if they do not impose similar obligations on all of our online competitors and decrease our future sales…. A successful assertion by one or more states or foreign countries that we should collect sales or other taxes on the sale of merchandise or services could result in substantial tax liabilities for past sales, decrease our ability to compete with traditional retailers, and otherwise harm our business.

Motivated by the above, we consider three alternative linguistic categories to capture the uncertainty in firms’ disclosure: (1) uncertainty words, (2) possibility words, and (3) strong words, following Loughran and McDonald (2011). Uncertainty words capture the general notion of future uncertainty. Common examples of uncertainty words include approximate, contingency, depend, fluctuate, and uncertain. Possibility (e.g., could, may, suggest, possibly, and possible) and strong (e.g., will, must, never, always, and definitely) words capture low and high levels of confidence, respectively. Loughran and McDonald (2011) show that these word categories significantly predict firms’ likelihood of shareholder class action lawsuits.

Internet Appendix Table 4, columns 1-3, reports the results of these additional tests. None of the estimated coefficients on each word category are significant, and the strong explanatory power of negative words remains robust and significant across the different columns. The absence of systematic pattern on other linguistic cues results suggests that our results are not driven by firms’ disclosure of tax risk. Together with the evidence presented earlier, the findings of all these different tests unanimously suggest that our results are unlikely to be driven by firms’ tax risk.
5. **Alternative Explanation: Expected and Realized Future Profitability**

We construct several proxies for future profitability, and augment each of them one at a time in our baseline regression from column 5 of Table II. The results are summarized in Internet Appendix Table 5. We also control for realized future profitability in our baseline regression from column 5 of Table II. We summarize the results in Internet Appendix Table 6. Our results remain robust.

6. **Alternative Explanation: Litigation Risk**

Perhaps our association between tax aggressiveness and negative words captures firms’ litigation risk. As firms pursue more aggressive tax planning activities, they could face higher likelihood of being sued by shareholders or other stakeholders. The IRS could also initiate tax litigation against firms that pursue aggressive tax avoidance strategies without economic substance. Firms could also initiate lawsuits against the tax authority to claim refunds of tax deficiencies pre-paid during audit. For instance, more than half of illegal tax shelter firms sue the commissioner or the U.S. government to preserve their tax treatments (Graham and Tucker 2006). Thus, if firms’ litigation is generally disclosed in negative tones (as evidenced in Altria’s example above), the use of negative words could simply capture their litigation risk.

Motivated by prior studies, we use firms’ (4) use of litigious words, (5) positive words, and (6) constraining words to quantify their litigation risk. Litigious words (e.g., claimant, deposition, interlocutory, testimony, and tort) primarily capture ongoing or settled lawsuits or reflect firm-specific litigious environments following Loughran and McDonald (2011), who find that their list of litigious words predicts the likelihood of being involved in class action lawsuits. Positive words measure the positive tones in firms’ annual reports. Rogers, van Buskirk, and Zechman (2011) find firms that use more optimistic words are more likely to be the targets of lawsuits, although in an early study Francis, Philbrick, and Schipper (1994) find no evidence that firms that are subject to shareholder lawsuit use more optimistic words in their pre-earnings disclosures. If firms use more positive words to suppress the negative aspects of their disclosure, we should observe a weaker explanatory power of negative words.
after controlling for the use of positive words. Constraining words measure firms’ contractual obligations and commitments (e.g., required, requirement, restricted, commitment, and covenant). Firms have higher litigation risk if they are not able to fulfill their existing contractual obligations and commitments.

Internet Appendix Table 4, columns 4-6, reports the estimated coefficients from these supplemental tests. None of the other linguistic cues are significantly (either economically or statistically) related to the UTB Balance, whereas Use of Negative Words continues to be positive and significant. The systematic patterns are robust under various specifications.

7. Alternative Explanation: Obfuscation and Readability

While firms could pursue aggressive tax planning activities, firm managers have incentives to suppress the negative information and obfuscate the readability of their annual reports (Bloomfield 2008; Wang 2011). Prior research also extensively studies the readability of corporate disclosure, generally showing that market participants incorporate the information into stock prices more completely if corporate filings are easier to read (You and Zhang 2009; Miller 2010; Lehavy, Li, and Merkley 2011). Because disclosing contingent tax liabilities or reserves may reveal tax strategies, firms tend to concede their existence but rarely disclose their values pre FIN-48 (Gleason and Mills 2002; Blouin and Tuna 2007).

Motivated by the above studies, we compute (7) the use of negation words (Use of Negation Words) and (8) number of words (Number of Words) to capture readability and obfuscation. Negation words (e.g., do not benefit) refer to one of the six words (no, not, none, neither, never, nobody) occurring within three words preceding a positive word. Number of Words is defined as the natural logarithm of the number of words in firm $j$’s annual 10-K filing in year $t$. Bloomfield (2008) conjectures that the increase in disclosure length correlates with firms’ litigation risk.

We also use several widely used readability measures: (9) Gunning Fog Index, (10) Flesch Reading Ease, and (11) Kincaid Readability. These different indices are all measures based on the readability of a firm’s annual 10-K filings, computed following Li (2008). Gunning Fog Index is defined as the average
number of words per sentence plus the percentage of complex words, multiplied by 0.4. Complex words have three or more syllables. Textual content with higher *Gunning Fog Index* is more difficult to read.

*Flesch Reading Ease* is calculated as \(206.835 - (1.015 \times \text{average number of word per sentence}) - (84.6 \times \text{average number of syllable per word})\). Textual content with higher *Flesch Reading Ease* is easier to read with a bound between zero (the most difficult to read) and one (the easiest to read). *Kincaid Readability* is defined as the sum of \(11.8 \times \text{average syllable per word}\) and \(0.39 \times \text{average number of word per sentence}\), minus 15.59. Textual content with higher *Kincaid Readability* is best understood by readers with more years of education.\(^1\)

Internet Appendix Table 4, columns 7-11 report the results of re-estimating our baseline regression from column 5 of Table II with these additional readability variables.\(^2\) The same patterns emerge. First, none of the readability measures explain unrecognized tax benefit balances. Second, the coefficients on *Use of Negative Words* continue to be positive and significant. Overall, across all eleven columns, *Use of Negative Words* continues to be positive and significant (with little variation ranging from 0.082, \(t\)-statistic: 3.94 to 0.100, \(t\)-statistic: 4.04) with the economic magnitudes similar—and sometimes even stronger—to those reported in Table II. However, perhaps future work can re-evaluate readability in connection with IRS tax return disclosures.

8. **Theoretical Construct of Linguistic Measures**

Prior studies sometimes interchangeably use different word categories in different contexts. For example, Nelson and Pritchard (2014) show that firms subject to greater litigation risk use more readable language. Hanley and Hoberg (2010) find that, while the risk factor section in an IPO prospectus contains a significant number of legal terms, the prospectus also contains a large number of positive and negative words. Hanley and Hoberg (2012) also find that firm managers who strategically disclose more risk factors in their IPO prospectus have lower probability of being sued in subsequent lawsuits. Firms

---

\(^1\) We thank Feng Li for graciously sharing these readability measures.

\(^2\) The number of observations decreases because Li (2008) requires at least 3,000 words in each filing to construct his measures.
with high litigation risk could also face incentives to reduce their disclosure when a significant amount of judgment is involved (Robinson and Schmidt 2013). Thus, the word categories could overlap, so we finish our tests by including all the linguistic measures together.

We report the results in Internet Appendix Table 4, column 12. We observe the same pattern: even including all additional eleven variables, Use of Negative Words is the only linguistic cue that explains UTB Balance. In the last column 13, we also drop Use of Negative Words to show that our results are not driven by collinearity between negative words and other linguistic measures. Again, none of the other linguistic cues are significantly related to UTB and the signs on these other cues are unchanged.

9. **Robustness Tests on Unrecognized Tax Benefit (UTB) Balance**

First, we re-estimate our baseline regression from column 5 of Table II using censored Tobit regression, which takes into account that UTB Balance is truncated at zero. The results in Internet Appendix Table 8, Panel A are robust. The economic magnitudes estimated under the censored Tobit regression are stronger than the ones reported under the OLS. The numbers in the brackets consecutively refer to the marginal effects estimated under (1) unconditional expected value (UEV), (2) conditional expected value (CEV), and (3) probability uncensored (PU).

In column 1, a one standard deviation increase in Use of Negative Words is associated with an increase in UTB Balance (conditional on observing positive UTB Balance) by 6.7% (5.8%) of the latter’s standard deviation. This also translates into a 4.5% increase in the probability of observing a positive UTB Balance. The economic significance is more salient among the top users of negative words, where in Panel A, column 4, as a one standard deviation increase in Use of Negative Words is 0.213 (t-statistic: 4.29) in column 1. Replacing Use of Negative Words with indicator variables based on Use of Negative Words gives similar results, as the estimated coefficient on quintile score in column 2 remains positive and statistically significant (coefficient: 0.164; t-statistic: 5.05). Columns 3 and 4 further consider the sensitivity with respect to the non-linearity in Use of Negative Words by using quintile indicators. Similar to the above, the systematic patterns appear to be similar to the baseline results reported earlier. Overall, all estimated coefficients of Use of Negative Words (or its variants) are all positive and statistically significant at the 1% level.

---

3 The estimated coefficient on Use of Negative Words is 0.213 (t-statistic: 4.29) in column 1. Replacing Use of Negative Words with indicator variables based on Use of Negative Words gives similar results, as the estimated coefficient on quintile score in column 2 remains positive and statistically significant (coefficient: 0.164; t-statistic: 5.05). Columns 3 and 4 further consider the sensitivity with respect to the non-linearity in Use of Negative Words by using quintile indicators. Similar to the above, the systematic patterns appear to be similar to the baseline results reported earlier. Overall, all estimated coefficients of Use of Negative Words (or its variants) are all positive and statistically significant at the 1% level.

4 UEV reports the marginal effect of a one standard deviation (unit) increase in the independent numeric (dummy) variables. CEV refers to the marginal effect of a standard deviation (unit) increase in the independent numeric (dummy) variables conditional on observing positive UTB Balance. PU reports the probability of observing a positive UTB Balance given a one standard deviation (unit) increase in the independent numeric (dummy) variables. Collectively, these marginal effects in brackets report the decomposition of different marginal effects estimated under the censored Tobit regression.
deviation increase in *Use of Negative Words* is associated with an increase in *UTB Balance* (conditional on observing positive *UTB Balance*) by 20% (17.3%) of the latter’s standard deviation, where the latter translates into a 13.5% increase in the probability of observing a positive *UTB Balance*.

10. **Robustness Tests on All Main Dependent Variables**

We also consider a number of alternative explanations that we position as supplementary tests, as these tests further limit our sample size. Internet Appendix Table 8, Panel B, reports the estimated coefficients for all five dependent variables used in our study, including (1) *UTB Balance*, (2) *Cash ETR*, (3) *GAAP ETR*, (4) *Use of Tax Havens*, and (5) *Number of Tax Havens*. In all cases, our results for *Use of Negatives Words* are not sensitive to including the variables below.

We examine if our results could reflect the heterogeneous sensitivity by firms’ executives to their compensation incentives (Jensen and Murphy 1990; Yermack 1995; Hall and Liebman 1998). We control for CEOs’ *Total Pay-for-Performance Sensitivities (Total PPS)*, which is defined as the natural logarithm of the dollar change in pay for a one-dollar increase in firm value divided by annual compensation (Core and Guay 2002; Edmans, Gabaix, and Landier 2009).

Prior research finds that corporate governance is an important determinant on firms’ tax avoidance behaviors (Desai and Dharmapala 2006; Cheng, Huang, Li, and Stanfield 2012). Thus, we include two measures of corporate governance into our baseline regressions. *Corporate Governance Index* refers to the average Gompers-Ishii-Metrick Index (Gompers, Ishii, and Metrick 2003) to capture the shareholder rights and corporate governance across firms from 1990 to 2006. *Institutional Ownership* refers to the latest quarterly level of institutional ownership from the Thomson-Reuters 13f database prior to fiscal year end. Our results in Tables II, V, and VI are robust to including these two governance measures.

A recent study finds that firms’ tax avoidance activities are associated with the IRS’s tax audit rate (Hoopes, Mescall, and Pittman 2012), where firms are less likely to pursue aggressive tax avoidance strategies when tax audit rate is higher. We construct *IRS Tax Audit Rate*, which refers to the IRS’s tax
audit rates based on the information collected by Transactional Records Access Clearinghouse and represents the proportion of corporate tax audits completed by the IRS in a specific year for firms within a specific range of asset-thresholds. We thus add IRS Tax Audit Rate to our baseline regressions.

We also investigate whether our results reflect firms’ geographical difference in religion-induced norm (Hilary and Hui 2009; Kumar, Page, and Spalt 2011; Boone, Khurana, and Raman 2013). We re-estimate our baseline regression by including Religiosity, defined as the number of religious adherents (as reported by the American Religion Data Archive) to the total population in a county (as reported by the Census Bureau). As the UTB data are from 2007 to 2011 whereas Religiosity data are from 1992 to 2006, the latest values of Religiosity in 2006 are used for the UTB analyses.\(^5\)

One potential concern is that our results reflect industry competition or concentration. To address this concern, we construct Herfindahl-Hirschman Index of net sales based on all firms with the same three-digit SIC in a year following Hou and Robinson (2006). Our results remain robust to including this index.

We then examine if the results are driven by industry classifications. We replace the Fama-French (1997) industry classifications with a finer 3-digit SIC. We also re-estimate our baseline regression using industry median-adjusted dependent variables. Specifically, we subtract the industry median based on the Fama-French industries in the same year. For the indicator dependent variable of Use of Tax Havens, we include additional (Fama-French industry × Fiscal year) indicator variables.\(^6\)

We examine if our results are driven by some small or hard-to-value firms. We exclude small firms with fewer than $50 million of total assets, firms in high technology industries, or young firms. We report robust results Excluding Small-Cap Firms, Excluding High Technology Firms, and Excluding Young Firms.

We also construct different versions of Use of Negative Words based on alternative definitions. First, instead of using 10-K, we construct Use of Negative Words based on McDonald’s count of the

\(^5\) We appreciate Alok Kumar and Oliver Spalt sharing the religiosity data for this robustness check.
\(^6\) In untabulated results, we re-estimate the specifications using industry mean-adjusted dependent variables. Our main results remain robust.
average of a firm’s 10-Q filings in a year. Interestingly, our results for 10-Q Quarterly filings are even stronger than those using 10-K filings. Second, we re-measure Use of Negative Words based on the difference between use of negative and positive words in Negative minus Positive Words. We also construct an alternative version by scaling the difference by the sum of both in Negative minus Positive Words, Divided by the Sum of Both. The results again remain robust and similar.

In all the specifications shown in Internet Appendix Table 8, Panel B, our results for Use of Negative Words are robust, and the signs, statistical significances and economic magnitudes are similar to those in our main tests. We conclude that our strong association between Use of Negative Words and tax aggressiveness measures is highly robust.
## Internet Appendix Table 1

### Determinants of Bank Loan Spreads

The dependent variables are Bank Loan Spread, Commitment Fee, Upfront Fee, Standby Fee, Annual Fee in basis points for each loan in DealScan database. Bank Loan Spread is the total annual spread in basis points paid over LIBOR for each dollar drawn from the loan net of upfront fee. Commitment Fee is the charge in basis points on the unused portion of the loan. Upfront Fee is a one-time charge in basis points collected at the loan initiation. Standby Fee is annual charge in basis points on the letter of credits issued. Annual Fee is the annual fee in basis points charged on the entire commitment loan amount. Number of Failed Banks is the annual number of failed banks in a firm’s headquarter-state obtained from the Federal Deposit Insurance Corporation’s website. Credit Rating is the S&P rating in cardinal scale with AAA=1, AA+=2,…C=21, D=22, and 23 otherwise. Commercial Paper Rating is an indicator that equals one if the borrower has a commercial paper rating. Private is an indicator that equals one if the borrower is not a publicly traded company. Previous Lending Relationship is an indicator that equals one if the borrower borrows from the same lead bank in the past three calendar years. Sale to Close (Ln) is the natural logarithm of the borrower’s sales at the loan initiation. Leverage is the industry median-adjusted book value of debt to total assets. Facility Amount (Ln) is the natural logarithm of the size of the largest facility in a loan package. Maturity is the maturity of the largest facility in a loan package. Number of Facilities is the number of facilities in a loan package. Collateral is an indicator that equals one if the loan is secured with collateral. Financial Covenants is an indicator that equals one if the loan has financial covenants. Performance Pricing is an indicator that equals one if the loan has performance pricing. Lead Bank Allocation is the share of the loan retained by the lead bank at loan origination. Prime Base Rate is an indicator that equals one if the base rate is prime rate. The sample period is from 2000 to 2011. Standard errors are clustered at loan level.

#### Panel A: Baselines

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Bank Loan Spread (1)</th>
<th>Commitment Fee (2)</th>
<th>Upfront Fee (3)</th>
<th>Standby Fee (4)</th>
<th>Annual Fee (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of failed banks</td>
<td>4.004</td>
<td>2.711</td>
<td>5.253</td>
<td>4.208</td>
<td>4.262</td>
</tr>
<tr>
<td>Credit rating</td>
<td>1.311</td>
<td>0.013</td>
<td>2.353</td>
<td>-0.117</td>
<td>1.267</td>
</tr>
<tr>
<td>Commercial paper rating</td>
<td>-57.340</td>
<td>-56.484</td>
<td>-82.752</td>
<td>-48.316</td>
<td>-32.907</td>
</tr>
<tr>
<td>Previous lending relationship</td>
<td>12.457</td>
<td>21.547</td>
<td>43.293</td>
<td>18.151</td>
<td>8.636</td>
</tr>
<tr>
<td>Sale to close (Ln)</td>
<td>-5.774</td>
<td>-3.876</td>
<td>-0.841</td>
<td>-3.281</td>
<td>-3.682</td>
</tr>
<tr>
<td>Leverage</td>
<td>1.87</td>
<td>0.93</td>
<td>1.77</td>
<td>1.38</td>
<td>1.39</td>
</tr>
<tr>
<td>ROA</td>
<td>107.721</td>
<td>107.006</td>
<td>83.547</td>
<td>98.177</td>
<td>64.634</td>
</tr>
<tr>
<td>Facility amount (Ln)</td>
<td>-162.923</td>
<td>-152.656</td>
<td>-101.660</td>
<td>-144.899</td>
<td>-142.634</td>
</tr>
<tr>
<td>Number of facilities</td>
<td>16.739</td>
<td>26.750</td>
<td>17.093</td>
<td>35.292</td>
<td>9.282</td>
</tr>
</tbody>
</table>

### Internet Appendix Table 2

#### Determinants of Bank Loan Spreads

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Bank Loan Spread (1)</th>
<th>Commitment Fee (2)</th>
<th>Upfront Fee (3)</th>
<th>Standby Fee (4)</th>
<th>Annual Fee (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of failed banks</td>
<td>4.004</td>
<td>2.711</td>
<td>5.253</td>
<td>4.208</td>
<td>4.262</td>
</tr>
<tr>
<td>Credit rating</td>
<td>1.311</td>
<td>0.013</td>
<td>2.353</td>
<td>-0.117</td>
<td>1.267</td>
</tr>
<tr>
<td>Commercial paper rating</td>
<td>-57.340</td>
<td>-56.484</td>
<td>-82.752</td>
<td>-48.316</td>
<td>-32.907</td>
</tr>
<tr>
<td>Previous lending relationship</td>
<td>12.457</td>
<td>21.547</td>
<td>43.293</td>
<td>18.151</td>
<td>8.636</td>
</tr>
<tr>
<td>Sale to close (Ln)</td>
<td>-5.774</td>
<td>-3.876</td>
<td>-0.841</td>
<td>-3.281</td>
<td>-3.682</td>
</tr>
<tr>
<td>Leverage</td>
<td>1.87</td>
<td>0.93</td>
<td>1.77</td>
<td>1.38</td>
<td>1.39</td>
</tr>
<tr>
<td>ROA</td>
<td>107.721</td>
<td>107.006</td>
<td>83.547</td>
<td>98.177</td>
<td>64.634</td>
</tr>
<tr>
<td>Facility amount (Ln)</td>
<td>-162.923</td>
<td>-152.656</td>
<td>-101.660</td>
<td>-144.899</td>
<td>-142.634</td>
</tr>
<tr>
<td>Number of facilities</td>
<td>16.739</td>
<td>26.750</td>
<td>17.093</td>
<td>35.292</td>
<td>9.282</td>
</tr>
</tbody>
</table>
### Internet Appendix Table 1 – Continued

**Determinants of Bank Loan Spreads**

#### Panel A: Baselines – Continued

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Dependent Variables (in basis points):</th>
<th>Bank Loan Spread</th>
<th>Commitment Fee</th>
<th>Upfront Fee</th>
<th>Standby Fee</th>
<th>Annual Fee</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Collateral</td>
<td></td>
<td>3.608</td>
<td>6.994</td>
<td>-0.893</td>
<td>8.073</td>
<td>-1.031</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.01)</td>
<td>(1.81)</td>
<td>(-0.11)</td>
<td>(2.20)</td>
<td>(-0.19)</td>
</tr>
<tr>
<td>Financial covenants</td>
<td></td>
<td>8.025</td>
<td>8.772</td>
<td>2.807</td>
<td>14.844</td>
<td>5.335</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.05)</td>
<td>(1.89)</td>
<td>(0.30)</td>
<td>(2.03)</td>
<td>(1.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-9.09)</td>
<td>(-4.72)</td>
<td>(-6.25)</td>
<td>(-4.10)</td>
<td>(-1.54)</td>
</tr>
<tr>
<td>Lead bank allocation</td>
<td></td>
<td>0.052</td>
<td>0.033</td>
<td>0.065</td>
<td>-0.021</td>
<td>0.207</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.97)</td>
<td>(0.56)</td>
<td>(0.59)</td>
<td>(-0.34)</td>
<td>(1.59)</td>
</tr>
<tr>
<td>Prime base rate</td>
<td></td>
<td>3.697</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4.78)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loan type fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Bank fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State fixed effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Number of observations</td>
<td>21,439</td>
<td>11,286</td>
<td>4,163</td>
<td>12,260</td>
<td>5,718</td>
<td></td>
</tr>
</tbody>
</table>

**Adjusted-R²**

|                | 0.539 | 0.467 | 0.612 | 0.513 | 0.665 |

**Statistics of Dep Variables**

| Mean          | 201.9 | 209.0 | 260.8 | 190.1 | 96.4  |
| Median        | 175.0 | 200.0 | 250.0 | 175.0 | 62.5  |

#### Panel B: Baselines with State Fixed Effects

| Number of failed banks | 4.649 | 3.485 | 5.313 | 5.122 | 5.306 |
| Number of observations | 20,790 | 10,938 | 4,058 | 11,913 | 5,622 |

**Controls/Fixed effects**

| State fixed effects | Yes | Yes | Yes | Yes | Yes |

**Adjusted-R²**

|                | 0.543 | 0.471 | 0.618 | 0.523 | 0.681 |
### Other Linguistic Cues, ETRs and Tax Havens

This table reports standardized, pooled OLS regression results, reporting the estimated coefficients on linguistic cues and readability measures. Additional details on all variables are in the Appendix. All estimated coefficients are beta coefficients from standardized regressions, where each variable is scaled to have zero mean and unit standard deviation. Robust standard errors are clustered at firm level, and two-tailed t-statistics are in parentheses.

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Dependent Variables:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cash ETRs</td>
<td>GAAP ETRs</td>
<td>Use of Tax Havens</td>
<td>Number of Tax Havens</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current Year</td>
<td>Next 3-Year</td>
<td>Next 5-Year</td>
<td>Current Year</td>
<td>Next 3-Year</td>
<td>Next 5-Year</td>
<td>Current Year</td>
<td>Next 3-Year</td>
<td>Next 5-Year</td>
<td>Current Year</td>
<td>Next 3-Year</td>
<td>Next 5-Year</td>
<td>Current Year</td>
<td>Next 3-Year</td>
</tr>
<tr>
<td>Use of negative words</td>
<td>-0.051</td>
<td>-0.042</td>
<td>-0.074</td>
<td>-0.074</td>
<td>-0.060</td>
<td>-0.081</td>
<td>0.025</td>
<td>0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.40)</td>
<td>(-1.90)</td>
<td>(-2.72)</td>
<td>(-4.70)</td>
<td>(-2.54)</td>
<td>(-2.92)</td>
<td>(2.45)</td>
<td>(2.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of uncertainty words</td>
<td>-0.012</td>
<td>-0.041</td>
<td>-0.039</td>
<td>0.002</td>
<td>0.009</td>
<td>-0.014</td>
<td>-0.002</td>
<td>-0.030</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.53)</td>
<td>(-1.17)</td>
<td>(-0.96)</td>
<td>(0.10)</td>
<td>(0.25)</td>
<td>(-0.34)</td>
<td>(-0.14)</td>
<td>(-1.75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of possibility words</td>
<td>-0.013</td>
<td>-0.015</td>
<td>-0.029</td>
<td>-0.012</td>
<td>-0.039</td>
<td>-0.017</td>
<td>-0.005</td>
<td>0.015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.71)</td>
<td>(-0.53)</td>
<td>(-0.90)</td>
<td>(-0.59)</td>
<td>(-1.35)</td>
<td>(-0.49)</td>
<td>(-0.42)</td>
<td>(1.07)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of strong words</td>
<td>-0.014</td>
<td>-0.026</td>
<td>-0.021</td>
<td>-0.006</td>
<td>-0.009</td>
<td>0.008</td>
<td>0.007</td>
<td>0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.25)</td>
<td>(-1.51)</td>
<td>(-0.91)</td>
<td>(-0.56)</td>
<td>(-0.51)</td>
<td>(0.39)</td>
<td>(1.00)</td>
<td>(1.44)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of litigious words</td>
<td>0.061</td>
<td>0.037</td>
<td>0.043</td>
<td>0.074</td>
<td>0.059</td>
<td>0.059</td>
<td>-0.008</td>
<td>-0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.80)</td>
<td>(1.13)</td>
<td>(1.11)</td>
<td>(3.31)</td>
<td>(1.80)</td>
<td>(1.54)</td>
<td>(-0.56)</td>
<td>(-1.45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of positive words</td>
<td>0.011</td>
<td>-0.044</td>
<td>-0.049</td>
<td>-0.003</td>
<td>-0.013</td>
<td>-0.005</td>
<td>-0.008</td>
<td>-0.015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.82)</td>
<td>(-2.29)</td>
<td>(-2.17)</td>
<td>(-0.21)</td>
<td>(-0.65)</td>
<td>(-0.23)</td>
<td>(-0.87)</td>
<td>(-1.43)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of constraining words</td>
<td>-0.015</td>
<td>-0.044</td>
<td>-0.047</td>
<td>0.000</td>
<td>-0.025</td>
<td>-0.008</td>
<td>0.010</td>
<td>0.020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.02)</td>
<td>(-2.06)</td>
<td>(-1.53)</td>
<td>(0.01)</td>
<td>(-1.25)</td>
<td>(-0.30)</td>
<td>(1.12)</td>
<td>(1.97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of negation words</td>
<td>-0.006</td>
<td>0.010</td>
<td>0.034</td>
<td>0.011</td>
<td>0.018</td>
<td>0.044</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.41)</td>
<td>(0.41)</td>
<td>(1.24)</td>
<td>(0.71)</td>
<td>(0.72)</td>
<td>(1.54)</td>
<td>(0.06)</td>
<td>(0.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of words</td>
<td>-0.078</td>
<td>-0.075</td>
<td>-0.089</td>
<td>-0.079</td>
<td>-0.053</td>
<td>-0.056</td>
<td>0.013</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.58)</td>
<td>(-2.23)</td>
<td>(-2.08)</td>
<td>(-3.64)</td>
<td>(-1.49)</td>
<td>(-1.29)</td>
<td>(0.81)</td>
<td>(0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gunning fog index</td>
<td>0.035</td>
<td>-0.004</td>
<td>0.006</td>
<td>0.088</td>
<td>-0.001</td>
<td>0.138</td>
<td>0.009</td>
<td>-0.019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.75)</td>
<td>(-0.05)</td>
<td>(0.08)</td>
<td>(1.81)</td>
<td>(-0.02)</td>
<td>(1.86)</td>
<td>(0.29)</td>
<td>(-0.55)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flesch reading ease</td>
<td>-0.015</td>
<td>-0.015</td>
<td>0.002</td>
<td>-0.019</td>
<td>0.007</td>
<td>0.003</td>
<td>-0.004</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.05)</td>
<td>(-0.69)</td>
<td>(0.08)</td>
<td>(-1.24)</td>
<td>(0.33)</td>
<td>(0.12)</td>
<td>(-0.46)</td>
<td>(-0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kincaid readability</td>
<td>-0.056</td>
<td>-0.025</td>
<td>0.021</td>
<td>-0.103</td>
<td>-0.024</td>
<td>-0.164</td>
<td>-0.013</td>
<td>0.029</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.08)</td>
<td>(-0.32)</td>
<td>(0.25)</td>
<td>(-1.94)</td>
<td>(-0.31)</td>
<td>(-1.90)</td>
<td>(-0.38)</td>
<td>(0.78)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Included controls/FEs</td>
<td>A1 of V</td>
<td>A2 of V</td>
<td>A3 of V</td>
<td>A6 of V</td>
<td>A7 of V</td>
<td>A8 of V</td>
<td>2 of VI</td>
<td>4 of VI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>10,729</td>
<td>5,607</td>
<td>3,777</td>
<td>10,729</td>
<td>5,607</td>
<td>3,777</td>
<td>10,729</td>
<td>10,729</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted-R²</td>
<td>0.110</td>
<td>0.133</td>
<td>0.137</td>
<td>0.157</td>
<td>0.132</td>
<td>0.131</td>
<td>0.407</td>
<td>0.737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Internet Appendix Table 3
Financial Constraints and Individual UTB Components

This table reports standardized, pooled OLS regression results, reporting the estimated beta coefficients on the individual components reconciling the beginning and ending UTB Balance. Panel A (B) uses measures in year $t$ ($t+1$) as dependent variables. Columns 1–4 refer to the movements in UTB during the year related to activities in current or prior years. Reduction in UTB due to Tax Settlements refers to the reduction in UTB Balance due to settlements with the tax authorities. Reduction in UTB Balance due to Lapsed Statute of Limitations refers to the reduction in UTB Balance resulting from the lapse of applicable statute of limitations. Amount of UTB Balance Would Affect ETR if Recognized refers to the amount of UTB Balance would affect firms’ effective tax rates if recognized. Penalties and Interest Relating to UTB refers to the amount of accrued penalties and interest related to UTB Balance. All estimated coefficients are beta coefficients from standardized regressions, where each variable is scaled to have zero mean and unit standard deviation. Robust standard errors are clustered at firm level, and two-tailed $t$-statistics are in parentheses.

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Panel A: Baselines</th>
<th>Others (Year $t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Addition in UTB Balance (Year $t$)</td>
<td>Reduction in UTB Balance (Year $t$)</td>
</tr>
<tr>
<td>Use of neg words ($\eta$)</td>
<td>Related to Current Year (Year $t$)</td>
<td>Related to Prior Years (Year $t$)</td>
</tr>
<tr>
<td>(1) Use of neg words ($\eta$)</td>
<td>0.033 (1.76)</td>
<td>0.041 (2.38)</td>
</tr>
</tbody>
</table>

Included controls/FEs

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of obs.</td>
<td>5,076</td>
<td>5,076</td>
</tr>
<tr>
<td>Adjusted-$R^2$</td>
<td>0.141</td>
<td>0.078</td>
</tr>
</tbody>
</table>
### Internet Appendix Table 3 – Continued
Financial Constraints and Individual UTB Components

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Addition in UTB Balance (Year ( t+1 ))</th>
<th>Reduction in UTB Balance (Year ( t+1 ))</th>
<th>Others (Year ( t+1 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Related to Current Year (Year ( t+1 ))</td>
<td>Related to Prior Years (Year ( t+1 ))</td>
<td>Due to Settlements (Year ( t+1 ))</td>
</tr>
<tr>
<td>Use of neg words (( t ))</td>
<td>0.037 (3.24)</td>
<td>0.021 (1.12)</td>
<td>0.010 (1.21)</td>
</tr>
<tr>
<td>Addition in UTB:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Re current year (( t ))</td>
<td>0.503 (16.13)</td>
<td>0.255 (10.70)</td>
<td></td>
</tr>
<tr>
<td>- Re prior years (( t ))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduction in UTB:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Re current year (( t ))</td>
<td></td>
<td></td>
<td>0.197 (5.21)</td>
</tr>
<tr>
<td>- Re prior years (( t ))</td>
<td></td>
<td></td>
<td>0.261 (10.20)</td>
</tr>
<tr>
<td>- Settlements (( t ))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Lapses (( t ))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Amount ETR rec. (( t ))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Penalties &amp; interest (( t ))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Included controls/FEs</td>
<td>Identical to column 5 of Table II with 19 control variables included</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of obs.</td>
<td>4,904</td>
<td>4,904</td>
<td>4,904</td>
</tr>
<tr>
<td>Adjusted-( R^2 )</td>
<td>0.391</td>
<td>0.159</td>
<td>0.040</td>
</tr>
</tbody>
</table>
Internet Appendix Table 4

Other Linguistic Cues and Unrecognized Tax Benefit

This table reports the standardized, pooled OLS regression results. The dependent variable is Unrecognized Tax Benefit. Use of Uncertainty Words, Use of Possibility Words (weak modal), Use of Strong Words (strong modal), Use of Litigious Words, Use of Positive Words, Use of Constraining Words, and Use of Negation Words are the number of category words divided by the total number of words. Number of Words is the natural logarithm of the number of words in firm j’s annual 10-K filing in year t. Use of Negation Words, Gunning Fog Index, Flesch Reading Ease, and Kincaid Readability are all measures based on the readability of firm’s annual 10-K filings from 2007 to 2011. Additional details on all variables are summarized in the Appendix. All estimated coefficients are beta coefficients from standardized regressions, where each variable (except indicator) is scaled to have zero mean and unit standard deviation. Robust standard errors are clustered at firm level. Two-tailed t-statistics are in parentheses.

Independent Variables

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of negative words</td>
<td>0.082</td>
<td>0.084</td>
<td>0.085</td>
<td>0.087</td>
<td>0.084</td>
<td>0.084</td>
<td>0.086</td>
<td>0.099</td>
<td>0.100</td>
<td>0.099</td>
<td>0.099</td>
<td>0.087</td>
<td></td>
</tr>
<tr>
<td>(3.94)</td>
<td>(4.01)</td>
<td>(4.06)</td>
<td>(4.06)</td>
<td>(4.05)</td>
<td>(4.07)</td>
<td>(4.03)</td>
<td>(4.04)</td>
<td>(4.03)</td>
<td>(3.38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of uncertainty words</td>
<td>0.014</td>
<td>0.029</td>
<td>0.067</td>
<td>0.067</td>
<td>0.067</td>
<td>0.067</td>
<td>0.067</td>
<td>0.067</td>
<td>0.067</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.86)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of possibility words</td>
<td>0.004</td>
<td>-0.002</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.20)</td>
<td>(-0.22)</td>
<td>(-0.22)</td>
<td>(-0.22)</td>
<td>(-0.22)</td>
<td>(-0.22)</td>
<td>(-0.22)</td>
<td>(-0.22)</td>
<td>(-0.22)</td>
<td>(-0.22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of strong words</td>
<td>-0.006</td>
<td>0.002</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-0.37)</td>
<td>(0.54)</td>
<td>(0.54)</td>
<td>(0.54)</td>
<td>(0.54)</td>
<td>(0.54)</td>
<td>(0.54)</td>
<td>(0.54)</td>
<td>(0.54)</td>
<td>(0.54)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of litigious words</td>
<td>-0.008</td>
<td>-0.010</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-0.57)</td>
<td>(-1.08)</td>
<td>(-1.08)</td>
<td>(-1.08)</td>
<td>(-1.08)</td>
<td>(-1.08)</td>
<td>(-1.08)</td>
<td>(-1.08)</td>
<td>(-1.08)</td>
<td>(-1.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of positive words</td>
<td>-0.012</td>
<td>-0.004</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-0.70)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of constraining words</td>
<td>0.006</td>
<td>0.021</td>
<td>0.027</td>
<td>0.027</td>
<td>0.027</td>
<td>0.027</td>
<td>0.027</td>
<td>0.027</td>
<td>0.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.35)</td>
<td>(1.24)</td>
<td>(1.24)</td>
<td>(1.24)</td>
<td>(1.24)</td>
<td>(1.24)</td>
<td>(1.24)</td>
<td>(1.24)</td>
<td>(1.24)</td>
<td>(1.24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of negation words</td>
<td>0.003</td>
<td>0.016</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.15)</td>
<td>(0.88)</td>
<td>(0.88)</td>
<td>(0.88)</td>
<td>(0.88)</td>
<td>(0.88)</td>
<td>(0.88)</td>
<td>(0.88)</td>
<td>(0.88)</td>
<td>(0.88)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of words</td>
<td>-0.008</td>
<td>0.022</td>
<td>0.023</td>
<td>0.023</td>
<td>0.023</td>
<td>0.023</td>
<td>0.023</td>
<td>0.023</td>
<td>0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-0.46)</td>
<td>(0.70)</td>
<td>(0.70)</td>
<td>(0.70)</td>
<td>(0.70)</td>
<td>(0.70)</td>
<td>(0.70)</td>
<td>(0.70)</td>
<td>(0.70)</td>
<td>(0.70)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fog index</td>
<td>0.007</td>
<td>0.081</td>
<td>0.105</td>
<td>0.105</td>
<td>0.105</td>
<td>0.105</td>
<td>0.105</td>
<td>0.105</td>
<td>0.105</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.48)</td>
<td>(1.04)</td>
<td>(1.04)</td>
<td>(1.04)</td>
<td>(1.04)</td>
<td>(1.04)</td>
<td>(1.04)</td>
<td>(1.04)</td>
<td>(1.04)</td>
<td>(1.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flesch reading ease</td>
<td>0.012</td>
<td>0.005</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.70)</td>
<td>(0.21)</td>
<td>(0.21)</td>
<td>(0.21)</td>
<td>(0.21)</td>
<td>(0.21)</td>
<td>(0.21)</td>
<td>(0.21)</td>
<td>(0.21)</td>
<td>(0.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kincaid readability</td>
<td>0.004</td>
<td>-0.074</td>
<td>-0.104</td>
<td>-0.104</td>
<td>-0.104</td>
<td>-0.104</td>
<td>-0.104</td>
<td>-0.104</td>
<td>-0.104</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.27)</td>
<td>(-0.71)</td>
<td>(-0.98)</td>
<td>(-0.98)</td>
<td>(-0.98)</td>
<td>(-0.98)</td>
<td>(-0.98)</td>
<td>(-0.98)</td>
<td>(-0.98)</td>
<td>(-0.98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent Variable: Unrecognized Tax Benefit (UTB) Balance

Included controls/FEs: Identical to column 5 of Table II with 19 control variables included

Number of observations: 5,076

Adjusted-R²: 0.179
### Internet Appendix Table 5

**Expected Firm Profitability**

The dependent variable is *Unrecognized Tax Benefit Balance*. *Accruals* refers to a firm’s accruals following Sloan (1996). *Positive Accruals Indicator* equals one when *Accruals* is positive, and zero otherwise. *Negative Accruals Indicator* equals one when *Accruals* is negative, and zero otherwise. *No Dividend Indicator* takes one when a firm declares dividend, and zero otherwise. *Dividend/Book equity* is the dividend divided by book equity following Fama and French (2006). *Asset Growth* is the change in total assets from year \( t-1 \) to year \( t \). *Loss Indicator* equals one when the net income of a firm is negative, and zero otherwise. *Ohlson Bankruptcy Score* is the bankruptcy risk score developed by Ohlson (1980). *Piotroski Firm Strength Score* is a composite measure measuring a firm’s financial strength developed by Piotroski (2000). *Change in Non-Cash Assets* refers to the change in non-cash assets from year \( t-1 \) to year \( t \). Non-cash assets is calculated as total assets minus cash and short-term investments following Dechow, Richardson, and Sloan (2008). *IBES One-Year-Ahead Mean (Median) EPS Consensus* is one-year-ahead earnings per share mean (median) consensus as at a firm’s fiscal year-end. *IBES Two-Year-Ahead Mean (Median) EPS Consensus* is two-year-ahead earnings per share mean (median) consensus as at a firm’s fiscal year-end. All estimated coefficients are beta coefficients from standardized regressions, where each variable (except indicator) is scaled to have zero mean and unit standard deviation. Robust standard errors are clustered at firm level. Two-tailed \( t \)-statistics are in parentheses.

<table>
<thead>
<tr>
<th>Controlling for:</th>
<th>Accruals</th>
<th>Accruals indicator</th>
<th>Negative accruals indicator</th>
<th>No dividend indicator</th>
<th>Dividend/Book equity</th>
<th>Asset growth</th>
<th>Loss indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Use of neg words</strong></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>Use of neg words</td>
<td>0.085</td>
<td>0.085</td>
<td>0.085</td>
<td>0.084</td>
<td>0.085</td>
<td>0.087</td>
<td>0.081</td>
</tr>
<tr>
<td>Num of obs.</td>
<td>5,076</td>
<td>5,076</td>
<td>5,076</td>
<td>5,076</td>
<td>5,075</td>
<td>5,076</td>
<td>5,076</td>
</tr>
<tr>
<td>Adjusted-R²</td>
<td>0.179</td>
<td>0.179</td>
<td>0.179</td>
<td>0.179</td>
<td>0.179</td>
<td>0.180</td>
<td>0.183</td>
</tr>
<tr>
<td>Included controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Identical to column 5 of Table II with 19 control variables included

<table>
<thead>
<tr>
<th>Controlling for:</th>
<th>Ohlson bankruptcy score</th>
<th>Piotroski firm strength score</th>
<th>Change in non-cash assets</th>
<th>IBES one-year-ahead mean EPS consensus</th>
<th>IBES one-year-ahead median EPS consensus</th>
<th>IBES two-year-ahead mean EPS consensus</th>
<th>IBES two-year-ahead median EPS consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Use of neg words</strong></td>
<td>(8)</td>
<td>(9)</td>
<td>(10)</td>
<td>(11)</td>
<td>(12)</td>
<td>(13)</td>
<td>(14)</td>
</tr>
<tr>
<td>Use of neg words</td>
<td>0.085</td>
<td>0.085</td>
<td>0.085</td>
<td>0.077</td>
<td>0.077</td>
<td>0.069</td>
<td>0.069</td>
</tr>
<tr>
<td>Num of obs.</td>
<td>5,076</td>
<td>5,076</td>
<td>5,076</td>
<td>4,053</td>
<td>4,053</td>
<td>3,150</td>
<td>3,150</td>
</tr>
<tr>
<td>Adjusted-R²</td>
<td>0.179</td>
<td>0.179</td>
<td>0.179</td>
<td>0.194</td>
<td>0.194</td>
<td>0.211</td>
<td>0.211</td>
</tr>
<tr>
<td>Included controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Identical to column 5 of Table II with 19 control variables included
Internet Appendix Table 6
Realized Future Firm Profitability

The dependent variable is *Unrecognized Tax Benefit Balance*. All estimated coefficients are beta coefficients from standardized regressions, where each variable (except indicator) is scaled to have zero mean and unit standard deviation. Robust standard errors are clustered at firm level. Two-tailed *t*-statistics are in parentheses.

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of negative words</td>
<td>0.079</td>
<td>0.072</td>
<td>0.074</td>
<td>0.072</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>(3.64)</td>
<td>(3.25)</td>
<td>(2.92)</td>
<td>(2.17)</td>
<td>(2.14)</td>
</tr>
</tbody>
</table>

Controlling for realized future firm profitability:

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return on assets (ROA) t+1</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Return on assets (ROA) t+2</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Return on assets (ROA) t+3</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Return on assets (ROA) t+4</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

| Number of observations | 4,589 | 4,169 | 2,547 | 1,032 | 1,032 |
| Adjusted-\(R^2\)       | 0.181 | 0.175 | 0.170 | 0.177 | 0.178 |
| Included controls       | Identical to column 5 of Table II with 19 control variables included |
Internet Appendix Table 7
Multinationals vs. Domestic Firms

The dependent variable is *Unrecognized Tax Benefit Balance*. *Foreign Income* is a firm’s foreign pre-tax income divided by total sales. *Number of Countries (Tax Havens)* is the natural logarithm of the number of countries (tax havens) reported in Exhibit 21 of a firm’s 10-K filing. Column 4 (5) reports the estimates based on a subsample of multinational (domestic) firms. *Multinational (Domestic)* are firms operating in more than one countries (only operating in the U.S.). All estimated coefficients are beta coefficients from standardized regressions, where each variable is scaled to have zero mean and unit standard deviation. Robust standard errors are clustered at firm level, and two-tailed t-statistics are in parentheses.

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of negative words</td>
<td>0.076</td>
<td>0.046</td>
<td>0.051</td>
<td>0.096</td>
<td>0.021</td>
</tr>
<tr>
<td>× Foreign income</td>
<td>(3.56)</td>
<td>(2.02)</td>
<td>(2.19)</td>
<td>(4.16)</td>
<td>(0.50)</td>
</tr>
<tr>
<td>× Number of countries</td>
<td>0.138</td>
<td>0.148</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>× Number of tax havens</td>
<td>(4.74)</td>
<td>(3.41)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Included controls/FEs</td>
<td>Identical to column 5 of Table II with 19 control variables included</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>5,076</td>
<td>5,076</td>
<td>5,076</td>
<td>4,174</td>
<td>902</td>
</tr>
<tr>
<td>Adjusted-R²</td>
<td>0.171</td>
<td>0.180</td>
<td>0.183</td>
<td>0.184</td>
<td>0.082</td>
</tr>
</tbody>
</table>
Internet Appendix Table 8
Robustness Checks

Panel A reports the standardized estimates from Tobit regressions where each variable (except indicator) is scaled to have zero mean and unit standard deviation. The numbers separated by dashes in brackets represent (1): unconditional expected value, (2) conditional expected value, and (3) probability uncensored values in percentages. Panel B reports standardized, pooled OLS regression results. The estimates in column 4 refer to the marginal probabilities of one standard deviation increase in the independent numeric variable, whereas other columns report the beta coefficients from standardized regression based on various robustness checks. Control for Gompers-Ishii-Metrick Index on Corporate Governance controls for the average Gompers-Ishii-Metrick index on corporate governance from 1990–2006. Control for Total Pay-for-Performance controls for CEOs’ total pay-for-performance sensitivity, defined as the dollar change in pay for a one-dollar increase in firm value, divided by annual pay (Core and Guay 2002 and Edmans, Gabaix, and Landier 2009). Control for Institutional Ownership controls for the latest total level of institutional ownership (in %) prior to fiscal year-end. Control for IRS Tax Audit Rate controls for the IRS face-to-face audit rate of returns filed by corporations from 1994 to 2010, where the data are obtained from Transactional Records Access Clearinghouse. Control for Religiosity controls for religiosity, defined as the number of religious adherents (as reported by the American Religion Data Archive) to the total population in a county (as reported by the Census Bureau), at a firm’s county level. Control for Industry Competition controls for the Herfindahl-Hirschman Index of net sales based on all firms with the same three-digit SIC in a year. Finer Industry Classification uses 3-digit SIC industry classifications. Industry-Adjusted adjusts the dependent variables using the industry-median. For indicator dependent variable, the specification is adjusted by including additional indicators for adjusting industry average. Excluding Small-Cap Firms excludes all firms with fewer than $50 million of total assets. Excluding High Technology Firms excludes all firms in high-technology industries with SIC codes of 357 and 737. Excluding Young Firms excludes all firms with fewer than five years of age since it has first appeared in CRSP. 10-Q Quarterly Filings alternatively constructs Use of Negative Words based on all 10-Q filings in a year. Negative minus Positive Words replaces Use of Negative Words with the difference between Use of Negative Words and Use of Positive Words. Negative minus Positive Words, Divided by the Sum of Both replaces Use of Negative Words with the difference between Use of Negative Words and Use of Positive Words, divided by the sum of two. Robust standard errors are clustered at firm level, and two-tailed t-statistics are reported in parentheses.

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Dependent Variable: Unrecognized Tax Benefit (UTB) Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Use of negative words</td>
<td>0.213</td>
</tr>
<tr>
<td>- Quintile score</td>
<td>(4.29)</td>
</tr>
<tr>
<td>- Top quintile indicator</td>
<td>[6.7-5.8-4.5]</td>
</tr>
<tr>
<td>- 4th quintile indicator</td>
<td></td>
</tr>
<tr>
<td>- 3rd quintile indicator</td>
<td></td>
</tr>
<tr>
<td>- 2nd quintile indicator</td>
<td></td>
</tr>
<tr>
<td>Included controls/FEs</td>
<td>Identical to column 5 of Table II with 19 control variables included</td>
</tr>
<tr>
<td>Num. of observations</td>
<td>5,076</td>
</tr>
<tr>
<td>Adjusted-/Pseudo-R²</td>
<td>0.106</td>
</tr>
</tbody>
</table>
### Internet Appendix Table 8

**Robustness Checks – Continued**

**Panel B: Various Robustness Tests**

<table>
<thead>
<tr>
<th>Descriptions</th>
<th>UTB Balance</th>
<th>Cash ETR</th>
<th>GAAP ETR</th>
<th>Use of Tax Havens</th>
<th>Number of Tax Havens</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Control for total pay-for-performance sensitivities</td>
<td>0.060</td>
<td>-0.068</td>
<td>-0.081</td>
<td>0.018</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>(2.56)</td>
<td>(-4.41)</td>
<td>(-4.97)</td>
<td>(2.62)</td>
<td>(2.10)</td>
</tr>
<tr>
<td>Control for Gompers-Ishii-Metrick Index on corporate governance</td>
<td>0.103</td>
<td>-0.055</td>
<td>-0.070</td>
<td>0.014</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(3.40)</td>
<td>(-3.54)</td>
<td>(-4.21)</td>
<td>(1.87)</td>
<td>(2.01)</td>
</tr>
<tr>
<td>Control for institutional ownership</td>
<td>0.084</td>
<td>-0.061</td>
<td>-0.062</td>
<td>0.025</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(4.05)</td>
<td>(-5.25)</td>
<td>(-5.21)</td>
<td>(3.07)</td>
<td>(2.51)</td>
</tr>
<tr>
<td>Control for IRS tax audit rate</td>
<td>0.082</td>
<td>-0.062</td>
<td>-0.063</td>
<td>0.025</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(3.97)</td>
<td>(-5.30)</td>
<td>(-5.32)</td>
<td>(3.13)</td>
<td>(2.63)</td>
</tr>
<tr>
<td>Control for religiosity</td>
<td>0.100</td>
<td>-0.075</td>
<td>-0.079</td>
<td>0.036</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>(4.26)</td>
<td>(-5.26)</td>
<td>(-5.33)</td>
<td>(3.67)</td>
<td>(3.07)</td>
</tr>
<tr>
<td>Control for industry competition</td>
<td>0.085</td>
<td>-0.061</td>
<td>-0.062</td>
<td>0.025</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(4.07)</td>
<td>(-5.26)</td>
<td>(-3.39)</td>
<td>(3.09)</td>
<td>(2.55)</td>
</tr>
<tr>
<td>Finer industry classification</td>
<td>0.076</td>
<td>-0.052</td>
<td>-0.053</td>
<td>0.016</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(3.43)</td>
<td>(-4.30)</td>
<td>(-4.23)</td>
<td>(1.91)</td>
<td>(2.34)</td>
</tr>
<tr>
<td>Industry median-adjusted</td>
<td>0.091</td>
<td>-0.060</td>
<td>-0.052</td>
<td>0.026</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>(4.45)</td>
<td>(-5.12)</td>
<td>(-4.86)</td>
<td>(3.23)</td>
<td>(3.37)</td>
</tr>
<tr>
<td>Excluding small-cap firms</td>
<td>0.085</td>
<td>-0.062</td>
<td>-0.059</td>
<td>0.024</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(3.97)</td>
<td>(-5.14)</td>
<td>(-4.74)</td>
<td>(3.00)</td>
<td>(2.53)</td>
</tr>
<tr>
<td>Excluding high technology firms</td>
<td>0.086</td>
<td>-0.061</td>
<td>-0.061</td>
<td>0.023</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(4.04)</td>
<td>(-4.85)</td>
<td>(-5.35)</td>
<td>(2.55)</td>
<td>(2.16)</td>
</tr>
<tr>
<td>Excluding young firms</td>
<td>0.094</td>
<td>-0.064</td>
<td>-0.068</td>
<td>0.024</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(3.88)</td>
<td>(-5.08)</td>
<td>(-5.29)</td>
<td>(2.78)</td>
<td>(2.11)</td>
</tr>
<tr>
<td>10-Q Quarterly filings</td>
<td>0.109</td>
<td>-0.066</td>
<td>-0.068</td>
<td>0.025</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>(5.05)</td>
<td>(-5.66)</td>
<td>(-5.49)</td>
<td>(3.02)</td>
<td>(3.16)</td>
</tr>
<tr>
<td>Negative minus positive words</td>
<td>0.083</td>
<td>-0.058</td>
<td>-0.056</td>
<td>0.028</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>(3.96)</td>
<td>(-5.16)</td>
<td>(-4.94)</td>
<td>(3.47)</td>
<td>(3.26)</td>
</tr>
<tr>
<td>Negative minus positive words, divided by the sum of both</td>
<td>0.072</td>
<td>-0.052</td>
<td>-0.047</td>
<td>0.024</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>(3.77)</td>
<td>(-4.79)</td>
<td>(-4.71)</td>
<td>(3.09)</td>
<td>(3.62)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controls/FEs/Constant identical to column of table:</th>
<th>5 of II</th>
<th>A1 of V</th>
<th>A6 of V</th>
<th>2 of VI</th>
<th>4 of VI</th>
</tr>
</thead>
</table>

22