10-K Filing Content and SEO Flotation Costs

Ali Sheikhbahaei
Department of Economics and Finance, Latrobe Business School, Latrobe University, Melbourne, Australia

Amalia Di Iorio
Department of Economics and Finance, Latrobe Business School, Latrobe University, Melbourne, Australia

Balasingham Balachandran
Department of Economics and Finance, Latrobe Business School, Latrobe University, Melbourne, Australia

Huu Nhan Duong
Department of Banking and Finance, Monash Business School, Monash University, Melbourne, Australia

Abstract

This study examines whether firms' 10-K disclosures, especially filing size and sentiment, affect the flotation costs of seasoned equity offerings. We find the flotation costs are higher when there is less information and more negative sentiment in 10-K filings. Using a cross-sectional approach, our study suggests that among a subsample of small firms, underwriters receive higher compensation through a wider gross spread from firms with a smaller volume of information in their 10-K filing. These results are robust given joint consideration of offer size and major information asymmetry components. We also show that a negative tone of a filing increases the probability of an offer withdrawal by underwriters. Overall, this study suggests that the size and sentiment of 10-K filings are important determinants of uncertainty for underwriters.

1.0 Introduction

As publicly available corporate disclosure documents, 10-Ks make a substantial contribution to firm valuation by conveying information about a firm's business cycle and performance (see, for example, Li (2008) and Biddle et al. (2009)). Prior studies indicate that investors under-react to firms that provide more complex and negatively written 10-K reports immediately after filing days (You and Zhang, 2008) and around major corporate events such as an IPO (Loughran and McDonald, 2013). Loughran and McDonald (2014) argue that the size of 10-Ks is a reliable proxy to gauge the text complexity in a financial context and it outperforms other commonly used readability measures such as the Fog-index. Although size-driven complexity makes it difficult for ordinary investors to interpret the 10-K, the question of how financial experts confront this complexity has surprisingly received little attention. Lang and Lundholm (1996) argue that firms with more informative disclosures attract more analysts, while Lehavy et al. (2011) show greater analyst following arises from more complex 10-K reports. However, the question of how underwriters (or investment bankers) interpret 10-K content to fulfill their requirements is still an open question. Particularly, whether underwriters translate 10-K filing size to complexity or in-depth information prior to their commitment to buy equity.

This paper examines the information content of 10-K disclosures from underwriters' perspective in the event of seasoned equity offerings (SEO) where they have no investment incentives but charge underwriting fees. We examine the impact of 10-K textual properties of size and sentiment on underwriters' uncertainty, and in particular, the impact on gross spread and the likelihood of offer withdrawal in the event of an SEO. We seek to provide an insight into (1) whether underwriters ask for less spread as a direct flotation cost when 10-K documents provide more in-depth information, and (2) whether the sentiment of 10-K disclosure affects the decision of the lead underwriter to cancel the equity offering. This study aims to answer

the question of whether the information content of 10-K filings alters information asymmetry between managers and, in particular, underwriters' uncertainty and their charges prior to an SEO.

Flotation costs associated with SEOs represent a significant economic portion of raised capital and are of considerable importance to managers. They incorporate the cost of asymmetric information either between issuers and underwriters (Baron, 1982) and/or between informed and un-informed investors (Rock, 1986). Consequently, reducing the information asymmetry can reduce flotation costs by a significant amount. It is broadly accepted that corporate disclosures can reflect conditions within firms revealed by managers who have better knowledge of their firm's business cycle than outsiders have. During the due diligence investigation ahead of an equity offering, the lead underwriter assesses the value of the issuing firm through available sources of information to state the fairness of the offering price. In their survey, Eckbo et al., (2007) find that underwriters use 10-K filings as a valuable source of information prior to their commitment to purchase the offering stocks, aiming to minimize uncertainty.

We argue that larger 10-K filings can provide greater in-depth information about different aspects of a firm's business cycle. For example, the *Kindered Healthcare INC.* 10-K filing for the year ended December 31, 2013 contains 104 pages of Annual Report and 58 pages of Consolidated Financial Statements whereas *I.D Systems INC.* filed for the financial year ended December 31, 2005 with 36 pages as Annual report and 52 pages as Consolidated Financial Statements. There is an unneglectable difference between the two filings in term of the Annual Report section which potentially could include a significant amount of qualitative information about the past, present, and forward looking business aspects and performance. What we ask is – should the file size be translated to readability or in-depth information about

a firm's business cycle. If the latter holds, then it is consistent with Levahy et al. (2011) who find a larger file size requires more analyst service and leads to less information asymmetry.

Using a sample of 1724 US SEO announcements, consisting of successful and cancelled offers between 1994 to 2014, we find that underwriters receive less compensation when there is a greater volume of information in a firm's 10-K filing. This result holds when controlling for offer size as the predominant factor of gross spread as well as major information asymmetry proxies. Further, our results reveal that there is a higher probability of offer withdrawal by underwriters when the 10-K filing provides a greater volume of information. In further analysis, we address the multi-collinearity issue between firm and offer size by employing principal component analysis in order to cluster the most variation explained by the firm size and information asymmetry measures. Checking the robustness of our estimation, we use two stage least square regression to estimate the gross spread using an endogenous estimation of offer size as suggested by Habib and Ljungqvist (2001). Our results show a consistent coefficient sign with statistically significance t-statistics of our main independent variable of file size while treating it together with principal components of information asymmetry and offers size. These results suggest that while offer size is a predominant determinant of gross spread, larger 10-K filings reduce the level of information asymmetry and lead to lower gross spread.

This study extends prior research on information content of the 10-K filing beyond investor (market) reaction. Prior studies show a strong link between textual properties (tone sentiment and readability) on filing day returns and volatility [see, for example, Jegadesh and Wu (2012) who study the effect of 10-K positive and negative sentiment on filing day returns; and Loughran and McDonald (2014) who look at the 10-K readability and subsequent return volatility]. These studies investigate the informativeness of 10-Ks from investors' perspective. To our knowledge, there has yet been no endeavor to investigate the impact of 10-K textual properties on financial experts' decision-making and uncertainty. We fill this gap by

investigating changes in volume of information within 10-K annual filing, proxied by file size, and their association to changes in the level of information asymmetry between underwriter(s) and managers. Contrary to Loughran and McDonald (2014), we suggest file size should not be interpreted as a readability measure when the readers (users) are financial experts, specifically underwriters, but should be considered in-depth information. Further, we investigate the relationship of 10-K tone with the probability of offer cancellation by the lead underwriter.

We also extend prior findings on information asymmetry and SEO flotation costs employing a new measure of information asymmetry. Since it is not directly observable, it is difficult to agree on the best measure of information asymmetry. In any case, existing different measures claim to have marginal effects on investors' decisions as well as investment banks' underwriting service charges and their compensation. Developing a new measure of information asymmetry, Lee and Masulis (2009) incorporate idiosyncratic risk into accruals quality so as to assess the quality of accruals and earning figures available to underwriters and outside investors. They find poor accounting quality is associated with higher flotation costs, including underwriters' gross spread and investors' reaction to offering stocks. We extend their work by assessing the in-depth information and sentiment of annual 10-K filings as an important source of corporate information to underwriters. Considering offer size as a predominant determinant of gross spread, and controlling for major information asymmetry measures, we examine the impact of size and sentiment of 10-Ks on direct and indirect flotation costs imposed by underwriters in a firm commitment SEO. We control for cross-sectional measures of asymmetric information between managers and outsiders, specifically underwriters, in order to observe the variation of uncertainty about the equity and the issuing firm, and respective flotation costs. We extend the work by Lee and Masulis (2009) on the information environment and gross spread by employing 10-K properties and applying econometric methods to prevent bias estimation of gross spread, and we address the

multicollinearity issue between firm size, offer size, and gross spread. Overall, we provide empirical evidence to show the importance of in-depth information of the business cycle to underwriters.

The reminder of this paper is structured as follow: Section 2 presents a literature review; Section 3 discusses the data, sample characteristics, and research design; Section 4 presents our empirical findings and reports on a robustness analysis; and Section 5 concludes the paper.

2.0 Literature Review and Hypothesis Development

In all equity offerings, firms seek to maximize the net proceeds by minimizing all direct and indirect flotation costs. According to a survey by Eckbo et al. (2007), direct costs include underwriting fees (in either cash, or warrants and options), underwriters' compensation, as well as other fees and out-of-pocket expenses, such as accountant and registration fees[1]. Indirect costs include announcement effects and issue delays or withdrawals. They find that underwriters receive compensation in several ways, but most importantly by paying a lower purchase price than the public offering price. The difference of the two prices is called underwriters gross spread which varies across stock and firm levels. In other words, the spread is paid to underwriters as a percentage of raised capital for selling the firm's issued stock.

In a typical firm commitment SEO, the most important underwriter's compensation is gross spread. Prior to each firm commitment SEO, the issuer contacts an investment banker to form and lead a syndicate to commence a process of due diligence. This involves appointing a lead underwriter in order to investigate the financial status of the company and confirm its registration statement. Once the due diligence is completed, the issuer negotiates an offer price

[1] Direct cost also includes fees to law firms, listing fees, printing, advertising, road show expenses, and management time. (Eckbo et al., 2007).

with the syndicate and the offer begins the following day (Eckbo et al., 2007). Meanwhile, in a firm commitment, all offering stock is bought by the underwriter from the issuer, therefore they seek reliable information to avoid putting their money at risk. In these circumstances, underwriters would be compensated for the exposure to greater risk by a higher gross spread so as to be encouraged to purchase the stock. It is rational to believe the greater the uncertainty that arises from the higher level of information asymmetry between insiders and outsiders, the higher expected gross spread. In a traditional firm commitment, a syndicate of investment bankers has sufficient time to assess the market demand and adjust their fees according to the respective risk. However, if the syndicate realizes poor or low demand for the offering stock, they can delay or even cancel the offer which bears an indirect costs with no proceeds at all. Notwithstanding this, expected flotation costs can vary in both cross-sectional and time-series analysis depending on either firm or security characteristics.

Lee and Masulis (2009) assert that measurable proxies such as analyst forecast dispersion (Marquardt and Wiedman, 1998), stock liquidity and stock return volatility (Corwin, 2003) and debt rating (Liu and Malatesta, 2006) cannot exogenously capture asymmetric information as they are influenced by other economic factors.

A consistent finding in SEO literature conveys that pre-offer risk as measured by stock return standard deviation, as well as offer size (inverse or in logarithm), can predominantly explain the gross spread in SEOs (see Denis, 1994; Altınkılıç and Hansen, 2000; Lee and Masulis, 2006). They show that gross spread is larger for firms with higher levels of systematic risk. More specifically, Altinkilic and Hansen (2000) reveal that lower quality offerings require higher spreads. They posit that a larger spread exists when the stock is more volatile and more underwriting services are required to execute greater marketing efforts.

While prior studies treat offer size (log of net proceed) as an exogenous variable, Lee and Masulis (2009) claim that it should be treated as an endogenous variable, following Habib and Ljungqvist (2001), since both gross spread and offer size can be affected by information asymmetry between managers and underwriters. Offers with larger proceeds are related to lower underwriter fees and compensation. Burch et al. (2005) find a significant association between the logarithm of gross proceeds and gross underwriting spread (lead management fees, management underwriting fees, and selling concession) using the producer price index[2]. Their results are consistent with Smith (1977) from the perspective of economies of scale, indicating larger offers require a lower underwriter spread. In addition to this view and consistent with Altinkilic and Hansen (2000), we also control for offer quality with other proxies such as analyst following and accrual quality.

Another major influential factor on gross spread that prior studies employ as a controlling variable is firm size, commonly measured by total assets (See Butler et. al., 2005; Lee and Masulis, 2006, 2009). Firm size is negatively related to gross spread because it can reflect asset diversification and the quality of information available to public (Eckbo et. al, 2007). We extend this discussion and argue that greater volumes of information can intuitively increase the quality of the information about different aspects of their business cycle and will reduce information asymmetry.

Corporate disclosures are valid channels through which financial (and other relevant) information is communicated between managers and outsiders, and they include both quantitative and qualitative information about firm performance, business cycle, and market pricing. Amongst all corporate disclosures, annual 10-K filings deliver primary textual properties of readability and sentiment to investors, creditors, bankers and analysts. Given the

[2] Converting gross proceed dollars of each offer and other financial variables to the benchmark time (January 2001).

complexity of disclosures for ordinary investors, prior studies demonstrate the relationship between readability and investors decision. For example, Lawrence (2013) investigates data from individual investors and reports that firms with more readable 10-K filings attract more investors. Li (2008) shows firms with less readable 10-Ks have lower earnings and those with easy to read filings have persistent positive earnings. All the above studies use the Fog index as a proxy to measure the readability of disclosures. However, Loughran and McDonald (2014) assert that the size of 10-K filing outperforms Fog index, and is correlated with other commonly used readability measures. Overall, prior studies describe how readability can affect decision-making in financial markets from the investor's point of view.

Other studies indicate the importance of a firm's in-depth disclosed information from financial experts' perspective where analysts are the users of such disclosures. Lang and Lundholm (1996) examine the informativeness of firms' disclosures and show that firms with more informative disclosures are followed by more analysts. They use scores[3] calculated from summarized evaluations as a proxy to capture the informative content. Notwithstanding these findings, their measure cannot endogenously capture a firm's in-depth information. In another similar study, Healy et al. (1999) use analysts' "*rating of firms*" and track their improvement over time to capture firms' disclosure quality. Their findings suggest that an increase in rating is followed by an increase in a firm's stock return and stock liquidity. More recently, Lehavy et al. (2011) find that analyst following and the informativeness of their report is greater for firms with less readable 10-K reports. They conclude that the demand for analysts' services increases for firms with more complex disclosures.

[3] They use FAF Report (1985-89) to measure the informativeness of a firm's disclosure. In the FAF report, analysts evaluate the complete range of a firm's disclosures, summarizing their evaluations by a score in each of three categories: annual published information, other published information (including quarterly filings, press releases and proxy statements) and investor relations.

Yet, the question of how readability can affect underwriters' concerns about the quality of a firm's offer remains unanswered. In this study, we shed light on filing size and its informativeness for underwriters.

H1: Larger 10-K filings can reduce the uncertainty about the issue quality and are associated with lower gross spreads.

If we accept this hypothesis, then we are consistent with the findings of Lehavy et al. (2011) that larger file size require more analyst service, which leads to less information asymmetry.

Measuring the sentiment and tone of corporate 10-K reports is also widely employed in prior literature through different methodologies. The main concern about the accuracy of measuring the tone in financial text refers to which dictionary is used in order to gauge the words' tone within sources. Prior studies commonly use the Diction platform to parse financial text and further assign sentiment weight to each word (for example: Rogers et al. (2011) tabulates parsed words by Diction into 3 categories of praise, satisfaction, and inspiration).

Loughran and McDonald (2011) claim that Diction optimistic and pessimistic word list is problematic when gauging financial text. They create a positive and negative dictionary specifically tailored for financial text and they find significant deviation between the Diction word list and the Loughran and McDonald (LM) word list. Further, they find a relationship between sentiment measures using the LM word list and market reaction around 10-K filing days, and find evidence that a negative list is associated with short-term returns around filing days. Loughran and McDonald (2013) employed the LM word list to investigate IPO offer day returns and subsequent stock volatility through S-1 mandatory pre-offer filing with SEC. They demonstrate a positive relationship between ex-ante uncertainties arising from S-1 filing tone and first day returns, as well as aftermarket volatility. However there is no prior study that examines the effect of filing sentiment on underwriters' decision to proceed or cancel the offer.

In our study we investigate the relationship between the sentiment of 10-Ks and the level of uncertainty around the offering from the underwriter's perspective.

H2: The negative tone of 10-K filing is associated to the probability of an offer withdrawal by underwriters.

Furthermore, prior studies highlight the following issues and firm characteristics as additional important controlling variables to estimate SEO flotation costs:

Leverage: Lee and Masulis (2009) argue a higher leverage ratio is associated with higher risk therefore underwriters avoid taking high risk unless they receive reasonable compensation. They show that higher leverage is significantly related to higher gross spread in higher information asymmetry circumstances.

Tobin Q: Market value of assets divided by total asset indicates the investment opportunity associated with a firm, and tends to be higher for growing firms. Consistent with expectations, Lee and Masulis (2009) find that Tobin Q is negatively related to gross spread and positively related to announcement returns. The higher the Tobin Q ratio, the more attractive to underwriters the investment is, and therefore they will expect less fees and compensation.

Accrual Quality: Francis et al. (2005) find a positive relationship between abnormal accruals and the cost of equity and debt. This anomaly decreases the quality of accruals and increases the degree of information asymmetry between managers and outside investors. Since managers use accruals as a source of earning management, outsiders may be concerned about the quality of the earning accruals as well as changes in operational or environmental factors during a given period of time. Lee and Masulis (2009) integrate firm fixed effect to estimate total current accruals controlling for time variation of firm characteristics. Unsurprisingly, they find that poor accrual quality is associated with high flotation costs.

Amihud Illiquidity: Corwin (2003) argues that a higher level of information asymmetry is associated with a larger bid-ask spread and uninformed traders can be compensated for bearing more risk to trade against informed investors. Uncertainty arises from information asymmetry and is positively associated with stock illiquidity which is defined as the average ratio of the daily absolute return to the (dollar) trading volume on that day (Amihud, 2002). We observe average daily illiquidity in a period of 50 trading days prior to equity offering. We expect to capture higher flotation costs when there is more illiquidity in a firm's stock arising from higher information asymmetry.

Analyst Following: Hansen and Torregrosa (1992) develop a theoretical model that predicts less underwriters' compensation when there is more capital market monitoring. Hong et al. (2000) and Brown et al. (2009) find empirical evidence that a greater analyst following is associated with reductions in the level of information asymmetry. We expect firms with greater analyst coverage experience less flotation costs because those analysts provide a better information environment for investors.

Underwriter rank: the rationale behind underwriter reputation is that higher ranked underwriters provide services to firms with less uncertainty and thereby they ask for less compensation. Li and Masulis (2007) find empirical evidence that supports this notion in a competitive underwriting market.

Secondary: occurs when a portion of offered stock is sold by existing shareholders rather than the issuing firm. Karpoff et al. (2012) argue that underwriters' guarantee the offered stock if they are confident that the price is fair after completing due diligence. Therefore, investors also are assured that they will not overpay insiders who have an information advantage. We expect less flotation costs attached to offers with a secondary component.

Shelf: Rule 415 shelf registration increases the competition between underwriters since an issuer can offer its securities at any point of time from the registration date through the next 3 years. Empirical results also show lowered underwriting fees since the implementation of shelf registration (Autore et. al, 2008). We expect shelf offers are related to a lower gross spread.

3.0 Data and Sample

Our initial sample includes 6299 US firm commitment offerings acquired from Security Data Company (SDC Platinum) from 1994 to 2014. Following prior studies on SEOs, sample selection criteria consists on common stocks in the U.S market listed on NASDAQ, NYSE, and AMEX excluding penny stocks (less than a dollar) with Book-Building pricing techniques. Further, we exclude offers with simultaneous corporate announcements and those with Merger and Acquisition announcements within a 6 month period prior to SEO announcement. Combined offers of Units and Warrants, Real Estate Investment Trusts (REITS) firms, closed-end financial firms were also excluded. Further, we merge our sample with Compustat and CRSP database to obtain financial statements, price, and return variables. We also merge our final sample with the WRDS SEC Analytics Suit database to get the 10-K filing Sentiment and Readability data. The number of analysts providing firm earnings forecasts in the last quarter before the SEO is acquired from the I/B/E/S database. We collect underwriter rankings from Jay. R Ritter's blog on the Warrington College of Business website[4]. All of the above

[4] IPO Underwriter Reputation Rankings (1980 – 2015) Appendix C of "Why Has IPO Underpricing Changed Over Time?" (Co-authored with Tim Loughran).

matchings and data availability lead result in a final sample of 1752 offers, including 1664 successful and 88 canceled offers.

A large body of literature raises doubt about the accuracy of SEO offering data in the SDC database (see, for example, Lease et al. (1991); Safieddine and Wilhem (1996)). They posit that if the offer is announced and executed after the official close of the stock market, the effective and real announcement date should be the next trading day. In order to address this issue, we follow Safieddine and Wilhem (1996) and Corwin (2003) to capture the actual effective issue date using the trading volume approach.

Table 1, Panel A shows the distribution of successful offers across the period. Generally, the number of offers increases over time except for a significant drop in 2008 which could be due to the Global Financial Crisis. Panel B shows the distribution of offers by individual firms and indicates that more than half of our sample (63.7%) are single SEOs by unique firms during the period. Panel C shows the frequency distribution of offer type, including pure primary and mixed offers, with a component of secondary shares being sold by current shareholders. On average, pure primary offers outnumber those with a 4.0 secondary component by more than 3.4 times (77.16% of sample consists of pure primary offers).

4.0 Results

4.1 Summary statistics:

Table 2 presents the sample summary statistics. In Panel A, we show descriptive statistics of the dependent variables employed in this study across the entire sample and in two sub-samples of lowest and highest quartiles of 10-K file size. The test of difference shows a statistically significant higher average gross spread for firms in the lowest 25% of file size. This result suggests that firms within the smallest filing size appear to pay significantly higher

compensation to underwriters. The results also show that there is a higher probability of offer withdrawal for firms in the lowest quantile of file size. The two primary results motivate us to examine the effect of filing size on gross spread cross-sectionally and control for important firm and issue variables.

Panel B describes the characteristics of the full sample and in the highest and lowest quartiles of 10-K filing size available on the SEC Edgar database. Firm and offer characteristics consist of controlling variables being used in previous studies to control for changes in flotation costs that vary across our sample. We test for the difference between firm characteristics depending on file size to investigate whether the relationship between file size and flotation costs is affected by other controlling variables. Our test of differences shows firms in the lowest quartile of file size raise significantly larger amounts of capital scaled to total assets. We also find that firms in largest quantile have statistically significant greater total assets and are leveraged more. Finally, firms in lowest quantile of file size have significantly higher ranked underwriters.

Overall, the results presented in both Panel A and Panel B motivate us to draw a hypothetical association between 10-K file size and each component of flotation costs (direct and indirect), controlling for other firm specific characteristics and market microstructural drivers.

4.2 Regression Results

4.2.1 Firms Size, Offer Proceed, and Gross Spread

We expect offer size increases with the increase in a firm's size measured by total assets. To graphically illustrate the pattern of SEO size and firm size, we plot scatter diagrams of the logarithm of total assets by their respective logarithm of net proceeds, which is scaled by total assets, in order to find the percentage of total assets to be covered by net proceeds. Prior studies

have demonstrated a negative association between offer size and gross spread. Figure 1 shows the negative association in our sample. Figure 2 shows a strong positive relationship between total assets and offer size (net proceeds) which signals a multicollinearity issue if we use them simultaneously in our regression models. We address this issue by first running separate models employing these two important explanatory variables (offer size and firm size). Next, we employ principal component analysis to better address the multicollinearity issue between our explanatory variables. Lastly, we use instrumental variables to run two stage least square regression in order to tackle limitations arising from ordinary least square regression.

Two different strands of literature exist for the treatment of net proceeds as an explanatory variable. Many studies employ offer proceeds as an exogenous variable whereas some studies claim that net proceeds is also affected by information asymmetry and should therefore be treated as an endogenous variable. As shown in Figure. 1 there is a negative correlation with an obvious pattern between the logarithm of net proceeds and gross spread which raises the question of whether endogeneity is an issue and, if it is, then how does it affect the robustness of the outcome of our hypothesis test. We will address this potential endogeneity issue in the last section of this study.

Table 3 shows the correlation coefficients between all variables used in this study. As discussed earlier, there is a strong positive correlation of 75 percent between net proceeds and total assets which motivates us to address the multicollinearity issue by using two separate approaches: (i) putting these two important explanatory variables into two separate regression models, and (ii) capturing the most variation of these variables by creating principal components of each. There is also a strong negative correlation between total assets and our main dependent variable of gross spread by -65 percent. We expect that this high level of correlation would affect the explanatory power of our main independent variable of file size.

Table 4 shows the results of our estimations for gross proceeds (offer size) by our main variables, file size and sentiments, as well as other controlling variables which prior studies highlight as important determinants of gross spread.

4.2.2 Gross spread and the size of 10-K filing

We develop several empirical models for each stage of our main analysis. In order to estimate gross spread as the first and the most important component of flotation costs, Table 5 represents OLS White heteroscedasticity robust standard error models. We control for this important issue and firm characteristics to find a robust estimation of gross spread by the size of 10-K filing. We also include highest and lowest quartile indicators of file size to capture the non-linear relationship in a separate model. We proxy the level of asymmetric information by measuring the volume of information within the 10-K filing. We argue that more information disclosed by a firm would lower the level of asymmetric information and therefore lower the gross spread. These results motivate us to undertake further investigation through indicator variables of the highest and lowest 20% quartiles of firm size and offer size in order to capture a probable existing non-linear relationship between gross spread and each variable.

Models 1 and 2 show that file size is significantly negatively associated with gross spread either with or without controlling for other firm and offer characteristics. Consistent with these findings, in Models 3 and 4 we find that the highest and lowest quartiles of file size are negatively and positively associated to gross spread respectively, which captures a potential non-linear association.

As a first approach to tackling the multi-collinearity issue, we run separate models to capture the effects of firm size and offer size including their interaction with our main dependent variable of file size. The following regressions exploit the cross-sectional changes

in gross spread controlling for firm and offer size separately. Equation (i) estimates the gross spread through cross sectional changes in file size and firm size. The variable of main interest is the interaction between file size and total asset, Log (File-Size) × Log (Total Asset). It estimates the effect of file size through different levels of firm size on gross spread.

$$\text{Gross spread} = \alpha_0 + \alpha_1 \text{Log(File-Size)} + \alpha_2 \text{Log(Total Asset)} + \alpha_3 \text{Log(File-Size)} \times \text{Log(Total Asset)} + \alpha_4 \text{Leverage} + \alpha_5 \text{Tobin q} + \alpha_6 \text{Accrual Quality} + \alpha_7 \text{Amihud Illiquidity} + \alpha_8 \text{Return volatility} + \alpha_9 \text{Underwriters Rank} + \alpha_{10} \text{Secondary\_dummy} + \alpha_{11} \text{Shelf\_dummy} + + \alpha_{12} \text{Hightech\_dummy} + \alpha_{13} \text{Analyst following} + \alpha_{14} \text{NYSE\_dummy} + \text{Year Dummy} + \text{Industry Dummy} + \varepsilon_i \quad \text{(i)}$$

Further, Equation (ii) estimates the gross spread through cross sectional changes in file size and offer size. Similar to Equation (i), we are mostly interested to find the effect of file size on gross spread through different values of offer size, β 3 Log(File-Size) × Log(net_proceed).

$$\text{Gross spread} = \beta_0 + \beta_1 \text{Log(File-Size)} + \beta_2 \text{Log(net\_proceed)} + \beta_3 \text{Log(File-Size)} \times \text{Log(net\_proceed)} + \beta_4 \text{Leverage} + \beta_5 \text{Tobin q} + \beta_6 \text{Accrual Quality} + \beta_7 \text{Amihud Illiquidity} + \beta_8 \text{Return volatility} + \beta_9 \text{Underwriters Rank} + \beta_{10} \text{Secondary\_dummy} + \beta_{11} \text{Shelf\_dummy} + + \beta_{12} \text{Hightech\_dummy} + \beta_{13} \text{Analyst following} + \beta_{14} \text{NYSE\_dummy} + \text{Year Dummy} + \text{Industry Dummy} + \varepsilon_{ii} \quad \text{(ii)}$$

Table 5 shows the results of the separate estimations above. Models 5 and 6 control for total assets and other explanatory and fixed effects, and Models 7 and 8 control for net proceeds and the remainder of the controlling and fixed effect variables.

Consistent with our expectations, the strong correlation between total assets and gross spread eliminates the relationship between file size and gross spread in Model 5. However, the effect of file size on gross spread appears to be through its interaction with total assets as shown in Model 6. In order to address this issue, we introduce two indicator variables representing the highest and the lowest quartiles of total assets in Table 6. File size also has a significant

negative association with gross spread when controlling for net proceeds and other explanatory variables.

Panel A of Table 6 shows that firms in the lowest 20% quartile of total assets experience higher gross spread, but the effect of file size is captured through its interaction with the highest quartile of total assets. This suggests that on average small size firms have higher gross spread, while firms with a larger file size which belong in the highest quartile of total assets, are significantly related to a lower gross spread. Panel B, shows the same justification applies to net proceeds. Since the correlation between net proceeds and gross spread is moderate, we expect to capture both predictive powers of file size and net proceeds simultaneously. The results show that file size is significantly negatively related to gross spread and also significantly negatively affects the gross spread through its interaction with the highest quartile of net proceeds. In other words, firms with more information content in their 10-K file that receive higher net proceeds pay less compensation to underwriters through gross spread.

Overall, our results on the gross spread as a significant portion of flotation costs suggest that firms with a larger amount of information in their annual 10-K filing will pay less compensation to underwriters in term of gross spread if they belong to the highest quartile of total assets and are among the highest quartile of net proceeds. However, we need to further investigate these findings by treating net proceeds as an endogenous variable. We will address this issue in the last section of our analysis.

4.2.3 Sentiment of 10-K and probability of offer cancellation

In this section, we investigate the probability of SEO withdrawal by measuring sentiment of 10-K filings. As discussed earlier, sentiment of public disclosure is an important influential factor in the decision-making of outsiders or external parties. Underwriters are among such

'outsiders' and use 10-K reports when they undertake due diligence in the process of book building the offer and evaluating the true value of the offering firm. We hypothesize that a more negatively written 10-K is associated with a higher probability of offer withdrawal. To test this hypothesis, we develop the following model with three major sentiment measures of Negative, Positive, and Uncertain word percentages introduced by LM, while controlling for all other important explanatory variables and fixed effects of year and industry:

$$\begin{aligned}\text{Cancell\_dummy} = {} & \gamma_0 + \gamma_1 \text{Positive(\%)} + \gamma_2 \text{ Negative(\%)} + \gamma_3 \text{ Uncertain(\%)} + \gamma_4 \text{ Leverage} + \gamma_5 \text{ Tobin q} + \gamma_6 \text{ Accrual Quality} + \gamma_7 \text{ Amihud Illiquidity} + \gamma_8 \text{ Return volatility} + \gamma_9 \text{ Underwriters Rank} + \gamma_{10} + \text{Secondary\_dummy} + \gamma_{11} \text{ Shelf\_dummy} + + \gamma_{12} \text{ Hightech\_dummy} + \gamma_{13} \text{Analyst following} + \gamma_{14} \text{ NYSE\_dummy} + \text{Year Dummy} + \text{Industry Dummy} + \varepsilon_{iii} \qquad \text{(iii)}\end{aligned}$$

Table 7 presents the results of the probit regression with the indicator of cancelled offer as the dependent variable. The results show that the percentage of negative words in a 10-K report has a significant relationship to the probability of an offer cancellation. It is also economically significant where a one-standard deviation increase in percentage of negative words increases the probability of offer cancellation by approximately 2%. This estimate appears to be statistically significantly in the case of firms with lower liquidity in 50 trading days pre-offer period, primary offerings, and non-shelf registered firms. However, positive and uncertain sentiments do not have a significant effect on the probability of an offer cancellation and marginal effects of -0.9% and 0.7% respectively.

5.0 Robustness Analysis

A multicollinearity issue arises from a strong correlation between totals assets and offer size, and poses an attenuation bias estimation of coefficients in the OLS model (see, for example, Lubotsky and Wittenberg, 2006; Boone et al, 2007). A simple approach to address

this issue, as shown in Table 5 and Table 6, is to treat the two highly correlated variables using two separate models. However, in order to show economic causality we need to control all firm and issue characteristics simultaneously (Maddala, 1983). Following Karpoff et al (2013), we construct a single variable of information asymmetry from the eight variables that do not exhibit multicollinearity, addressing the bias estimation of coefficients. We employ principal component analysis to aggregate the variances of those eight variables as shown in Table 8. We use the components which drive the most explanatory power in the rest of the analysis as a robustness check. Further, following Lee and Masulis (2009), we use 2SLS analysis controlling for principal components to address the endogeneity issue of estimating gross spread with offer size.

Table 8 shows principal component statistics using the eight individual information asymmetry measures we employ in our study. Out of all eight components constructed from eight primitive measures, three components have Eigenvalues of greater than one - 2.26, 1.21, and 1.09 respectively (as shown in Appendix B), and we observe statistically significant variation of our eight primitive measures of information asymmetry. The overall KMO rate is also over 0.5 which allows us to apply the PCA econometric method in our study. Component 1 captures significant (above 0.3) variation loadings of total assets, illiquidity, number of analysts following, return volatility, and NYSE indicator. The correlation sign is consistent with predictions though all show a reverse sign. Therefore, we multiply this component by (-1) and use it in our regression models as they show information symmetry (Karpoff et al, 2013). Components 1 and 2 also show acceptable Eigenvalues (very close to 1), however they are relatively less reliable components compared to first component. Theoretically, we can use all three components in our analysis as our eight measures of information asymmetry are clustered between our three components and all Eigenvalues are greater than 1. Component 2 captures

variance loading of accrual quality, return volatility and shelf indicator, whereas component 3 captures variance loading of accrual quality, illiquidity, and underwriter rank.

Table 9 presents our robust estimation results of gross spread using principal components. We run two stage least square analysis by estimating offer size using two instrumental variables. In the first stage, we use *secondary indicator* and *non-High tech indicator*. In the second stage, we use predicted value of offer proceeds as a controlling variable. In this stage, unlike our initial results from Tables 5 and 6, we replace individual information asymmetry measures with principal components of information asymmetry. In order to show the bias OLS estimation, Table 9 reports both OLS and 2SLS next to each other. Model 1 shows that, multicollinearity between offer proceeds and component 1 (which represents variance loading of total asset), leads us to a bias coefficient estimation of file size. However, Model 2 shows that if we treat offer size as an endogenous determinant of gross spread, we have an unbiased estimation of both coefficients of file size and component 1.

Models 3 and 4 show the results of estimating gross spread with file size and the two principal components respectively. PC2 which captures the most variations of accrual quality, return volatility, NYSE indicator, and Shelf indicator does not show statistically significant association to gross spread. However, Model 4 indicates that PC3 as a principal component of accrual quality, illiquidity and underwriter rank shows significant predictive power. Finally, we put all the principal components and other controlling variables in our 2SLS estimation. The results in Model 5 indicates that file firms with larger file size, and lower information asymmetry which is captured by PC1 and PC3, pay less compensation to underwriters through gross spread.

6.0 Conclusion

This paper assesses how size and sentiment of 10-K filings affects underwriters' spread and their uncertainty about the offering. A survey by Eckbo et al. (2007) reveals that underwriters use 10-K filings as a valuable source of information to assess the quality of an offer. Loughran and McDonald (2014) suggest that 10-K file size is a better proxy to determine complexity for investors rather than other commonly used readability measures. The fundamental question this study investigates is whether the 10-K file size triggers complexity or in-depth information for underwriters. Initially, if in-depth information of 10-Ks decreases the level of uncertainty around the offer, underwriters receive a lower spread and also there is a less likelihood that they cancel the offer.

Table 1
Offers Characteristics. The sample requires the firm to undergo a firm commitment offering with book-build offering technique being listed in Nasdaq, Amex , or NYSE with offers greater than $1 offering price per share, also excludes REITS, ADRs, close-end funds, simultaneous offers of other securities and offers with M&A announcement in prior 6 month period.

Panel A	**Distribution across years**		
Year	Frequency	Percent	Cumulative frequency
1994	17	1.02	1.02
1995	49	2.94	3.97
1996	61	3.67	7.63
1997	70	4.21	11.84
1998	60	3.61	15.44
1999	74	4.45	19.89
2000	76	4.57	24.46
2001	67	4.03	28.49
2002	76	4.57	33.05
2003	87	5.23	38.28
2004	117	7.03	45.31
2005	86	5.17	50.48
2006	71	4.27	54.75
2007	60	3.61	58.35
2008	33	1.98	60.34
2009	129	7.75	68.09
2010	101	6.07	74.16
2011	96	5.77	79.93
2012	94	5.65	85.58
2013	135	8.11	93.69
2014	105	6.31	100
Total	1,664		
Panel B	**Frequency of SEO's by individual issuers**		
Number of offering	Frequency	Percent	Cumulative frequency
1	1,060	63.7	63.7
2	369	22.18	85.88
3	136	8.17	94.05
4	56	3.37	97.42
5	23	1.38	98.8
6	11	0.66	99.46
7	5	0.3	99.76
8	3	0.24	100
Total	1,664		
Panel C	**Frequency of primary and secondary offering**		
Offer type	Frequency	Percent	Cumulative frequency
Primary	1,284	77.16	77.16
Mixed(secondary)	380	22.84	100
Total	1,664		

Table 2

Firm and issue Characteristics. The sample requires the firm to undergo a firm commitment offering with book-build offering technique being listed in Nasdaq, Amex , or NYSE with offers greater than $1 offering price per share, also excludes REITS, ADRs, close-end funds, simultaneous offers of other securities and offers with M&A announcement in prior 6 month period. The results are presented in full sample, subsample of first quarter of 10-K file size and the last quarter of 10-K file size. T-test of differences between means of the two subsamples are shown in the last column. Panel A represents the sample and subsamples statistics of our main dependent variables being employed in this study. Gross spread as a measure of direct flotation cost and offer cancellation as an indicator of indirect flotation cost. Panel B represents sample and subsample statistics of all controlling variables being used in regression analysis.

Panel A

	All Sample				FS_Q1				FS_Q4				Test of Difference
Variable	Mean	Std.	Median	N	Mean	Std. Dev.	Median	N	Mean	Std. Dev.	Median	N	(t-value)
Gross Spread	4.87	1.39	5	1664	5.03	1.14	5.07	417	4.67	1.73	5	417	(3.45)***
Probability of Cancelation	0.03			1752	0.07				0.01				

Panel B

Variable	Mean	Std.	Median	N	Mean	Std. Dev.	Median	N	Mean	Std. Dev.	Median	N	(t-value)
Net_Proceed/TA	0.51	1.12	0.28	1666	0.70	1.84	0.34	417	0.39	0.55	0.23	417	(3.30)***
Total Asset(m)	1193.0	2602.5	253.70	1725	761.62	1631.04	212.51	432	2019.4	3605.12	363.47	431	(-6.48)***
Leverage	0.27	0.26	0.23	1725	0.24	0.22	0.22	432	0.30	0.32	0.26	431	(-2.99)***
TobinQ	2.88	2.52	1.92	1725	3.14	2.80	2.02	432	2.88	2.71	1.85	431	(1.36)
Accrual Quality	0.13	0.10	0.10	1725	0.12	0.10	0.09	432	0.12	0.10	0.10	431	(-0.54)
Illiquidity	0.01	0.01	0.001	1725	0.01	0.02	0.001	432	0.005	0.02	0.001	431	(1.87)*
Return Volatility	0.04	0.02	0.03	1725	0.04	0.02	0.034	432	0.03	0.02	0.03	431	(0.79)
Analyst	0.61	0.68	0.35	1725	0.60	0.73	0.13	432	0.62	0.68	0.31	431	(-0.38)
Underwriter Rank	7.35	2.16	1	1725	7.73	1.78	0	432	6.99	2.24	8.00	431	(5.28)***

Figure 1. Scatter plot of gross spreads and proceeds.

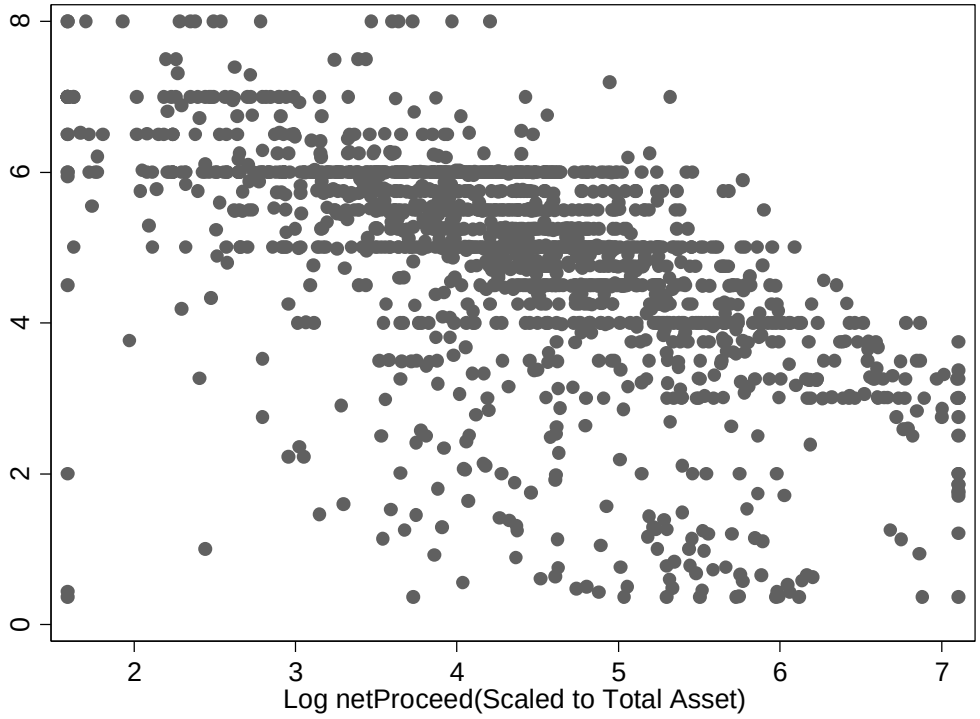


Figure 2. Scatter plot of Total Assets and proceeds.

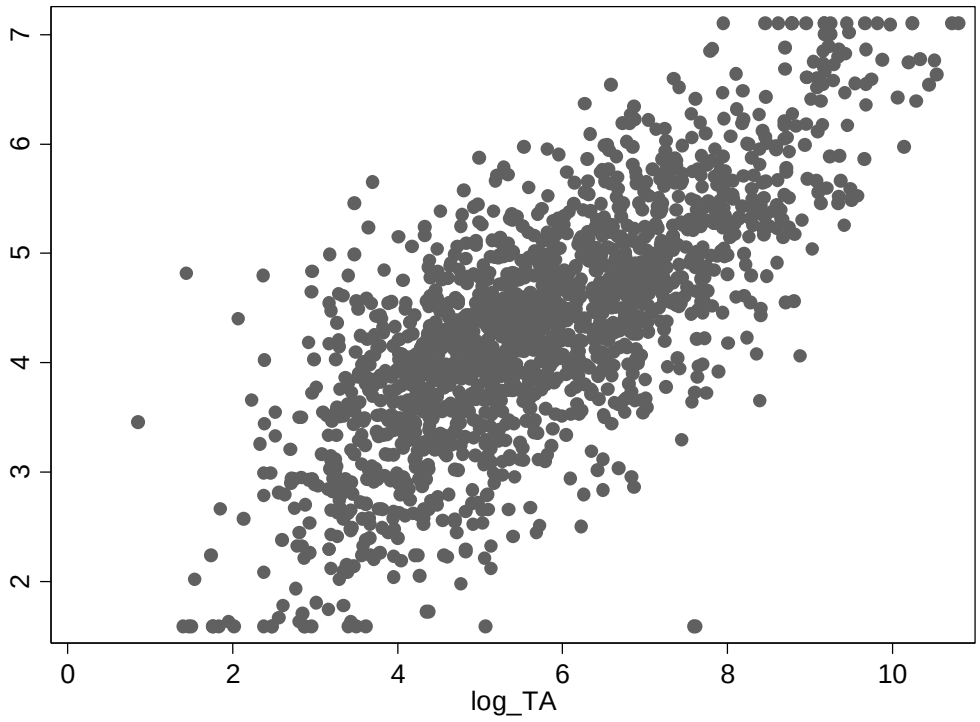


Table 3

Correlation coefficients across all variables. Section I presents the dependent variable of gross spread, section II represents the 10-K filing properties of size and sentiments, and section III presents all controlling variables being used in our regression analysis.

			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
I	(1)	Gross Spread	1												
II	(2)	File Size	-0.154	1											
	(3)	Positive (%)	0.1339	0.0241	1										
	(4)	Negative (%)	-0.016	0.2935	0.1957	1									
	(5)	Uncertain (%)	0.0596	0.2988	0.3456	0.6025	1								
III	(6)	Net_Proceed	-0.5761	0.1007	-0.0748	0.0249	-0.028	1							
	(7)	TA	-0.65	0.1749	-0.2275	0.0121	-0.091	0.7469	1						
	(8)	Leverage	-0.1949	0.078	-0.1281	0.0095	-0.0778	0.1757	0.3526	1					
	(9)	TobinQ	0.1519	-0.0506	0.2924	0.0605	0.1269	-0.0523	-0.4427	-0.193	1				
	(10)	Accruals Quality	0.0867	-0.0512	0.0924	0.0785	0.1428	-0.0079	-0.0667	0.0038	0.0881	1			
	(11)	ILLIQ	0.2643	-0.0494	-0.0317	-0.0843	-0.0555	-0.3612	-0.3039	-0.0418	-0.0389	0.0062	1		
	(12)	Return Volatility	0.3429	-0.1191	0.2424	0.2042	0.2079	-0.3334	-0.4354	-0.0845	0.2079	0.0622	0.1283	1	
	(13)	Analyst	-0.3438	0.0448	0.0116	0.005	-0.0055	0.3717	0.3697	0.0232	-0.0111	-0.0197	-0.1991	-0.175	1
	(14)	Underwriter Rank	-0.1918	-0.0939	0.0051	-0.0557	-0.1191	0.2759	0.2266	0.1238	-0.0436	0.0399	-0.1976	-0.120	0.1239

Table 4

Estimates of the offer size and 10-K information content. This table presents regression results of issuer 10-K characteristics on the log of net proceed from seasoned equity offering (SEO). The sample requires the firm to undergo a firm commitment offering with book-build offering technique being listed in Nasdaq, Amex ,or NYSE with offers greater than $1 offering price per share, also excludes REITS, ADRs, close-end funds, simultaneous offers of other securities and offers with M&A announcement in prior 6 month period. All dependent and controlling variables are explained in Appendix A. year and industry fixed effects are included in all regressions. All models are estimated using OLS regression with White heteroscedasticity-consistent standard error. T statistics are given in parentheses. ***, **, and * represent 1%, 5%, and 10% significance, respectively.

	1	2	3	4	5
Log File Size	0.2514				
	(7.46)***				
FS_Q1		-0.0707			
		(-1.10)			
FS_Q4		0.4938			
		(6.47)***			
Negative			22.0431		
			(4.07)***		
Positive				9.3414	
				(0.83)	
Uncertainty					27.5429
					(3.44)***
Leverage	0.1342	0.1825	0.2095	0.2235	0.2387
	(1.61)	(2.20)**	(2.51)**	(2.66)***	(2.85)***
TobinQ	0.0438	0.0395	0.0386	0.0367	0.0379
	(4.53)***	(4.08)***	(3.96)***	(3.75)***	(3.88)***
Accruals Quality	0.0794	0.1107	0.1314	0.1655	0.0973
	(0.40)	(0.55)	(0.65)	-0.82	-0.48
Illiquidity	-9.3646	-9.7044	-9.758	-10.0067	-9.9034
	(-8.85)***	(-9.16)***	(-9.13)***	(-9.33)***	(-9.26)***
Return Volatility	-10.4519	-10.4574	-11.3728	-10.5252	-10.9641
	(-6.98)***	(-6.95)***	(-7.43)***	(-6.90)***	(-7.19)***
Underwriter Rank	0.0703	0.0706	0.0699	0.0713	0.0714
	(7.29)***	(7.30)***	(7.16)***	(7.26)***	(7.31)***
secondary	0.2029	0.2169	0.2072	0.2182	0.2213
	(3.84)***	(4.09)***	(3.87)***	(4.05)***	(4.13)***
shelf	0.0054	0.0248	0.0006	0.0142	0.007
	(0.10)	(0.45)	(0.01)	(0.25)	(0.13)
Non-High_tech	-0.1029	-0.0809	-0.098	-0.0982	-0.0844
	(-1.44)	(-1.13)	(-1.35)	(-1.35)	(-1.16)
Analyst	0.342	0.3505	0.3394	0.3435	0.3376
	(10.97)***	(11.17)***	(10.74)***	(10.80)***	(10.65)***
NYSE	0.6108	0.6235	0.6642	0.6736	0.6907
	(11.10)***	(11.30)***	(12.05)***	(12.13)***	(12.44)***
Constant	0.7746	2.9899	3.2145	3.3726	3.084
	(0.90)	(3.74)***	(3.99)***	(4.17)***	(3.81)***
Industry Effect	Yes	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes	Yes
R2	0.4926	0.4887	0.4803	0.4751	0.4787
F	21.45	20.82	20.42	20.00	20.29

Prob>F	0.0000	0.0000	0.0000	0.0000	0.0000
N	1664	1664	1664	1664	1664

Table 5

(Estimate for the gross spread as dependent variable. This table presents regression results of issuer 10-K file size on the gross spread of seasoned equity offering (SEO) controlling for other firm and issue characteristics. The sample requires the firm to undergo a firm commitment offering with book-build offering technique being listed in Nasdaq, Amex ,or NYSE with offers greater than $1 offering price per share, also excludes REITS, ADRs, close-end funds, simultaneous offers of other securities and offers with M&A announcement in prior 6 month period. All dependent and controlling variables are explained in Appendix A. year and industry fixed effects are included in all regressions. All models are estimated using OLS regression with White heteroscedasticity-consistent standard error. T statistics are given in parentheses. ***, **, and * represent 1%, 5%, and 10% significance, respectively

	1	2	3	4	5	6	7	8
Log File Size	-0.4727	-0.2465			-0.0462	0.112	-0.1278	0.0211
	(-8.95)***	(-5.21)***			(-1.05)	(1.39)	(-2.81)***	(0.23)
FS_Q1			0.2919	0.0633				
			(2.76)***	(0.70)				
FS_Q4			-0.5971	-0.3539				
			(-4.81)***	(-3.30)***				
Log_TA					-0.5183	-0.3286		
					(-19.22)***	(-3.85)***		
Ln_Proceed							-0.4725	-0.2255
							(-14.23)***	(-1.66)*
FileSize X TA						-0.0252		
						(-2.34)**		
FileSize X Ln_Proceed								-0.0333
								(-1.87)*
Leverage		-0.3687		-0.4255	0.0595	0.058	-0.3053	-0.2954
		(-3.15)***		(-3.64)***	(0.55)	(0.54)	(-2.77)***	(-2.68)***
TobinQ		-0.0248		-0.02	-0.0919	-0.0893	-0.0042	-0.0041
		(-1.83)*		(-1.47)	(-7.22)***	(-7.00)***	(-0.32)	(-0.32)
Accruals Quality		0.4989		0.4536	0.5537	0.5749	0.5364	0.5605
		(1.78)*		(1.61)	(2.19)**	(2.28)**	(2.03)**	(2.12)**
Illiquidity		7.1791		7.5888	1.6819	1.6651	2.7546	2.8319
		(4.83)***		(5.09)***	(1.23)	(1.22)	(1.92)*	(1.97)**
Return Volatility		12.5582		12.5736	2.0996	2.1858	7.62	7.5642
		(5.96)***		(5.94)***	(1.06)	(1.11)	(3.78)***	(3.76)***
Underwriter Rank		-0.0522		-0.0529	-0.0196	-0.0201	-0.019	-0.02
		(-3.85)***		(-3.88)***	(-1.59)	(-1.63)	(-1.47)	(-1.54)
secondary		-0.6106		-0.624	-0.5623	-0.5517	-0.5148	-0.5037
		(-8.22)***		(-8.36)***	(-8.39)***	(-8.23)***	(-7.32)***	(-7.14)***
shelf		-0.3032		-0.3187	-0.1642	-0.1881	-0.3006	-0.3122
		(-3.91)***		(-4.08)***	(-2.34)**	(-2.65)***	(-4.12)***	(-4.26)***
Non-High_tech		-0.2279		-0.245	-0.2568	-0.2542	-0.2765	-0.2727
		(-2.27)**		(-2.42)**	(-2.83)***	(-2.81)***	(-2.92)***	(-2.88)***
Analyst		-0.3694		-0.3761	-0.1124	-0.1228	-0.2079	-0.2139
		(-8.43)***		(-8.52)***	(-2.70)***	(-2.93)***	(-4.85)***	(-4.98)***
NYSE		-0.6562		-0.6805	-0.1001	-0.1077	-0.3676	-0.3721
		(-8.48)***		(-8.77)***	(-1.33)	(-1.43)	(-4.86)***	(-4.92)***
Constant	5.7468	5.9441	2.7811	4.1317	6.7233	5.5673	6.3101	5.2447
	(4.13)***	(4.93)***	(2.12)**	(3.67)***	(6.18)***	(4.67)***	(5.55)***	(4.13)***

Industry Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.1789	0.4006	0.1543	0.3947	0.5136	0.5153	0.4683	0.4695
F	5.72	14.77	4.71	14.2	23	22.83	19.18	19
Prob>F	0	0	0	0	0	0	0	0
N	1664	1664	1664	1664	1664	1664	1664	1664

Table 6

(Estimate for the gross spread as dependent variable) This table presents regression results of issuer 10-K characteristics on the log of gross spread of seasoned equity offering (SEO). Panel A, we estimate gross spread using file size and controlling for indicator variables of highest and lowest 20% quantiles of Total Asset. Panel B, we estimate gross spread using file size and controlling for indicator variables of highest and lowest 20% quantiles of Net proceed. The sample requires the firm to undergo a firm commitment offering with book-build offering technique being listed in Nasdaq, Amex ,or NYSE with offers greater than $1 offering price per share, also excludes REITS, ADRs, close-end funds, simultaneous offers of other securities and offers with M&A announcement in prior 6 month period. All dependent and controlling variables are explained in Appendix A. year and industry fixed effects are included in all regressions. All models are estimated using OLS regression with White heteroscedasticity-consistent standard error. T statistics are given in parentheses. ***, **, and * represent 1%, 5%, and 10% significance, respectively.

Panel A	1	2	3
Log_FS	-0.1264	-0.2441	-0.0839
	(-2.78)***	(-5.17)***	(-1.74)*
TA_Q1	0.5354	-0.006	
	(6.51)***	(-0.02)	
TA_Q5	-1.1072		0.2855
	(-13.13)***		(0.77)
Log_Fs X TA_Q1		0.0677	
		(1.34)	
Log_Fs X TA_Q5			-0.1818
			(-3.78)***
Leverage	-0.2059	-0.2933	-0.2874
	(-1.85)*	(-2.51)**	(-2.58)***
TobinQ	-0.0516	-0.0466	-0.0259
	(-3.86)***	(-3.31)***	(-2.01)**
Accruals Quality	0.6528	0.4478	0.7339
	(2.47)**	-1.61	(2.75)***
Illiquidity	5.2543	5.3162	7.2542
	(3.66)***	(3.52)***	(5.14)***
Return Volatility	7.5292	10.5454	9.8359
	(3.72)***	(4.98)***	(4.89)***
Underwriter Rank	-0.039	-0.0455	-0.0461
	(-3.04)***	(-3.38)***	(-3.58)***
secondary	-0.6048	-0.5892	-0.6085
	(-8.64)***	(-7.99)***	(-8.61)***
shelf	-0.2401	-0.2875	-0.3005
	(-3.28)***	(-3.73)***	(-4.04)***
Non-High_tech	-0.2348	-0.2234	-0.2218
	(-2.48)**	(-2.24)**	(-2.32)**
Analyst	-0.2326	-0.3223	-0.2985
	(-5.46)***	(-7.27)***	(-7.08)***
NYSE	-0.3072	-0.645	-0.3279
	(-3.98)***	(-8.40)***	(-4.21)***
Constant	4.2576	5.8597	3.912
	(3.73)***	(4.90)***	(3.37)***
Industry Effect	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes
R2	0.4692	0.4124	0.4599
F	18.98	15.07	18.29
Prob>F	0	0	0
N	1664	1664	1664

Panel B	1	2	3
Log_FS	-0.1391	-0.2317	-0.1173
	(-3.12)***	(-4.86)***	(-2.53)**
LProceed_Q1	0.3684	-0.1029	
	(4.84)***	(-0.28)	
LProceed _Q5	-1.0891		-0.0287
	(-14.68)***		(-0.08)
Log_Fs X LProceed _Q1		0.0777	
		(1.56)	
Log_Fs X LProceed _Q5			-0.1472
			(-3.22)***
Leverage	-0.2711	-0.3273	-0.293
	(-2.49)**	(-2.81)***	(-2.68)***
TobinQ	-0.0081	-0.0198	-0.0115
	(-0.64)	(-1.47)	(-0.90)
Accruals Quality	0.558	0.4573	0.6327
	(2.14)**	(1.65)*	(2.42)**
Illiquidity	5.555	5.4315	6.9693
	(3.93)***	(3.61)***	(5.02)***
Return Volatility	8.4814	11.1237	9.4117
	(4.28)***	(5.30)***	(4.76)***
Underwriter Rank	-0.0328	-0.0394	-0.0431
	(-2.57)**	(-2.89)***	(-3.40)***
secondary	-0.5511	-0.5743	-0.5593
	(-7.96)***	(-7.76)***	(-8.04)***
shelf	-0.283	-0.3228	-0.2925
	(-3.92)***	(-4.20)***	(-4.02)***
Non-High_tech	-0.2752	-0.2334	-0.2575
	(-2.94)***	(-2.34)**	(-2.74)***
Analyst	-0.2602	-0.3316	-0.2969
	(-6.26)***	(-7.55)***	(-7.19)***
NYSE	-0.3668	-0.6078	-0.393
	(-4.94)***	(-7.89)***	(-5.29)***
Constant	4.1623	5.6264	4.0923
	(3.69)***	(4.70)***	(3.60)***
Industry Effect	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes
R2	0.4825	0.4132	0.4783
F	20.02	15.12	19.68
Prob>F	0.0000	0.0000	0.0000
N	1664	1664	1664

Table 7

Estimate for the offer withdrawal and sentiment measures. Cancel dummy as dependent variable.This table presents regression results of issuer 10-K sentiments on the probability of offer withdrawal. The sample requires the firm to undergo a firm commitment offering with book-build offering technique being listed in Nasdaq, Amex ,or NYSE with offers greater than $1 offering price per share, also excludes REITS, ADRs, close-end funds, simultaneous offers of other securities and offers with M&A announcement in prior 6 month period. All dependent and controlling variables are explained in Appendix A. year and industry fixed effects are included in all regressions. All models are estimated using OLS regression with White heteroscedasticity-consistent standard error. T statistics are given in parentheses. ***, **, and * represent 1%, 5%, and 10% significance, respectively.

	1	2	3	4	5	6	7
Negative (%)	1.7275	2.4672					2.5463
	(1.46)	(2.09)**					(2.04)**
Positive (%)			-3.0308	-2.8895			-2.7019
			(-1.22)	(-1.17)			(-1.08)
Uncertainty (%)					0.2282	0.4403	-0.4718
					(0.13)	(0.25)	(-0.25)
Leverage		0.0349		0.0355		0.0368	0.0336
		(1.91)*		(1.94)*		(2.00)**	(1.83)*
TobinQ		-0.0023		-0.0024		-0.0025	-0.0023
		(-1.10)		(-1.14)		(-1.16)	(-1.08)
Accruals Quality		-0.0567		-0.0515		-0.0537	-0.0546
		(-1.27)		(-1.16)		(-1.20)	(-1.22)
Illiquidity		0.4736		0.4463		0.4505	0.4703
		(2.14)**		(2.02)**		(2.04)**	(2.13)**
Return Volatility		0.379		0.4974		0.4766	0.3962
		(1.12)		(1.49)		(1.42)	(1.17)
Underwriter Rank		-0.0025		-0.0022		-0.0023	-0.0024
		(-1.16)		(-1.02)		(-1.09)	(-1.10)
secondary		-0.0268		-0.0267		-0.0258	-0.0277
		(-2.29)**		(-2.28)**		(-2.20)**	(-2.36)**
shelf		-0.0874		-0.086		-0.0861	-0.0873
		(-7.16)***		(-7.05)***		(-7.06)***	(-7.15)***
Non-High_tech		-0.0098		-0.0087		-0.0092	-0.0094
		(-0.62)		(-0.55)		(-0.58)	(-0.59)
Analyst		-0.0046		-0.0031		-0.004	-0.0039
		(-0.66)		(-0.45)		(-0.56)	(-0.55)
NYSE		-0.0149		-0.0152		-0.0139	-0.0161
		(-1.23)		(-1.25)		(-1.14)	(-1.32)
Constant	-0.0223	0.0482	0.0081	0.0773	-0.0069	0.0627	0.0616
	(-0.12)	(0.26)	(0.04)	(0.41)	(-0.04)	(0.33)	(0.33)

Industry Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.0675	0.1166	0.0671	0.115	0.0663	0.1143	0.1173
F	1.97	3.03	1.96	2.98	1.94	2.96	2.96
Prob>F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	1725	1725	1725	1725	1725	1725	1725

Table 8

Principal components of information asymmetry. This table describes the loading of the first three principal components measured to conduct principal component analysis. We reduce the number of variables to measure information asymmetry. The scree plot in Figure 3 shows the eigenvalues of components that capture the variation in eight primitive measures of information asymmetry. Component 1 captures a substantial amount of variation, followed by 2 and 3 with marginal eigenvalues, but the rest do not. The component loading have reverse signs as predicted, thereby, we multiply this component by -1 in our regression analysis.

	Predicted correlation with information	Comp 1	Comp 2	Comp 3	Unexplained	KMO
Log_TA	-	**0.5360**	0.1779	-0.2085	0.2652	0.6178
Accruals	+	-0.0711	**0.4032**	**0.4403**	0.5806	0.6008
Illiquidity	+	**-0.3289**	-0.2046	**-0.3743**	0.5522	0.7570
Analyst	-	**0.3578**	0.0912	0.1265	0.6832	0.8294
Return Volatility	+	**-0.3992**	**0.3674**	0.0964	0.4665	0.7329
Underwriter Rank	-	0.2188	-0.1868	**0.7048**	0.3081	0.6595
NYSE	-	**0.4606**	**-0.3026**	-0.1756	0.3764	0.6673
Shelf	-	0.2280	**0.7029**	-0.26385	0.2088	0.4471
Overall KMO						**0.6523**
Eigenvalue		2.2591	1.2101	1.0899		

Table 9

Estimates of the gross spread by10-K file size and principal components of information asymmetry this table presents 2SLS estimates of gross spread. First stage estimates the L_proceed and second stage incorporates the estimated L_Proceed to estimate Gross Spread controlling file size and principal components of information asymmetry. The sample requires the firm to undergo a firm commitment offering with book-build offering technique being listed in Nasdaq, Amex ,or NYSE with offers greater than $1 offering price per share, also excludes REITS, ADRs, close-end funds, simultaneous offers of other securities and offers with M&A announcement in prior 6 month period. Columns 1,3 and 4 are based on ordinary least square (OLS) and column 2 and 5 are based on two-stage least square(2SLS),where the dependent variable in the first stage regression is log of gross proceed and using instruments of non-high tech indicator and secondary indicator. All dependent and controlling variables are explained in Appendix A. year and industry fixed effects are included in all regressions. All models are estimated using OLS regression with White heteroscedasticity-consistent standard error. T statistics are given in parentheses. ***, **, and * represent 1%, 5%, and 10% significance, respectively.

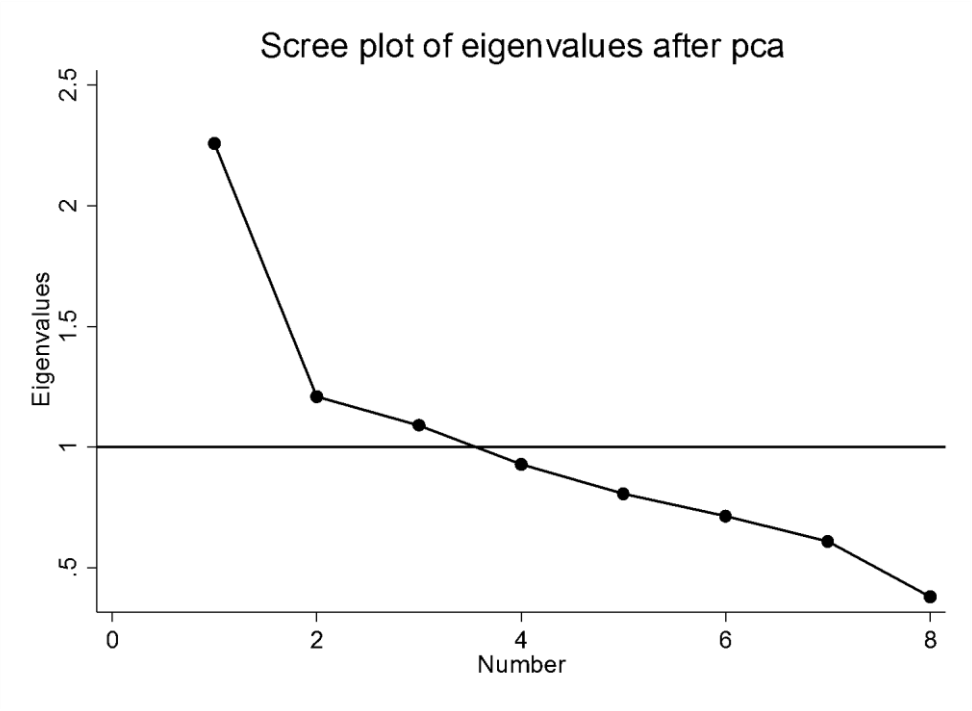
	OLS	2SLS	OLS	OLS	2SLS
	1	2	3	4	5
Log_FS	0.0412	-0.3115	-0.0722	-0.0531	-0.3120
	(2.04)**	(-2.76)***	(-3.09)***	(-2.50)**	(-2.52)**
PC1	0.405	-1.1172			-1.2051
	(14.91)***	(-2.38)**			(-2.49)**
PC2			0.0111		0.0724
			(-0.37)		(1.03)
PC3				0.0595	0.1975
				(2.11)**	(2.80)***
L_Proceed	-0.3377	-3.1064	-0.6776	-0.6863	-3.2854
	(-10.09)***	(-3.69)***	(-26.01)***	(-26.07)***	(-3.78)***
leverage	-0.2354	0.11206	-0.3631	-0.3694	0.1328
	(-2.28)**	(0.48)	(-3.31)***	(-3.38)***	(0.51)
Tobin Q	0.0014	0.2037	0.0433	0.041	0.1972
	(0.12)	(3.07)***	(3.67)***	(3.47)***	(3.03)***
secondary	-0.4667		-0.4949	-0.5022	
	(-7.54)***		(-7.38)***	(-7.63)***	
Non-High_tech	-0.1074		0.0764	0.0602	
	(-1.84)*		(1.24)	(0.98)	
Constant	6.2785	19.8727	8.3654	8.2851	20.6599
	(28.46)***	(4.76)***	(41.10)***	(44.87)***	(4.76)***
Industry Effect	Yes	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes	Yes
R2	0.4544	0.5075	0.3812	0.3828	
F	197.05		145.76	146.75	
Wald Chi^2		284.99			230.00
Prob>F	0.0000		0.0000	0.0000	
Prob>Chi^2		0.0000			0.0000
N	1664	1664	1664	1664	1664

Appendix A

Definition of all variables used in this study.

Variable	Definition
Dependent variables	
Announcement Return	Cumulative abnormal return for the period[-1,+1]
Cancelled SEO	Indicator variable equals to 1 if the offer is cancelled and zero if completed. Reported by SDC Platinum.
Gross Spread	The difference between underwriters' purchase price and SEO offer price. Reported by SDC Platinum.
Ln_Proceed	Natural logarithm of net proceed after paying all fees. Reported by SDC Platinum.
Dif_Illiquidity	Difference between Amihud illiquidity measure before and after SEO.
Controlling variables	
Ln_FS	Natural logarithm of 10-K file size.
FS_Q1	Indicator variable of lowest quartile of file size
FS_Q4	Indicator variable of highest quartile of file size
Negative (%)	Percentage of negative words used in 10-K
Positive (%)	Percentage of positive words used in 10-K
Uncertainty (%)	Percentage of uncertain words used in 10-K
Ln_TA	Natural logarithm of total asset
TA_Q1	Indicator variable of lowest 20% quantile of total asset
TA_Q5	Indicator variable of highest 20% quantile of total asset
Proceed_Q1	Indicator variable of lowest 20% quantile of net proceed
Proceed_Q5	Indicator variable of highest 20% quantile of net proceed
PC1	First principal Component captures most variation of total asset, illiquidity, analyst, return volatility, and NYSE indicator.
PC2	Second principal Component captures most variation of accrual quality, return volatility, NYSE indicator, and shelf indicator.
PC3	Third principal Component captures most variation of accrual quality, illiquidity, and underwriter rank.
Leverage	Long term debt plus debt in current liabilities scaled by total asset
TobinQ	Total asset minus book value of common equity plus market value of shares outstanding scaled by total asset.
Accruals Quality	Standard deviation of residuals in FDD model developed by Lee and Masulis(2009).
Illiquidity	The natural logarithm of stock illiquidity, calculated as the average of daily Amihud (2002) illiquidity measures over the period [-250, -51]; the daily illiquidity measure is defined as the ratio of the daily absolute return to the dollar trading volume on that day, multiplied by 100,000
Return Volatility	The standard deviation of daily stock return during the trading period (-90, -11) prior to the announcement day of SEO.
Underwriter Rank	From Ritter website.
secondary	Indicator variable equals to 1 if offer includes portion of current shareholders selling their share and zero if it's all offered by firm as primary shares.
shelf	Indicator variable equals to 1 if firm is using shelf registration and zero if not.
Non-High_tech	Indicator variable equals to 1 if firm is a non-high tech company and zero otherwise.
Analyst	Natural logarithm of number of analyst following the firm in the last quarter prior to offering
NYSE	Indicator variable if firm is listed in NYSE and zero if not.

Appendix B


Scree plot of eigenvalues after pca
Eigenvalues
.5
1
1.5
2
2.5
0
2
4
6
8
Number

References

1. Li, F., *Annual report readability, current earnings, and earnings persistence.* Journal of Accounting and economics, 2008. **45**(2): p. 221-247.
2. Loughran, T. and B. McDonald, *Measuring readability in financial disclosures.* The Journal of Finance, 2014. **69**(4): p. 1643-1671.
3. Biddle, G.C., G. Hilary, and R.S. Verdi, *How does financial reporting quality relate to investment efficiency?* Journal of Accounting and Economics, 2009. **48**(2): p. 112-131.
4. Baron, D.P., *A model of the demand for investment banking advising and distribution services for new issues.* The Journal of Finance, 1982. **37**(4): p. 955-976.
5. Rock, K., *Why new issues are underpriced.* Journal of financial economics, 1986. **15**(1-2): p. 187-212.
6. Eckbo, B.E., R.W. Masulis, and O. Norli, *Security offerings.* 2007.
7. Jegadeesh, N. and D. Wu, *Word power: A new approach for content analysis.* Journal of Financial Economics, 2013. **110**(3): p. 712-729.
8. Lee, G. and R.W. Masulis, *Seasoned equity offerings: Quality of accounting information and expected flotation costs.* Journal of Financial Economics, 2009. **92**(3): p. 443-469.
9. Marquardt, C. and C.I. Wiedman, *Voluntary disclosure, information asymmetry, and insider selling through secondary equity offerings.* Contemporary Accounting Research, 1998. **15**(4): p. 505-537.
10. Corwin, S.A., *The determinants of underpricing for seasoned equity offers.* The Journal of Finance, 2003. **58**(5): p. 2249-2279.
11. Liu, Y. and P.H. Malatesta, *Credit ratings and the pricing of seasoned equity offerings.* Unpublished working paper (University of Washington), 2006.
12. Denis, D.J., *Investment opportunities and the market reaction to equity offerings.* Journal of Financial and Quantitative Analysis, 1994. **29**(02): p. 159-177.
13. Altınkılıç, O. and R.S. Hansen, *Are there economies of scale in underwriting fees? Evidence of rising external financing costs.* Review of financial Studies, 2000. **13**(1): p. 191-218.
14. Habib, M.A. and A.P. Ljungqvist, *Underpricing and entrepreneurial wealth losses in IPOs: Theory and evidence.* Review of Financial Studies, 2001. **14**(2): p. 433-458.
15. Burch, T.R., V. Nanda, and V. Warther, *Does it pay to be loyal? An empirical analysis of underwriting relationships and fees.* Journal of Financial Economics, 2005. **77**(3): p. 673-699.
16. Smith, C.W., *Alternative methods for raising capital: Rights versus underwritten offerings.* Journal of financial economics, 1977. **5**(3): p. 273-307.
17. Butler, A.W., G. Grullon, and J.P. Weston, *Stock market liquidity and the cost of issuing equity.* Journal of Financial and Quantitative Analysis, 2005. **40**(02): p. 331-348.
18. Lawrence, A., *Individual investors and financial disclosure.* Journal of Accounting and Economics, 2013. **56**(1): p. 130-147.
19. Lang, M.H. and R.J. Lundholm, *Corporate disclosure policy and analyst behavior.* Accounting review, 1996: p. 467-492.
20. Healy, P.M., A.P. Hutton, and K.G. Palepu, *Stock performance and intermediation changes surrounding sustained increases in disclosure.* Contemporary accounting research, 1999. **16**(3): p. 485-520.
21. Lehavy, R., F. Li, and K. Merkley, *The effect of annual report readability on analyst following and the properties of their earnings forecasts.* The Accounting Review, 2011. **86**(3): p. 1087-1115.
22. Loughran, T. and B. McDonald, *When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks.* The Journal of Finance, 2011. **66**(1): p. 35-65.
23. Loughran, T. and B. McDonald, *IPO first-day returns, offer price revisions, volatility, and form S-1 language.* Journal of Financial Economics, 2013. **109**(2): p. 307-326.
24. Francis, J., et al., *The market pricing of accruals quality.* Journal of accounting and economics, 2005. **39**(2): p. 295-327.

25. Amihud, Y., *Illiquidity and stock returns: cross-section and time-series effects.* Journal of financial markets, 2002. **5**(1): p. 31-56.
26. Hansen, R.S. and P. Torregrosa, *Underwriter compensation and corporate monitoring.* The Journal of Finance, 1992. **47**(4): p. 1537-1555.
27. Hong, H., T. Lim, and J.C. Stein, *Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies.* The Journal of Finance, 2000. **55**(1): p. 265-295.
28. Brown, S., S.A. Hillegeist, and K. Lo, *The effect of earnings surprises on information asymmetry.* Journal of Accounting and Economics, 2009. **47**(3): p. 208-225.
29. Li, X. and R.W. Masulis, *How do venture investments by different classes of financial institutions affect the equity underwriting process?* Available at SSRN 891738, 2007.
30. Karpoff, J.M., M. Barnett, and T. Pollock, *Does reputation work to discipline corporate misconduct*. 2012: Oxford: Oxford University Press.
31. Kumar, R. and D.K. Shome, *The revival of shelf-registered corporate equity offerings.* Journal of Corporate Finance, 2008. **14**(1): p. 32-50.
32. Lease, R.C., R.W. Masulis, and J.R. Page, *An investigation of market microstructure impacts on event study returns.* The Journal of Finance, 1991. **46**(4): p. 1523-1536.
33. Safieddine, A. and W.J. Wilhelm, *An empirical investigation of short-selling activity prior to seasoned equity offerings.* The Journal of Finance, 1996. **51**(2): p. 729-749.
34. Lubotsky, D. and M. Wittenberg, *Interpretation of regressions with multiple proxies.* The Review of Economics and Statistics, 2006. **88**(3): p. 549-562.
35. Boone, A.L., et al., *The determinants of corporate board size and composition: An empirical analysis.* Journal of Financial Economics, 2007. **85**(1): p. 66-101.
36. Maddala, G.S., *A perspective on the use of limited-dependent and qualitative variables models in accounting research.* The Accounting Review, 1991. **66**(4): p. 788-807.
37. Karpoff, J.M., G. Lee, and R.W. Masulis, *Lockup agreements in seasoned equity offerings: evidence of optimal contracting.* Journal of Financial Economics (JFE), Forthcoming, 2012.