Disclosure Crowdsourcing by Lawyers

Michael Drake
Brigham Young University

Kenneth Merkley
Indiana University

Chase Potter
Indiana University

John Treu
West Virginia University

Preliminary draft
Please do not cite without permission

January 2020

We thank Jeff McMullin and Melissa Western-Lewis for helpful comments and suggestions. We acknowledge and thank research assistants as Brigham Young University and graduate assistants Brian George and Brandon Nicholas for assistance with data collection in support of this research.

Disclosure Crowdsourcing by Lawyers

ABSTRACT

Lawyers play an important advisory role in the IPO process, but our understanding of how they influence the process is limited. We provide evidence on the extent to which the external legal counsel of a company undergoing an IPO engages in "disclosure crowdsourcing" by obtaining information from other companies' IPO disclosures and the impact of such crowdsourcing on important IPO outcomes. We document an increase in disclosure crowdsourcing that begins approximately ten weeks prior to an IPO's first S-1 filing and declines after thereafter. We find that disclosure crowdsourcing is associated with more efficient interactions with the SEC during the pre-IPO filing review period, increased speed of price formation after the offering, and a decreased likelihood of litigation. Overall, the evidence suggests that companies experience several favorable IPO-related outcomes when their legal counsel engages in greater disclosure crowdsourcing.

1. Introduction

Financial disclosures play an important role in capital markets as a solution to the "lemons problem" (Akerlof 1970) by reducing the information asymmetry between managers and providers of capital. The role of disclosure in reducing this information gap is highly salient during the IPO process when a broad set of financial disclosures are required to be filed with the Securities and Exchange Commission. Generally, the first public disclosure issued by a company in the IPO process is the Form S-1 Registration Statement (hereafter "S-1"). The S-1 provides investors with their first look into the financial position and performance of a private company and thus can trigger considerable attention from investors. For example, our data indicate that when Facebook filed their S-1 on February 1, 2012, it was downloaded approximately 262,000 times in the first 24-hours that it was available in EDGAR. The initial S-1 filing is typically revised through subsequent amendments as part of the SEC review process (e.g., S-1/A filings) and the final version constitutes the IPO prospectus.

Given the importance of the initial S-1 filing to the success of the IPO, companies invest considerable time and resources into the preparation of this filing, which typically includes hiring external legal counsel ("IPO Lawyers" hereafter).[1] IPO Lawyers play an important advisory role in the production of initial S-1 filings with a particular focus on the narrative sections of the S-1 filing, which include the management discussion and analysis, uses of proceeds, risk factors sections, and notes (Beatty and Welch 1996). In order to advise on the extent, nature, and content of any particular financial disclosure, IPO Lawyers may examine the S-1 filings of other

[1] The US Office of Management and Budget estimates that the average S-1 filing involves approximately 972.32 hours of preparation and only 25% of these hours are performed by the issuer (Federal Register/ Vol. 79, No. 150/Tuesday, August 5, 2014).

companies who recently went public to identify disclosures norms and best practices. We term this practice “disclosure crowdsourcing.”

Our objective is to investigate the influence of disclosure crowdsourcing by IPO Lawyers on the responses to the IPO of regulators, capital providers, and courts of law. More specifically, we predict that greater disclosure crowdsourcing by IPO Lawyers will be associated with more favorable outcomes with respect to the efficiency of the SEC’s pre-IPO approval process as well as improvements in two important post-IPO outcomes: the speed of price formation after the IPO and the likelihood of litigation. Our rationale is that IPO Lawyers that crowdsource disclosures to a greater degree will help produce a set of disclosures that are more consistent with the disclosure expectations of regulators and investors based on past IPOs. The alignment of disclosure and expectations will result in more efficient vetting and processing of the disclosures by these parties. Furthermore, under Section 11 of the Securities Act of 1933, investors have the right to sue issuers for subsequent declines in value arising from material misstatements or omissions in the S-1. Disclosure crowdsourcing may thus enable IPO Lawyers to help their clients avoid such issues and, in turn, also lower the likelihood of subsequent litigation.

Our predictions are not without tension, however, as disclosure crowdsourcing could lead to excessive use of boilerplate language that is more efficient from the perspective of the disclosure drafting process and benefits lawyers in the event of a malpractice claim, but is also less informative. Badawi (2017) observes that S-1 filings have a fair amount of boilerplate language and practice manuals recommend that companies begin drafting the S-1 by looking at other prospectuses to “take the drudgery out of the chore” (see, Bartlett 1999, p. 137).[2] Disclosure crowdsourcing may benefit lawyers should a malpractice lawsuit arise. The standard of care in

---

[2] Consistent with this recommendation, Hanley and Hoberg (2010) find that IPO prospectuses have greater content similarities when they are developed by the same underwriters and lawyers.

attorney malpractice cases is based on whether the lawyer provided counsel in a manner that is consistent with other lawyers in their line of work and so disclosure crowdsourcing simply be a precautionary measure for attorneys that provides little or no benefit to the companies they represent. Also, it is also possible that high levels of disclosure crowdsourcing is associated with settings where IPO lawyers feel particularly inexperienced, and therefore, need to leverage the work of others more heavily in the draft process. If this is the case, then the overall quality of the filings may be lower.

Our empirical analyses are based on a novel dataset of EDGAR S-1 filing downloads made by law firms that advise IPO clients during the 2003 to 2017 period. We measure disclosure crowdsourcing as the number of times that an IPO lawyer accessed an S-1 or S-1/A filing of another company in the months prior to the S-1 filing of their client. We first document the incidence of disclosure crowdsourcing and find that IPO Lawyers significantly increase their downloads of other firms' S-1 filings beginning about 13 weeks prior to the initial S-1 filing. In the week prior to the initial S-1 filing, IPO Lawyers download 53 S-1 filings, on average, compared to 16 downloads by control lawyers (i.e., lawyers not preparing an S-1 in the near future). This level of increased S-1 download activity continues for another 15 weeks during the SEC's S-1 review period. This pattern of download activity is consistent with the notion that IPO Lawyers conduct disclosure crowdsourcing as part of their work to help their clients prepare IPO filings.

We next examine the association between disclosure crowdsourcing and three important IPO-related outcomes. Our models include controls for attributes of the IPO firm (e.g., size, leverage, performance, sales, and age), the equity offering (venture capital backing, insider retention rate, proportion of secondary shares sold, prestige of underwriters , and the total amount of legal fees the IPO lawyers charge through the IPO process. When appropriate, we also include law firm, industry and year fixed effects to control for unobserved, time-invariant heterogeneity.

We begin by examining the efficiency of an IPO's SEC review that occurs before the offering goes public. The pre-IPO review process involves back-and-forth correspondence between the companies and the SEC via comment letters. Prior research finds that the average IPO goes through several rounds of comment letters (Ertimur and Nondorf 2006; Li and Liu 2017). The comment letter process is costly in terms of both management time and resources. We measure the efficiency with which an IPO completes the SEC comment letter process using two constructs: (1) the degree to which the SEC disagrees with the content of the initial S-1 filing (as measured by the number of issues identified in the SEC's initial comment letter) and (2) how quickly the a company is able to satisfy the SEC's concerns (as measured by the time and number of subsequent amendments to the initial S-1). We find some, albeit modest, evidence that disclosure crowdsourcing is negatively associated with how quickly the SEC's concerns are resolved but not with the degree of SEC disagreement. Our evidence is consistent with law firms that conduct more disclosure crowdsourcing being able to help their clients to more efficiently resolve disclosure issues with the SEC.

We next test whether disclosure crowdsourcing is associated with the speed of price formation post-IPO. Following prior research (Butler, Kraft, and Weiss 2007; Bushman, Smith, and Wittenberg-Moerman 2010; Twedt 2016), we measure the speed of price formation using the intraperiod timeliness metric over the 20 trading days (approximately one month) after the IPO. Consistent with our predictions, we find that if advising IPO Lawyers engage in greater disclosure crowdsourcing then the companies' they advise experience significantly faster post-IPO price formation. This finding is consistent with capital providers being better able to process the information contained in the IPO disclosures when advising lawyers spend more time reviewing the disclosures of other IPOs.

Finally, we test whether disclosure crowdsourcing relates to the likelihood of shareholder litigation within three calendar years after the IPO. We find that the likelihood that a company will be sued by its shareholders following an IPO is significantly lower when advising lawyers conduct more disclosure crowdsourcing. In terms of economic significance, we find that a one standard deviation change in crowdsourcing is associated with a 5 percent reduction in the likelihood of subsequent litigation, which represents about a 50% reduction relative to the sample mean. This finding is consistent with disclosure crowdsourcing significantly reducing companies' exposure to lawsuits.

It is important to recognize that the first of the IPO outcomes we examine, efficiency of SEC interactions, occurs *before* the IPO and that the interactions with the SEC could spur additional disclosure crowdsourcing during the review period. Furthermore, both of these activities – efficiency of SEC interactions and disclosure crowdsourcing during the filing review period – could potentially influence the other two post-IPO outcomes we examine (speed of price formation and litigation), which occur *after* the IPO.[3] This raises the possibility of the existence of indirect channels through which disclosure crowdsourcing prior the initial S-1 influences post-IPO outcomes. We examine this possibility using a system of equations (path regression) designed to examine the direct and indirect effects of IPO law firm disclosure crowdsourcing on post-IPO outcomes. While we find some evidence that disclosure crowdsourcing is associated with our mediating variables (SEC review efficiency and review period disclosure crowdsourcing), we do

[3] Ertimur and Nondorf (2006) provide some evidence of a positive relation between the "smoothness" of the SEC comment letter process and the quality of the firm's information environment at the IPO (as measured by information asymmetry proxies). Li and Liu (2017) find that IPO firms tend to receive their initial offer price when they receive more comment letters.

not find evidence of any significant indirect paths. This evidence is consistent with the idea that the benefits of disclosure crowdsourcing associated with the speed of price formation and litigation risk are reflected in the initial S-1 filing and not in subsequent revisions to the filing that result from the SEC review.

Our study makes several contributions to the literature. First, it contributes to our understanding of the IPO process, the role of disclosures in this process, and the influence of various parties that assist firms in this process. While numerous prior studies provide evidence on the relation between disclosure and IPO pricing (Beatty and Ritter 1986; Ljungqvist and Wilhelm 2003; Schrand and Verrecchia 2005; Leone, Rock, and Willenborg 2007), there is relatively little empirical evidence on the role that IPO Lawyers play in the process. We are aware of only two empirical studies in this area. Beatty and Welch (1996) provide evidence that well-paid IPO Lawyers reduce underpricing at the offering. Hanley and Hoberg (2010) examine the information content of IPO prospectus and find that the information content of prospectus is positively associated with lawyer fees. This suggests that effort in the due diligence process can influence the informativeness of the disclosures. We contribute to this growing line of research by providing evidence that the information-gathering activities of IPO Lawyers are associated with important economic outcomes to their clients. More specifically, we find that higher levels of crowdsourcing by IPO Lawyers is associated with a more efficient vetting process, faster price formation, and lower litigation risk.

Second, our paper contributes to the growing literature on the role of external legal counsel in financial reporting in general. While prior studies tend to focus on the influence of legal experts inside the firm (Kwak et al. 2012; Bird et al. 2014; Hopkins et al. Venkatachalam 2015), empirical evidence on the role of external counsel is an emerging stream of literature (Hanley and Hoberg

2010, 2012, Dechow and Tan 2017; Bozanic et al. 2019). While these studies tend to focus on how lawyers influence financial reporting, our study provides the first empirical evidence on how lawyers' use of financial reports influences important client outcomes. Additionally, the introduction of our unique lawyer crowdsourcing measure provides this literature with a new way to directly observe the activities of legal advisors. This is an important improvement over measures used in extant research (e.g., is a legal professional one of the company's top officers) because it provides time-series and cross-sectional variation in the actual activities of the lawyers. This measure may be useful to future research seeking to better understand the relation between external legal counsel advising and financial disclosures outside the IPO setting.

Third, our study contributes to the emerging literature that examines the practice of disclosure crowdsourcing by various parties that participate in the disclosure production process. Drake et al. (2019) examine the external auditor disclosures crowdsourcing in the context of 10-K filings and also provide evidence of positive, real outcomes (e.g., greater levels of financial statement disaggregation). They do not, however, examine whether the practice is associated with market or legal outcomes, which is an important contribution of our study. Finally, we extend the literature on the uses of financial disclosures by various players including tax authorities (Bozanic, Hoopes, Thornock, and Williams 2017), regulators (Stice-Lawrence 2017), equity research analysts (Gibbon et al. 2018), and professional investors (Drake et al. 2020). These findings should be of particular interest to standard setters and regulators given the important interactions of law firms with capital market participants.

2. Background and Empirical Predictions

2.1 Institutional Background

Companies that want access to public capital markets in the United States must first obtain the approval of the SEC by filing a new securities registration statement (see 1 Federal Securities Act of 1933 § 7.03 (2019)). Prior to the initial public offering (IPO) of shares, companies typically file a Form S-1 with the SEC by electronic transmission through EDGAR to start the formal approval process. The initial S-1 is reviewed by the SEC's Division of Corporation Finance, which assigns an attorney, an accountant, and a financial analyst to review the filing and provide comments. These comment letters include questions, possible deficiencies for review, and suggestions regarding revisions. The company issues a response to the SEC's comments with an amended S-1 filing (S-1/A) and the process repeats until the SEC issues a final approval letter. The final, approved document is known as the prospectus.

Following this approval, the company markets the offering (i.e., conducts roadshows and interacts with institutional investors) and finalizes the offering documents, including the effective date and offering price. The offering is then brought to market and the stock is formally listed on an exchange. To help navigate the complexities of the registration process and the subsequent market offering, companies involve the help of various external parties including lawyers, investment banks, underwriters, and auditors.

We focus on the role of lawyers in the IPO process for three reasons. First, companies spend significant resources employing lawyers in the IPO process. Fees for IPO Lawyers constitute a significant portion of the direct costs companies bear when conducting an IPO. Excluding the underwriter discount that is based on a percentage of gross proceeds from the IPO, legal fees are the single largest IPO expense, accounting for 46% of the remaining direct costs.[4] Evidence also suggests that these fees are increasing over time. Audit Analytics reports that legal fees as a percentage of the IPO proceeds nearly tripled from 2008 to 2016 going from approximately 0.5 percent to nearly 1.5 percent (Bradford and Bogdan 2017; also see Dambra et al. 2015). They further report that the average legal fee is approximately $1.7 million in 2016.[5] Beatty and Welch (1996) examine IPO legal fees and find that they are negatively associated with IPO underpricing. Thus, at least one observed benefit of paying for legal advice is an increase in the proceeds received.

Second, IPO Lawyers play an important role in the actual drafting of the S-1. In fact, McClane (2015) state that "counsel for the issuer typically *takes the lead* in drafting the prospectus and thus has a large amount of control over the draft" (p. 140, italics added for emphasis). Prior research provides some empirical evidence that IPO Lawyers have a significant impact on the content of S-1 filings. Hanley and Hoberg (2010) find that legal fees are positively associated with the information content of the S-1. Specifically, they find that a one standard deviation increase in the information content of the document is associated with a 25 percent increase in legal fees. This suggests that greater IPO Lawyer due diligence can improve the information provided in the filing.

Third, IPO Lawyers have significant expertise in helping their clients identify and mitigate risks associated with the IPO process, including regulatory and litigation risk. Navigating the SEC approval process can be challenging and time-consuming given the technical nature of the registration filings and the need for, often repeated, correspondence with SEC officials (Bozanic

[4] A study by PWC identifies the costs for accounting, legal, printing, registration and miscellaneous (See PWC, 2015). In addition to these direct costs there is typically an underwriter discount of 4% to 7% of the gross proceeds from the IPO.

[5] PwC reports similar average legal fees for drafting the registration statement and providing related advice cost of $1.8 million, on average (PwC, 2015).

et al. 2019). Delays in this process could cause management to allocate more time and resources to the IPO process at the expense of managing operations. Furthermore, misstatements or omissions in these filings can result in subsequent fines or penalties from the SEC, but also open the door to potential shareholder litigation. Section 10(b) and Rule 10b-5 of the Securities and Exchange Act permits shareholders to file a lawsuit for damages arising from the disclosure by a publicly traded company of misleading or incomplete information. Cornerstone Research notes that increased IPO activity appears to be correlated with increased levels of shareholder lawsuits in the ensuing years (Cornerstone 2018). Prior research further suggests that the risk of litigation by shareholders constrains financial disclosures in the IPO setting (Fan, 2007; Ball and Shivakumar 2008) and that aggressive pre-IPO reporting can result in shareholder lawsuits (Billings and Lewis-Western 2016).[6]

Given the significant costs and risks of the IPO process, it is not surprising that companies hire lawyers to play an important advisory role in the production of S-1 filings and the subsequent SEC comment letter response process. However, little has been documented empirically about the process these lawyers take for preparing these important disclosures and the degree to which their involvement is associated with observable economic benefits. This lack of understanding is due, in part, to data limitations as internal lawyer processes are typically protected by attorney-client privilege and the work product doctrine. While Hanley and Hobert (2010) suggest that IPO Lawyers turn to S-1 filings already in the public domain to assist in the process, they do not provide any direct tests of this practice. We shed light on the role of external legal counsel by focusing on one potentially important element of the S-1 preparation process — disclosure crowdsourcing.

[6] For example, Billings and Lewis-Western (2016) find that earnings inflation in the pre-IPO period triggers shareholder lawsuits in situations where investors are more likely to rely on such disclosures because other sources of information are lacking or less reliable.

2.2 Disclosure Crowdsourcing

IPO Lawyers routinely examine the S-1 filings of other companies who recently went public as such disclosures establish disclosure norms for the drafting process.[7] Hanley and Hoberg (2010) find that the language in new prospectus filings contains content from other prospectuses already in the public domain, which provides indirect evidence of disclosure crowdsourcing. They also provide some evidence, again indirect, of cost efficiencies associated with IPO filings that include more *standard* information content contained in these other prospectuses. Specifically, they find that a one standard deviation increase in standard S-1 content is associated with a modest reduction in legal fees (nearly 8 percent).

While this evidence suggests that the practice of disclosure crowdsourcing may result in cost efficiencies, it is less clear whether the practice influences any real, post-IPO economic outcomes. We argue that disclosure crowdsourcing by IPO Lawyers could be associated with three positive IPO-related outcomes including more efficient interactions with the SEC during the filing review period, improved pricing dynamics, and reduced litigation risk. We motivate our predictions related to each of these IPO-related outcomes in turn.

Once an S-1 is filed with the SEC, it undergoes an extensive review that generally involves back-and-forth correspondence between the company and the SEC via comment letters. The SEC issues a comment letter when it believes the filing is inadequate in some area(s) that could be improved. While prior research finds that the average IPO goes through several rounds of comment letters with the SEC (generally three to four rounds; Ertimur and Nondorf 2006; Li and Liu 2017), some IPOs receive none. Thus, substantial variation exists in the number of comment letters received.

[7] Badawi (2017) notes that lawyers likely use databases of prior S-1 filings as a guide when preparing new S-1 filings (also see Bartlett (1999)).

The comment letter process is costly and requires not only company resources to resolve, but also the time and attention of management. We argue that disclosure crowdsourcing by lawyers can help companies provide IPO disclosures that are more in line with SEC standards and expectations. Consistent with this argument, Bozanic et al. (2019) examine the role of lawyers in the comment letter process for 10-K filings and find that securities lawyer involvement is associated with positive disclosure outcomes, such as improved readability, more cautionary language, and fewer restatements and comment letters in future 10-K filings. This discussion leads to our first prediction that disclosure crowdsourcing by IPO Lawyers increases the efficiency of the SEC comment letter process

With respect to IPO pricing, companies are motivated to provide disclosures that are consistent with the expectations of capital providers and crowdsourcing disclosures may also help to ensure that the company's disclosures comply with such expectations. Disclosure crowdsourcing can help lawyers identify and produce the types of IPO disclosures that investors have processed in the past. Consistent with this idea, Drake et al. (2019) find that the acquisition of non-client, peer company financial disclosures by auditors is positively associated with the comparability of the financial disclosures across peer companies. Providing typical disclosures, in standard formats, could potentially improve the speed and efficiency of the price formation process by reducing information processing frictions associated with heterogeneous disclosure. This leads to our second prediction that IPO lawyer disclosure crowdsourcing will be positively associated with speed of price formation post-IPO.

Finally, regarding litigation risk, disclosure crowdsourcing could reduce the likelihood that the IPO disclosures trigger lawsuits if it leads to disclosures generally observed in the IPO market. Cazier, McMullin, and Treu (2019) find that boilerplate risk factor language is associated with more favorable outcomes under judicial and regulatory review, suggesting this crowdsourcing practice may provide protection in courts of law and benefits in the SEC review process. Similar legal benefits may be found in the IPO disclosure setting. Because companies can be sued for stock price declines due to material omissions in the IPO prospectus, Hanley and Hoberg (2012) suggest *enhancing disclosure* as an alternative way to hedge against the likelihood lawsuits related to material omissions.[8] We argue that disclosure crowdsourcing by IPO Lawyers is one way that a company can help to ensure their disclosures are complete, thus reducing their legal exposure. These arguments lead to our third prediction that IPO lawyer disclosure crowdsourcing will be negatively associated with the likelihood of post-IPO litigation.

Our predictions are not without tension as not all instances of external counsel involvement positively affect companies' financial disclosure environment. External law firms may utilize the disclosures of other companies as a template simply because doing so is more efficient than drafting such disclosures from scratch (Badawi 2017). While this approach is efficient, it may not result in an overall increase in useful information if this approach leads to more standardized (Hanley and Hoberg 2010) and less company-specific information (Campbell, Chen, Dhaliwal, Lu, and Steele 2014; Hope Hu, and Lu. 2016). Law firms may also employ disclosure crowdsourcing to protect themselves, not the client, against lawsuits. Malpractice standards for securities counsel promote compliance with a standard of care based on other similarly-situated lawyers (Mallen and Smith, 2004) and so crowdsourcing disclosures may benefit the external legal counsel should a malpractice suit subsequently arise based on such disclosures. It may also be the case that high levels of disclosure crowdsourcing indicates setting where IPO lawyers lack

[8] Some researchers also argue that firms can hedge against this risk by purposefully underpricing the offering (Ibbotson 1975; Tinic 1998). However, others are skeptical of the idea that underpricing is an effective hedge against lawsuits (see, e.g., Ritter and Welch 2002) and Hughes and Thakor (1992) develop a model which shows that litigation risk alone is not sufficient to produce underpricing.

experience or expertise, and thus, need to rely more heavily on the work of others in the drafting process.

It may also be the case that disclosure crowdsourcing results in less transparent client disclosures making such disclosures more difficult for the market to digest. For example, Dechow and Tan (2017) find that the practice of companies avoiding reporting compensation expense by engaging in stock option backdating likely spread through shared external legal counsel. They predict that executives engaged in backdating because they were desensitized to its inappropriateness when they learned through their external legal counsel that other companies were engaging in this practice. While the mechanism is different in our setting, if disclosure crowdsourcing is used to provide generalized rather than specific disclosures during the IPO process then the practice may contribute to slower price formation in the post IPO period.

3. Data and Variables

3.1 Data and Sample

We begin by creating a dataset that links companies' IPOs to their IPO Lawyers using data compiled from S-1 filings in the Lexis Securities Mosaic Law Firm Relationships database. These data were obtained from the S-1 filings themselves since S-1 filings include the contact information for the company to receive subsequent correspondence from the SEC including the SEC comment letters. The Mosaic dataset provides the name of the filer, the name of the advising law firm, and the accession number and type of filing that the law firm helped to prepare.

Next, we construct a dataset that tracks each time these law firms access (download or view) an S-1 or S-1/A filing in EDGAR.[9] These data are available from the EDGAR server that

[9] The term "EDGAR" is an acronym for Electronic Data Gathering, Analysis, and Retrieval. The SEC describes and provides access to EDGAR at the following location: https://www.sec.gov/edgar/searchedgar/companysearch.html.

tracks all user activity in the database. Specifically, the log provides a partial anonymized IP address for the user, the date and time of their request, the Central Index Key (CIK) of the SEC registrant, and a link to the filing being requested. While the SEC anonymizes the final octet of the IP addresses, we follow prior literature (Bozanic et al. 2017; Li et al. 2017; Drake et al. 2019) and identify the owner of the IP addresses by matching the first three octet in the EDGAR log file with the first three octet of the IP addresses listed in the Registry for Internet Numbers (ARIN). This matching process is appropriate because IP addresses are typically purchased in blocks that can be uniquely identified using the first three octet only. We then manually reviewed each IP address ownership to identify law firms with data available on both Lexis Securities Mosaic and EDGAR. The intersection of our IPO database with the EDGAR database results in a sample of 51 law firms during the 2003 to 2017 period.[10] This process eliminates plaintiff law firms and law firms that are not involved in IPOs and ensures that our sample is a homogenous group of law firms that both prepares S-1 filings and uses EDGAR.

To investigate the effects of disclosure crowdsourcing on post-IPO outcomes, we merge in data from several additional sources. We obtain IPO-related data from SCD, financial statement data from Compustat, and SEC comment letter data from Audit Analytics. We use data from CRSP to construct our intra-period timeliness measure. We also obtain data on the prestige of the lead underwriter for the IPO from Jay Ritter's webpage. Finally, the litigation data that we use comes from the Securities Class Action Clearinghouse. Merging these six datasets into the IPO and EDGAR datasets results in a final sample of 43 law firms that advise 332 unique IPOs for a total of 408 IPO-law firm observations. This is the main sample for our analysis.[11]

[10] We begin our sample period in 2003 because that is the first year in which EDGAR web traffic data is available.
[11] When we compare our sample of IPO companies to IPO companies not in our sample (with GVKEYs in Compustat), we find that the IPO proceeds from the offerings in our sample are larger. The average proceeds in our sample is $302

3.2 Disclosure Crowdsourcing Variable

Our measure of disclosure crowdsourcing is based on the number of times that the law firm advising the IPO accesses the S-1 and S-1/A filings of other companies in the EDGAR system before the initial S-1 is filed. Several conversations with lawyers involved with IPOs suggest that the disclosure crowdsourcing of other S-1 filings begins approximately 8 to 10 weeks prior to the initial S-1 filing. We define our primary variable of interest, *S-1 Prep Crowdsourcing*, as the natural logarithm of the number of times that an IP address owned by company *i*'s lawyer accessed a S-1 or S-1/A filing from another company during the 8-week period that ends on the initial S-1 filing date.[12]

3.3 SEC Response variables

We examine two constructs that capture different dimensions of the SEC's response to the S-1 filing. The first construct relates to the extent of the SEC's concerns with the initial S-1. We capture this construct using the number of issues identified in their comment letter. More specifically, we define *Phrase Count* as the natural log of one plus the number of issues identified by the SEC in their comment letter to the S-1 filing as recorded in the Audit Analytics dataset. The second construct relates to the speed with which the company is able to address the SEC's concerns with the IPO process, which we measure using two proxies. We define *Conv Count* as the natural log of one plus the number of amended S-1s the company files between the initial S-1 filing and the IPO date as recorded in the Audit Analytics dataset. We define *Time* as the natural logarithm

million, compared to $211 million for the IPOs not in our sample. Thus, we acknowledge that a large offering bias could weaken the generalizability of our results.

[12] Specifically, we use the natural log of 1 plus the total number of daily unique S-1 or S-1/A EDGAR downloads from the IPO firm's legal counsel for the 56 days prior to the IPO firm's initial S-1. Our results are qualitatively similar if we use the total number of downloads.

of one plus the number of days between the initial S-1 filing and the IPO date. Both proxies capture how quickly the company is able to satisfy the concerns of the SEC, where a lower value implies a quicker resolution of the SEC's concerns.

3.4 Control variables

Disclosure crowdsourcing is only one of many activities done by the company's legal counsel during the IPO process. To help ensure that our disclosure crowdsourcing variables are not capturing variation in the overall effort of legal counsel, we include the variable *Legal Fees* in our analysis. *Legal Fees* is the natural logarithm of the total dollar amount of legal fees as reported by in the company in their final S-1/A.

We also include a broad set of control variables that capture different dimensions of the company itself. These controls include measures of size (*Size*), financial leverage (*Leverage*), accounting performance (i.e., *ROA*), and asset efficiency (*Sales-to-Assets*). We measure all of these variables in the quarter prior to the IPO.

It is possible for a company to engage multiple law firms to be legal counsel during the IPO process. Because disclosure crowdsourcing incentives might differ if there are multiple law firms engaged in the IPO process, we include an indicator variable, *Multiple Law Firms*, set equal to one if the company has more than one law firm on its S-1 filings and both of those law firms have S-1 download activity in the EDGAR log files, and set to zero, otherwise.

We follow Billings and Lewis-Western (2016) in selecting additional control variables for the IPO itself. We include five variables that capture different dimensions of IPO quality. We include an indicator variable, *VC*, equal to one if the IPO has venture capital backing. We also include the percentage of the company's shares retained by insiders after the IPO (*Overhang%*), the proportion of secondary shares sold in the issue (*Sec %*), and the company's age from its founding date (*Age*). All of these variables come from Thomson Reuters SDC Platinum database. Finally, we include the reputation of the lead underwriting bank for the IPO issue from Jay Ritter's webpage (*UW Rank*). We provide formal variable definitions in Appendix A.

Finally, we include three dimensions of fixed effects in our models. We use industry fixed effects measured at the Fama-French 10 Industry Classification to control for time-invariant characteristics that could jointly affect our outcome variables and the amount of disclosure crowdsourcing done by legal counsel. We include year of IPO fixed effects in order to control for general time trends in the use of EDGAR and the number of IPOs generally. We also use law firm fixed effects because we are interested in the how the level of disclosure crowdsourcing affects the SEC's response and the speed of price formation once the company is public. Including law firm fixed effects allows us to subsume unobservable, time-invariant law firm characteristics, such as law firm reputation, resources, quality, or experience that could attenuate the relation between these constructs.

3.5 IPO Outcome variables

Speed of price formation

We measure the speed of price formation after the IPO using the intraperiod timeliness (IPT, hereafter) metric used in prior research (Butler et al. 2007; Bushman et al. 2010; Twedt 2016). The IPT metric uses an area-under-the-curve methodology to capture the speed with which a particular information event is impounded into price over a specified period. Our measurement window begins on the IPO date (day 0) and continues for 20 trading days (approximately one month). To measure IPT, we first calculate the daily proportion of abnormal returns realized up to and including a given trading day. Then, for each day in a measurement window, we calculate the cumulative buy-and-hold abnormal return from day 0 to that day, scaled by the cumulative buy-and-hold abnormal return for the entire 20-day window. More specifically, for each day, we compute the proportion of the total-period abnormal return that accumulated up to that day. This means that at the end of the event window, the proportion of the realized return during the measurement window will be equal to one. We then plot the daily proportion of abnormal returns and calculate the area under the curve for each IPO observation as follows:

$$IPT = \tfrac{1}{2} \sum^{20}_{t=0} (ABRET_{t-1} + ABRET_t)/ABRET_{20} = \sum^{19}_{t=0} (ABRET_t/ABRET_{20}) + 0.5. \qquad (1)$$

A larger area under the curve indicates a faster (i.e., more efficient) price formation process post-IPO. Prior research generally ranks the IPT measure in order to reduce the influence of noise and outliers. The measure we include in our models, *Rank IPT*, is the ranked value of IPT constructed such that higher ranks indicates a faster price formation process. It is important to note that the IPT measures is designed to measure the speed of price formation and not the overall magnitude of the price change over a specified period. The denominator in the metric is the overall return over the period, which effectively controls for this overall price change.

Post-IPO litigation

We identify shareholder litigation against companies using data from the Stanford Law School/Cornerstone Research Securities Class Action Clearinghouse. This database contains all securities class actions filed in federal courts starting in 1996. As noted in Billings and Lewis-Western (2016, p. 381), “IPO firms encounter added legal concerns associated with lawsuits filed under section 11 of the Securities Act of 1933 … shareholders can (and do) target IPO managers under section 11’s relaxed legal requirements, where a section 10b-5 claim may be dismissed”. Section 11 claims must be filed within three years from the IPO date (Schiller and Murage 2001). Accordingly, we set an indicator variable, *Litigation*, equal to one if the Stanford Law

School/Cornerstone Research Securities Class Action Clearinghouse has a record of a class lawsuit against the company within three years of the IPO date, and to zero otherwise.

4. Disclosure Crowdsourcing Activity around Initial S-1 Filings

We first examine the extent of disclosure crowdsourcing activity by IPO Lawyers. To do this, we estimate a regression where the dependent variable is the natural logarithm of the number of weekly S-1 downloads by IPO Lawyers. The panel of data we use to estimate this model is organized in event time relative to the initial S-1 filing in week zero. The variables of interest are denoted *D(x)*, where *x* represents the number of weeks from the initial S-1 filing. *D(x)* is an indicator variable that is set equal to one in the week of *x*, and to zero otherwise. For example, when the initial S-1 filing is 8 weeks away *D(-8)* is equal to one. We omit the indicator for *D(-26)* from the regression so that the coefficients on these indictor variables are interpreted relative to the first week of the time-series. It is important to note that the sample is aggregated in event time such that each law firm only has one observation per week and that observation is the aggregated amount of S-1 downloads for all the IPOs that the law firm engages in during our sample period.

Table 1 presents the results. We note that a clear pattern emerges where there is no statistical difference in the number of weekly downloads until 13 weeks prior to the initial S-1 filing. At that point, we observe a marked increase in the number of downloads that is statistically significant in all but three of the weeks that occur prior to the initial S-1 filing. In terms of economic significance, the results indicate that about 8 weeks prior to the S-1 filing of their client, IPO Lawyers increase downloads of the S-1 filings of other companies by about 25.8 percent ($e^{0.23} - 1 = 0.258$). These results are consistent with the disclosure crowdsourcing being part of the S-1 filing drafting process.

Figure 1 provides additional descriptive evidence of whether there is more disclosure crowdsourcing when a law firm is engaged to be legal counsel on an IPO. This figure displays the weekly average unique S-1 downloads for two groups of law firms. The solid, blue line in the figure depicts the downloads of IPO Lawyers. The red line depicts the downloads of control law firms (non-IPO Lawyers). We impose two criteria to be selected as a control law firm. First, we require that the law firm prepare S-1 filings but does not file an initial S-1 for the 44 weeks surrounding the initial S-1 that the IPO Lawyers file. Second, we take the closest match based on the prior calendar year's EDGAR download activity with replacement. This matching process ensures that the both the IPO Lawyers and the non-IPO Lawyers are similar in their level of disclosure crowdsourcing in the prior year, but that there is a difference in disclosure crowdsourcing incentives. The vertical, dotted green line depicted in the Figure 1 represents the initial filing of an S-1 in preparation for an IPO for the IPO Lawyers. We note that there is no distinguishable difference in the disclosure crowdsourcing between our IPO and non-IPO law firms until about ten weeks from the initial S-1 filing. We also observe a clear spike in disclosure crowdsourcing for IPO Lawyers, relative to the non-IPO Lawyers, that begins about 6 weeks prior the initial S-1 filing. The magnitude of this spike represents about a 20 percent increase and it peaks about one week before the initial S-1 filing.

After the initial S-1 filing, the SEC typically responds with a comment letter stating their concerns about the S-1 filing. The vertical, dotted blue line depicted in Figure 1 represents the average time to the SEC's first comment letter on the initial S-1 filing which averages 17 days from the initial S-1 filing in our sample. We note that the disclosure crowdsourcing again temporarily increases after the SEC comment letter for our IPO Lawyers.

Overall, the evidence suggests that engaging in an S-1 filing for an upcoming IPO is associated with an increase in S-1 download activity on EDGAR, consistent with greater disclosure crowdsourcing activity. Furthermore, the pattern of marked increases and decreases is consistent with our anecdotal evidence from securities attorneys on how and when they use other companies' S-1 filings in their efforts to prepare filings for an IPO.

5. Disclosure Crowdsourcing and IPO Outcomes

5.1 Sample Selection

We next examine whether the disclosure crowdsourcing by IPO Lawyers is associated with important IPO-related outcomes. We require data from several sources for observations to be included in these analyses. We begin with the intersection between the EDGAR log files and the relationship data between legal counsel and publicly traded companies provided by Lexis Nexis Mosaic. We then merge in IPO data from Compustat/CRSP, SEC comment letter data from Audit Analytics, and IPO data from SDC. These data requirements generate some sample attrition and so our sample for the main empirical analyses comes from 43 law firms that advise 332 IPOs for the 2003 to 2017 period. There are instances where multiple law firms are the legal counsel for the same IPO and so the main sample is 408 IPO-law firm observations.

5.2 Descriptive Statistics

Table 2, Panel A presents descriptive statistics for our main sample. The mean of *S-1 Prep Crowdsourcing* is 726. This indicates that the IPO law firm is downloading approximately 13 unique S-1 or S-1/A filings per day during the 8 weeks prior to the initial S-1 filing. We find that our sample includes 41 law firm-IPO observations (10% of sample firms) that have class-action litigation within the three years after the IPO. We also find that the SEC typically finds several issues in the initial S-1 filing that require further attention by management. For example, on average the SEC discusses 15.1 issues in their comment letter on the initial S-1 filing. It then generally takes companies several rounds of amendments, 7.3 on average, to get final approval. We also find that there is approximately 83 days on average between the initial S-1 filing date and the IPO date. Finally, the companies in our sample spend $1.85 million in legal fees and have $2.98 billion in assets, $0.34 million in sales, and negative ROA prior to the IPO, on average.

Table 2, Panel B presents disclosure crowdsourcing behavior based on the magnitude of legal fees (quintiles are presented) to assess the degree to which the level of due diligence reflected in the practice of disclosure crowdsourcing is priced into fees. The smallest quintile includes those observations where the total legal fees for the IPO was $1 million or less. Law firms in this quintile download an average 489 unique filings in the S-1 Prep Period. Law firms in the highest quintile, where the average legal fee for the IPO is $3.8 million, download an average of 740 unique S-1 filings during the S-1 Prep Period.[13]

5.3 Univariate correlations

Table 3 presents Pearson and Spearman univariate correlations. We find that *S-1 Prep Crowdsourcing* is positively correlated with *S-1 Review Crowdsourcing*. We find that *S-1 Prep Crowdsourcing* is negatively correlated with the three proxies for the efficiency of the SEC review (*Time*, *Conv Count*, and *Phrase Count*), though the correlation is only statistically significant with *Phrase Count.* We also find that it is positively, but not significantly, correlated with *IPT* and *Litigation*. Finally, we note that the degree of correlation among the other variables in the table do not suggest that multicollinearity is likely to be an issue in our models.

5.4 Primary tests of our predictions

[13] An untabulated test indicates that the difference in legal between between the top and bottom quintile is statistically significant.

We next examine our predictions that disclosure crowdsourcing by IPO Lawyers is associated with positive outcomes using the following OLS model:

$$\begin{aligned} IPO\ Outcome = {} & \beta\ S\text{-}1\ Prep\ Crowdsourcing + \gamma\ Controls + \mu\ Law\ Firm\ FE \\ & + \theta\ Year\ FE + \delta\ Industry\ FE + \varepsilon \end{aligned} \tag{2}$$

where *IPO Outcome* is one of five alternative dependent variables. The first three outcome variables *Phrase Count, Conv Count,* or *Time*, relate to the SEC's response to the S-1. The fourth variable, *Rank IPT20*, captures the speed of price formation after the IPO. The fifth variable, *Litigation*, indicates whether the IPO results in a lawsuit. The vector (γ) of control variables includes all variables discussed in Section 3.4. We also include law firm, year, and IPO industry fixed effects. More details on these measures and their data sources are provided in Appendix A.

Table 4 presents the estimation results for model (2) using the three alternative SEC response variables. In column (2), we find a negative and significant (at the 1% level) relation between *S-1 Prep Crowdsourcing* and one of the three proxies, *ConvCount*. This suggests that SEC review process requires fewer rounds of amendments to complete when the IPO Lawyers engage in more disclosure crowdsourcing. We also note that *S-1 Prep Crowdsourcing* is one of only two variables with significant coefficients in column (2). *Legal Fees,* which also potentially captures the extent of lawyer due diligence in the IPO process is not significant nor are several broad firm characteristic such as firm size, leverage and sales. This highlights the unique explanatory power of disclosure crowdsourcing on the SEC review process.

In columns (1) and (3), we find no evidence that the practice of disclosure crowdsourcing is significantly related to the number of issues identified by the SEC (*Phrase Count)* or the overall time it takes to address and resolve the issues (*Time*). Thus, we find modest support for our first prediction that IPO Lawyers that engage in more crowdsourcing are better able to help their clients resolve comment letter issues with the SEC.

We next examine our second prediction, which is that IPO lawyer disclosure crowdsourcing is positively associated with speed of price formation post-IPO. Table 5 presents the model (2) estimation results when *Rank IPT20* is the dependent variable. Consistent with our prediction, we find a positive and highly significant coefficient (at the 1% level) on *S-1 Prep Crowdsourcing*. This suggests that when IPO Lawyers engage in higher levels of disclosure crowdsourcing that prices move more quickly to a post-IPO price measured one month later. This finding is consistent with the notion that IPO Lawyer crowdsourcing produces disclosure that better meet the needs and expectations of investors. We also find a positive and significant coefficient on the *D(Multi Law Firms*). Here again, we find that *S-1 Prep Crowdsourcing* is one of only a few significant explanatory variables in the model.

Finally, we examine our third prediction, which is that IPO lawyer disclosure crowdsourcing will be negatively associated with the likelihood of post-IPO litigation. Given the dichotomous nature of the dependent variable, and our inclusion of various fixed effects (year and industry), we test this prediction using a linear probability model (model 2).[14] In untabulated tests, we confirm that our inferences are unchanged if we estimate the model using a logit model.

We present the model (2) estimation results using *Litigation* as the dependent variable in Table 6. Consistent with our prediction, we find a negative and significant (at the 10% level) relation between *S-1 Prep Crowdsourcing* and post-IPO litigation, indicating that disclosure crowdsourcing prior to the IPO is associated with fewer subsequent shareholder lawsuits. In terms

[14] We do not include law firm fixed effects in these analyses for two reasons. First, logit and probit models have known issues with including fixed effects of this kind. And second, given that only 10% firms in our sample has at least one incidence of shareholder litigation, the inclusion of law firm fixed effects would results in low power test using variation from only a very small number of sample IPOs.

of economic significance, we find that a one standard deviation change in *S-1 Prep Crowdsourcing* is associated with about a 5 percent reduction in the likelihood of subsequent litigation, which represents about a 50% reduction relative to the sample mean. This finding is consistent with disclosure crowdsourcing significantly reducing companies' exposure to future lawsuits.

5.5 Additional analyses

It is important to acknowledge that the IPO process has a clearly defined chronology that is not considered in our primary tests. The S-1 preparation period happens prior to the initial S-1 filing. Then, the SEC reviews the S-1 and responds with concerns. The company then addresses the concerns and files an Amended S-1 and the process repeats until the S-l is approved. Finally, the company goes public. At that point, the stock begins trading and the possibility of IPO-related litigation exists. In this subsection, we consider whether the defined chronology just described influences our results.

Our primary predictions focus on the IPO lawyer disclosure crowdsourcing during the S-1 preparation period. However, as illustrated in Figure 1, IPO Lawyers also engage in the practice *during* the SEC review period, which is the period between the initial S-1 filing and the IPO date. The extent of disclosure crowdsourcing during the SEC review period could potentially be influenced by the SEC's response to the initial S-1 filing. Further, the extent of review period disclosure crowdsourcing and the SEC's response to the S-1 could each potentially affect the other outcome variables we examine (price formation and litigation).

We conduct two sets of additional analyses to examine these possibilities. In the first set of tests, we re-estimate model (2) after controlling for the natural logarithm of the number of times that an IP address owned by company *i*'s lawyer accessed a S-1 or S-1/A filing from another company during the SEC review period. We label this variable *S-1 Review Crowdsourcing*.

Including *S-1 Review Crowdsourcing* in our analysis has two benefits. First, we view the S-1 Review period as distinct from the S-1 Preparation period because it is likely that lawyers' disclosure crowdsourcing during the review period is targeted to the specific issues that the SEC identifies in their initial comment letter. Second, disclosure crowdsourcing in the preparation and review period are likely to be positively correlated with one another and information gained in the initial search will likely aid IPO Lawyers in their post-initial S-1 response.

We present the results in Table 7 for all five dependent variables. The results are similar to those presented in Tables 4 through 6 with one exception. As presented in column (3), we now find a negative and significant association between *S-1 Prep Crowdsourcing* and the length of time it takes to resolve the comments from the SEC (*Time*). We continue to find that *S-1 Prep Crowdsourcing* is negatively associated with the number of S-1 amendments that are filed (column 2) and post-IPO litigation (column 5), and positively associated with the speed of price formation post-IPO (column 4).

In the second set of tests, we more formally model the sequence of IPO-related outcomes using a system of equations. As discussed previously, disclosure crowdsourcing during the S-1 preparation period influences the SEC response, which in turn, can influence the level of disclosure crowdsourcing employed to respond to the SEC comments during the review period. All of these factors can then potentially impact the post-IPO outcomes related to price formation and litigation. Figures 2 and 3 present the many possible direct and indirect links between these various constructs using *Rank IPT20* and *Litigation*, respectively. We formally model each of the paths depicted in Figures 2 and 3 using a structural equation model (SEM) that includes the control variables and that allows the error terms in each model to be correlated. Figures 2 and 3 also present the standardized path coefficients. We provide the full SEM estimation results for *Rank IPT20* and *Litigation* in Tables 8 and 9, respectively.

The results reveal two key findings. First, we find that the direct path between *S-1 Prep Crowdsourcing* and *Rank IPT20* (Figure 2) or *Litigation* (Figure 3) remain significant using the SEM. Second, we find no evidence of any significant indirect paths between *S-1 Prep Crowdsourcing* and the two post-IPO outcomes that operate via the SEC review process.

6. Conclusion

Lawyers play an important advisory role in the IPO process, but little is understood about their role and the methods employed in providing such advisory services. In order to advise on the extent, nature, and content of any particular financial disclosure, IPO Lawyers may examine the S-1 filings of other companies who recently went public to identify disclosures norms and best practices. We term this practice “disclosure crowdsourcing” and provide evidence on both extent of this behavior and its association with IPO-related outcomes.

First, we document an increase in “disclosure crowdsourcing” that begins approximately ten weeks prior to the IPO’s first S-1 filing, consistent with IPO Lawyers conducting crowdsourcing to assist their clients in the preparation of IPO prospectuses. Second, we examine the association between disclosure crowdsourcing and IPO-related outcomes. We find that disclosure crowdsourcing is positively associated with the efficiency of interactions with the SEC (as measured by the number of interactions with the SEC) and the speed of price formation following the IPO and is negatively associated with the likelihood of subsequent shareholder securities litigation. This evidence suggests that companies experience several favorable IPO-related outcomes when their legal counsel processes more financial reporting data prior to the issuance of the first S-1.

This study contributes to our understanding of the IPO process, the role of disclosures in this process, and the influence of various parties that assist companies in this process. Several prior studies provide evidence on the relation between disclosure and IPO pricing (Beatty and Ritter 1986; Ljungqvist and Wilhelm 2003; Schrand and Verrecchia 2005; Leone et al. 2007), but evidence is sparse on the role that external consultants such as lawyers play in the process (Beatty and Welch 1996; Hanley and Hoberg 2010). We contribute to this line of research by providing evidence that the information-gathering activities of lawyers in the IPO process are associated with real, economic outcomes to their clients. We also contribute to the growing literature on the role of external legal counsel on financial reporting in general (Hanley and Hoberg 2010, 2012, Dechow and Tan 2017; Bozanic et al. 2019). We highlight the role of crowdsourcing as a channel through which external law firms obtain knowledge to help their clients.

Lastly, our study contributes to the emerging literature that examines the practice of disclosure crowdsourcing (see, e.g., Drake et al. 2019) and provides information regarding the use of financial disclosures by external legal counsel. These findings should be of particular interest to standard setters and regulators given the important interactions of law firms with capital market participants.

References

Akerlof, G. 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics* 84(3): 488-500.

Ball, R., and L. Shivakumar. 2008. Earnings quality at initial public offerings. *Journal of Accounting and Economics* 45 (2): 324–49.

Badawi, A. 2017. Lawyers, Law Firms, and the Production of Legal Information. Working paper. https://www.law.berkeley.edu/wp-content/ uploads/2017/10/Badawi_Law_Firms.pdf

Bartlett, J. 1999. *Fundamentals of Venture Capital*. Madison Books.

Beatty, R.P., and J.R. Ritter. 1986. Investment banking, reputation, and the underpricing of initial public offerings. *Journal of Financial Economics* 15(1-2): 213-232.

Beatty, R. P., and Welch, I.. 1996. Issuer expenses and legal liability in initial public offerings. The Journal of Law and Economics 39.2: 545-602.

Billings, N., and M. Lewis-Western. 2016. When Does Pre-IPO Financial Reporting Trigger Post-IPO Legal Consequences? *Contemporary Accounting Research* 33(1): 378-411.

Bird, R.C., P.A. Borochin, and J.D. Knopf. 2014. The value of the chief legal officer to the firm. *Journal of Corporate Finance*, forthcoming.

Bozanic, Z., Choudhary, P. and Merkley K. 2019. Securities Law Expertise and Corporate Disclosure. *The Accounting Review,* forthcoming.

Bozanic, Z., J.L. Hoopes, J.R. Thornock, and B.M. Williams. 2019. IRS Attention. *Journal of Accounting Research* 55(1): 79.

Bushman, R., A. Smith, and R. Wittenberg-Moerman. 2010. Price discovery and dissemination of private information by loan syndicate participants. *Journal of Accounting Research* 48: 921-972.

Butler, M., A. Kraft, and I. Weiss. 2007. The effect of reporting frequency on the timeliness of earnings: The cases of voluntary and mandatory interim reports. *Journal of Accounting and Economics* 43: 181-217.

Bradford, J. and M. Bogdon. 2017. IPO Accounting and Legal Fees: 2016 Update. Audit Analytics. Available at https://blog.auditanalytics.com/ipo-accounting-and-legal-fees-2016-update/ (last visited 1/3/2019).

Campbell, J., H. Chen, D. Dhaliwal, H. Lu, and L. Steele. 2014. The Information Content of Mandatory Risk Factor Disclosures in Corporate Filings. *Review Of Accounting Studies* 19(1):396-455.

Cazier, R., McMullin, J., and Treu, J. 2019. Are Lengthy and Boilerplate Risk Factor Disclosures Inadequate? An Examination of Judicial and Regulatory Assessments of Risk Factor Language. Working Paper.

Cornerstone Research. 2019. Securities Class Action Filings: 2019 Year in Review. Boston, MA: Cornerstone Research, Inc.

Dambra, Michael, Laura Casares Field, and Matthew T. Gustafson. 2015. The JOBS Act and IPO volume: Evidence that disclosure costs affect the IPO decision. *Journal of Financial Economics* 116.1 (2015): 121-143.

Drake, M., B. Johnson, D. Roulstone, J. Thornock. 2020. Is there information content in information acquisition. *The Accounting Review*, forthcoming.

Drake, M., P. Lamoreaux, P. Quinn, and J. Thornock. 2019. Auditor benchmarking of client discloures. *Review of Accounting Studies*, 24(2): 393-425.

Dechow, P., and S. Tan. 2017. How do accounting practices spread? An examination of law firm networks and stock option backdating. Working paper.

Ertimur, Y., and M. Nondorf. 2006. IPO firms and the SEC comment letter process. Working Paper.

Fan, Q. 2007. Earnings management and ownership retention for initial public offering firms: Theory and evidence. *The Accounting Review* 82 (1): 27–64.

Gibbons, B., P. Iliev, and J. Kalodimos. 2018. Analyst information acquisition via EDGAR. *Working Paper*, Pennsylvania State University. Hanley, K.W., and G. Hoberg (2010). The information content of IPO Prospectuses. *Review of Financial Studies* 23(7): 2821–64.

Hanley, K.W., and G. Hoberg. 2010. The information content of IPO prospectuses. *Review of Financial Studies* 23: 2821-2864.

Hanley, K.W., and G. Hoberg. 2012. Litigation Risk, Strategic Disclosure and the Underpricing of Initial Public Offerings. Journal of Financial Economics 103(2): 235–54.

Hope, O.K., D. Hu, and H. Lu. 2016. The benefits of specific risk-factor disclosures. Review of Accounting Studies 21(4): 1005-1045.

Hopkins, J., E.L. Maydew, and M. Venkatachalam. 2015. Corporate general counsel and financial reporting quality. *Management Science* 61(1): 129-145.

Hughes, P.J., and A.V. Thakor. 1992. Litigation risk, intermediation, and the underpricing of initial public offerings. *The Review of Financial Studies* 5.4: 709-742.

Kwak, B., B.T. Ro, and I. Suk. 2012. The composition of top management with general counsel and voluntary information disclosure. *Journal of Accounting and Economics* 54(1): 19-41.

Leone, A.J., S. Rock, and M. Willenborg. 2007. Disclosure of Intended Use of Proceeds and Underpricing in Initial Public Offerings. *Journal of Accounting Research* 45: 111-153

Li, B., and Z. Liu. 2017. The oversight role of regulators: evidence from SEC comment letters in the IPO process. *Review of Accounting Studies* 22: 1229-1260.

Ljundqvist, A., and W.J. Wilhelm. 2003. IPO Pricing in the Dot-com bubble. *The Journal of Finance* 58(2): 723-752.

Mallen, R.E., and J.M. Smith. 2004. Legal Malpractice. West Publishing Co. (5th ed. 2000 & 2004 Supp.).

McClane, J. R. 2015. The Sum of Its Parts: The Lawyer-Client Relationship in Initial Public Offerings. *Fordham Law Review* 84: 131.

PWC. 2015. Considering an IPO? An insight into the costs post-JOBS act.

Rogers, J., A. Van Buskirk, and S. Zechman. 2011. Disclosure Tone and Shareholder Litigation. *The Accounting Review* 86 (6): 2155-2183.

Schiller, M. S., and H. W. Murage. 2001. Who has standing to sue under section 11 of the Securities Act? *Securities Reform Act Litigation Reporter*, March

Schrand, C.M., and R.E. Verrecchia. 2005. Information Disclosure and Adverse Selection Explanations for IPO Underpricing. Working Paper.

Sobel, M. E. 1982. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology*, *13*, 290-312.

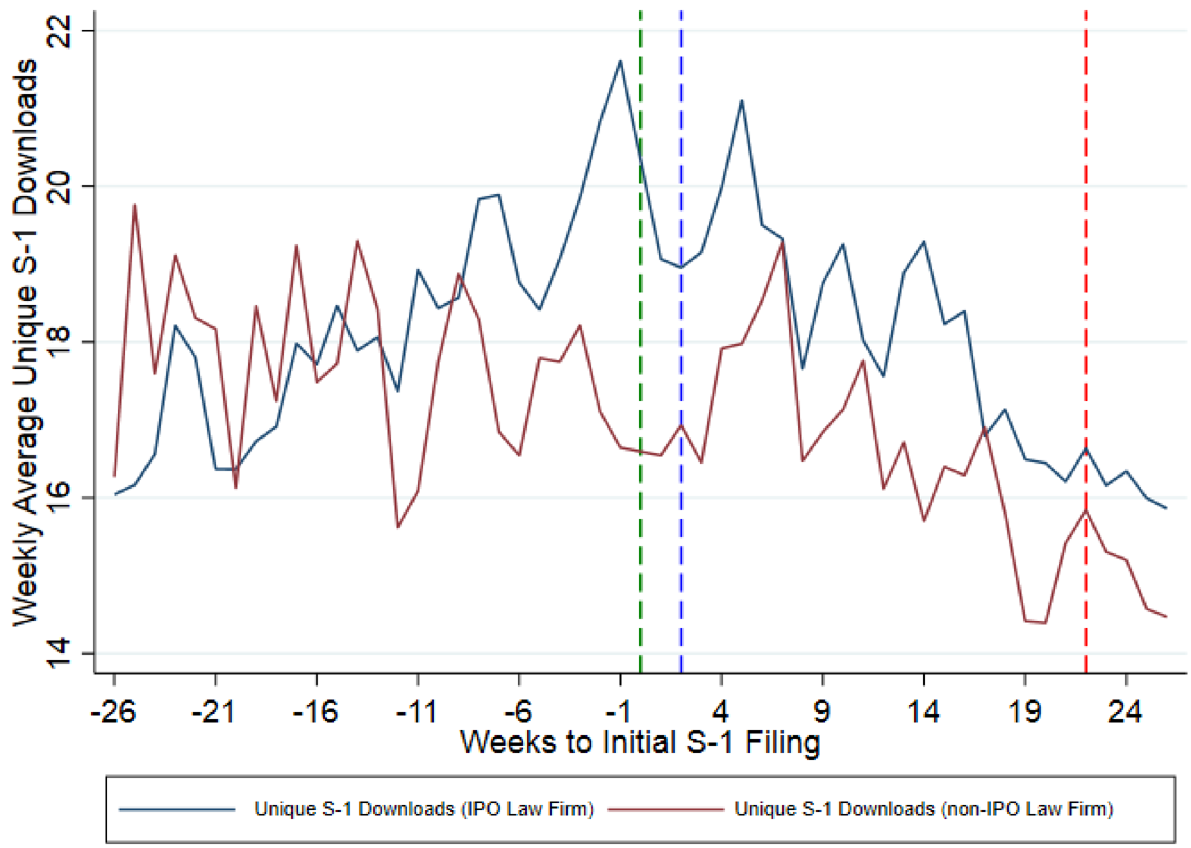
Stice-Lawrence, L. 2017. Too Small to Regulate: International Evidence on Regulatory Monitoring. Working Paper.

Twedt, B. 2016. Spreading the word: price discovery and newswire dissemination of management earnings guidance. *The Accounting Review* 91(1): 317-346.

Appendix A: Variable definitions

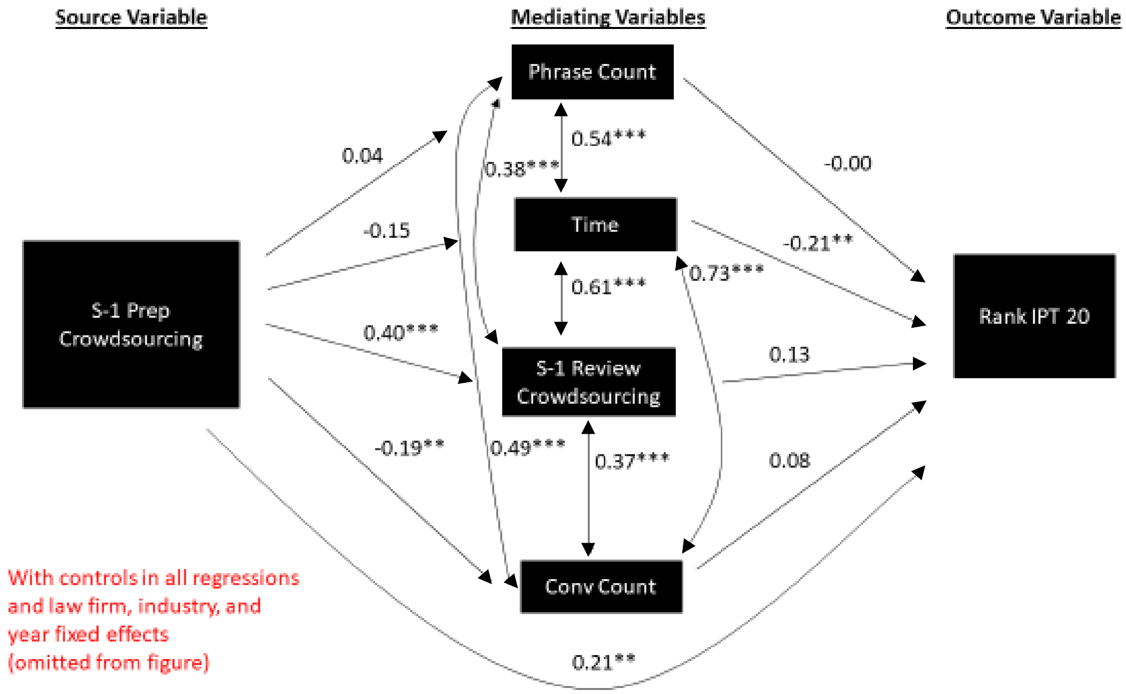
Variable Label	Definition
Outcome Variables	
Rank IPT 20	The raw IPT value is the daily proportion of size-adjusted abnormal returns realized up to and including a given day, starting on the IPO date and continuing through day 20. For each day, we calculate the cumulative buy-and-hold return from IPO day to that day, scaled by the cumulative abnormal return for the entire 20 days. We then estimate the area under this curve, thus larger area indicates that information is more quickly impounded into price. We then rank all IPT values where a higher rank is associated with higher IPT values. (Source: CRSP)
D(Litigation)	An indicator equal to one if the IPO firm is sued within three calendar years of the IPO date, zero otherwise (Source: Stanford Law School/Cornerstone Research Securities Class Action Clearinghouse)
Source Variables	
S-1 Prep Crowdsourcing	The natural log of 1 plus the total number of daily unique S-1 or S-1/A EDGAR downloads from the IPO firm's legal counsel for the 56 days prior to the IPO firm's initial S-1. (Source: EDGAR log files)
Mediating Variables	
S-1 Review Crowdsourcing	Natural log of 1 plus the total number of daily unique S-1 or S-1/A EDGAR downloads from the IPO firm's legal counsel beginning on the day after the IPO firm's initial S-1 and ending on the day of the IPO. (Source: EDGAR log files)
Phrase Count	Natural log of one plus the number of phrases (issues) identified by the SEC in their response to the initial S-1 filing. (Source: Audit Analytics)
Conv Count	Natural log of one plus the number of amended S-1s the IPO firm files between the initial S-1 filing and the IPO date. (Source: Audit Analytics)
Time	Natural log of one plus the number of days between the initial S-1 filing date and the IPO date. (Source: Audit Analytics)
Control Variables	
D(Multi-Lawfirms)	An indicator variable equal to one if the IPO firm has more than one law firm on its S-1 filings and both of those law firms have S-1 download activity in the EDGAR log files. (Source: EDGAR log files)
Size	Natural logarithm of total assets in the quarter prior to the IPO. (Source: Compustat)
Leverage	Ratio of total liabilities to total assets in the quarter prior to the IPO. (Source: Compustat)
ROA	Ratio of net income to total assets in the quarter prior to the IPO. (Source: Compustat)
Sales	Ratio of total sales to total assets in the quarter prior to the IPO. (Source: Compustat)
D(VC)	Indicator variable equal to one if the IPO has venture capital backing. (Source: SDC)
Overhang %	Percentage of the firm retained by insiders after the IPO. (Source: SDC)
Sec %	Proportion of secondary shares sold at the IPO. (Source: SDC)
UW Rank	Rank of the lead underwriting bank's reputation. (Source: Jay Ritter's website - https://site.warrington.ufl.edu/ritter/ipo-data/)
Age	Natural log of one plus the number of years between the firm's founding date and the IPO date. (Source: SDC)

Figure 1 – Disclosure Crowdsourcing Patterns



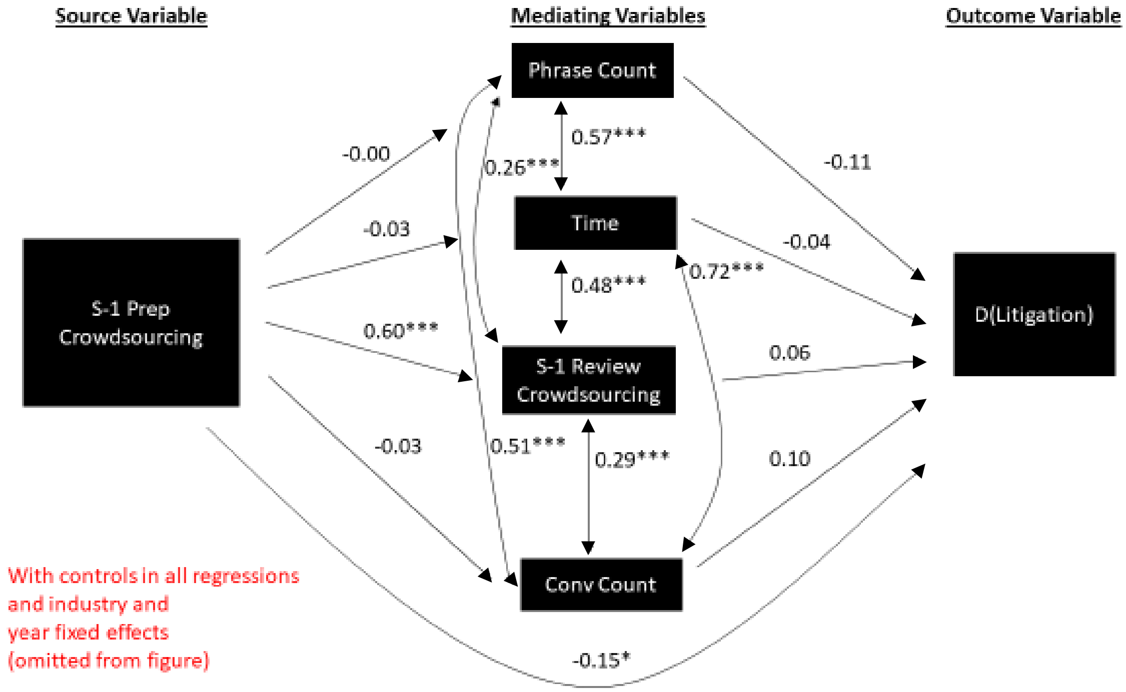
This figure displays the weekly average unique S-1 downloads for two groups of law firms. The blue represents those law firms that are engaged to be the legal counsel for an upcoming IPO. The red line represents those law firms that are not engaged to be the legal counsel for an upcoming IPO but have been legal counsel for an IPO in the past and have identifiable download activity on the EDGAR log database. The vertical, dotted green line represents the initial filing of an S-1 in preparation for an IPO. The vertical, dotted blue line represents the average time to the SEC's first comment letter on the initial S-1 filing which is on average 17 days after the initial S-1 filing. The vertical, dotted red line represents the average time from the initial S-1 filing to the actual IPO date which on average is 154 days after the initial S-1 filing.

Figure 2 – Path Analysis of Disclosure Crowdsourcing on Speed of Price Formation



The figure shows the results of the path analysis of the direct and indirect association between disclosure crowdsourcing during the preparation period and the speed that information is impounded into price during the first 20 days after the IPO. We estimate a structural equation model of the direct effect of disclosure crowdsourcing in the preparation period on the market based measure of the speed of price formation as well as the indirect effects of disclosure crowdsourcing in the preparation period through the number of issues identified by the SEC in their response to the initial S-1 (*Phrase Count*), the length of time of the SEC review process (*Time*), the number of S-1 amendments (*Conv Count*) and the S-1 download activity in the review period (*S-1 Review Crowdsourcing*). The equations in the structural equation model include a regression of the outcome variable (*Rank IPT 20*) on the mediating variables (*Phrase Count*, *Time*, *S-1 Review Crowdsourcing*, *Conv Count* and all control variables), and regressions of each mediating variable *Phrase Count, Time, S-1 Review Crowdsourcing*, and *Conv Count* on the source variable, *S-1 Prep Crowdsourcing.* We present the standardized path coefficient with ***, **, * indicating significance at the 1%, 5%, and 10% levels, respectively.

Figure 3 – Path Analysis of Disclosure Crowdsourcing on Post-IPO Litigation



The figure shows the results of the path analysis of the direct and indirect association between disclosure crowdsourcing during the preparation period and the likelihood of litigation during the first three years post-IPO. We estimate a structural equation model of the direct effect of disclosure crowdsourcing in the preparation period on a litigation indicator as well as the indirect effects of disclosure crowdsourcing in the preparation period through the number of issues identified by the SEC in their response to the initial S-1 (*Phrase Count*), the length of time of the SEC review process (*Time*), the number of S-1 amendments (*Conv Count*) and the S-1 download activity in the review period (*S-1 Review Crowdsourcing*). The equations in the structural equation model include a regression of the outcome variable (*D(Litigation)*) on the mediating variables (*Phrase Count*, *Time*, *S-1 Review Crowdsourcing*, *Conv Count* and all control variables), and regressions of each mediating variable *Phrase Count*, *S-1 Review Crowdsourcing*, and *Conv Count* on the source variable, *S-1 Prep Crowdsourcing.* We present the standardized path coefficient with ***, **, * indicating significance at the 1%, 5%, and 10% levels, respectively.

Table 1 –Analysis of download activity for IPO law firms

Dependent Variable: Log(Total S-1 Downloads)					
Pre S-1 Indicators			Post S-1 Indicators		
Variable	**β**	**T-stat**	**Variable**	**β**	**T-stat**
D(-25)	0.0083	(0.10)	D(0)	0.1948**	(2.43)
D(-24)	0.0691	(0.86)	D(1)	0.1574**	(1.96)
D(-23)	0.0986	(1.23)	D(2)	0.1803**	(2.25)
D(-22)	-0.0305	(-0.38)	D(3)	0.2397***	(2.99)
D(-21)	-0.0314	(-0.39)	D(4)	0.1961**	(2.44)
D(-20)	-0.0691	(-0.86)	D(5)	0.1817**	(2.26)
D(-19)	-0.0440	(-0.55)	D(6)	0.2318***	(2.89)
D(-18)	0.0338	(0.42)	D(7)	0.2170***	(2.70)
D(-17)	0.0968	(1.21)	D(8)	0.2670***	(3.33)
D(-16)	0.0191	(0.24)	D(9)	0.1870**	(2.33)
D(-15)	0.0665	(0.83)	D(10)	0.2788***	(3.47)
D(-14)	0.0970	(1.21)	D(11)	0.2130***	(2.65)
D(-13)	0.1405*	(1.75)	D(12)	0.1284	(1.60)
D(-12)	0.1464*	(1.82)	D(13)	0.1735**	(2.16)
D(-11)	0.0869	(1.08)	D(14)	0.1468*	(1.83)
D(-10)	0.2104***	(2.62)	D(15)	0.1260	(1.57)
D(-9)	0.2458***	(3.06)	D(16)	0.0736	(0.92)
D(-8)	0.2320***	(2.89)	D(17)	0.0648	(0.81)
D(-7)	0.1564*	(1.95)	D(18)	0.0961	(1.20)
D(-6)	0.2734***	(3.41)	D(19)	0.0919	(1.14)
D(-5)	0.1183	(1.47)	D(20)	-0.0133	(-0.17)
D(-4)	0.2768***	(3.45)	D(21)	0.0202	(0.25)
D(-3)	0.0314	(0.39)	D(22)	0.0506	(0.63)
D(-2)	0.1625**	(2.02)	D(23)	0.0965	(1.20)
D(-1)	0.1872**	(2.33)	D(24)	0.0197	(0.25)
			D(25)	0.0310	(0.39)
Adjusted R2		0.978			
Observations		2,652			

This table presents regression results on the difference between the download activities of law firms that are engaged to be the legal counsel for an IPO company over time. The sample comes from the 51 law firms that we are able to i) positively identify at least one download on the EDGAR log database and ii) IPO legal counsel and publicly traded company relationship data is available through Lexis Securities Mosaic. Each law firm has 52 weekly observations for a total N of 2,652. The variables of interest are indicators denoted *D(x)*, where *x* represents how many weeks from the initial S-1 filing. Week -26 is the omitted reference category. The regression includes law firm fixed effects. *, **, and *** represent significance at 10%, 5%, and 1%, respectively (two-tailed). t-values are presented beside the coefficient estimates in parentheses.

Table 2 – Descriptive Statistics

Panel A: Crowdsourcing Summary Statistics

Variable	N	Mean	Std Dev	p25	p50	p75
S-1 Prep Crowdsourcing (unlogged)	408	726	652	161	473	1,347
S-1 Review Crowdsourcing (unlogged)	408	1,255	1,922	110	698	1,650
Rank IPT20	408	297.68	170.01	150.50	303.00	441.50
D(Litigation)	408	0.10	0.30	0.00	0.00	0.00
Phrase Count (unlogged)	408	15.14	9.61	4.00	21.00	24.00
Conv Count (unlogged)	408	7.29	4.36	4.00	7.00	9.00
Time (unlogged)	408	83.34	116.40	32.00	57.00	90.00
Legal Fees (unlogged)	408	1.85	1.25	1.05	1.50	2.40
Size (unlogged)	408	2,979	16,724	45	180	1,022
Leverage	408	0.78	0.65	0.38	0.68	0.93
ROA	408	-0.16	0.51	-0.10	-0.00	0.01
Sales	408	0.34	0.49	0.06	0.21	0.44
D(Venture Capital)	408	0.40	0.49	0.00	0.00	1.00
Overhang %	408	33.20	25.88	6.39	31.70	53.82
Secondary %	408	6.33	12.32	0.00	0.00	7.50
UW Rank	408	8.48	1.29	8.00	9.00	9.00
Age	408	7.83	6.79	4.00	6.00	10.00

This table provides summary statistics for the 408 law firm-IPO observations included in our sample. The sample contains all observations where i) IPO legal counsel has at least one download on the EDGAR log file database, ii) IPO legal counsel and publicly traded company relationship data is available through Lexis Securities Mosaic, iii) company financial and trading data is available from Compustat and CRSP and, iv) IPO data available through SDC for the period 2003 through 2017. *Phrase Count, Conv Count, Time,* and *Size* are presented in their unlogged form in this table. All variables are defined in Appendix A.

Panel B: S-1 Crowdsourcing Behavior by Quintile of Legal Fees

	N	Mean	Std Dev	p25	p50	p75
Quintile 1 (Smallest Legal Fees)						
S-1 Prep Crowdsourcing	98	489	625	73	226	555
S-1 Review Crowdsourcing	98	844	1,263	78	370	892
Quintile 2						
S-1 Prep Crowdsourcing	66	902	718	267	725	1,450
S-1 Review Crowdsourcing	66	1,258	2,084	124	886	1,572
Quintile 3						
S-1 Prep Crowdsourcing	81	927	679	326	863	1,452
S-1 Review Crowdsourcing	81	1,126	1,227	302	790	1,552
Quintile 4						
S-1 Prep Crowdsourcing	82	655	590	179	436	1,353
S-1 Review Crowdsourcing	82	1,062	1,583	110	576	1,258
Quintile 5 (Largest Legal Fees)						
S-1 Prep Crowdsourcing	81	740	560	240	715	1,203
S-1 Review Crowdsourcing	81	2,073	2,889	172	859	2,660

This table presents the S-1 Disclosure Crowdsourcing behavior in both the S-1 Prep period and the S-1 Review period by quintile of Legal Fees. Quintile 1 has the smallest legal fees where legal fees associated with the IPO are less than $1 million. Quintile 5 has the largest legal fees with the average legal fees associated with the IPO are $3.8 million.

Table 3 – Pearson (below) and Spearman (above) Correlation Table

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
(1) S-1 Prep Crowdsourcing		**-0.23**	-0.05	**-0.13**	0.04	0.02	**0.13**	-0.07	-0.06	**-0.20**	**-0.12**	**0.21**	-0.09	**-0.10**	0.03	0.01	**0.64**
(2) Phrase Count	**-0.30**		**0.46**	**0.53**	-0.04	-0.02	0.00	**0.24**	0.04	**0.30**	**0.28**	**-0.21**	-0.06	**0.32**	**0.11**	-0.02	**0.19**
(3) Conv Count	-0.08	**0.43**		**0.65**	-0.04	0.07	**0.10**	**0.12**	0.00	0.04	0.06	-0.08	-0.00	**0.18**	0.07	-0.08	**0.25**
(4) Time	-0.06	**0.28**	**0.48**		-0.08	0.01	0.09	**0.16**	0.04	**0.10**	**0.11**	**-0.14**	**-0.13**	**0.14**	0.07	0.01	**0.34**
(5) Rank IPT20	0.02	-0.04	-0.08	-0.06		0.04	0.09	0.07	-0.02	0.01	-0.01	-0.04	-0.05	0.02	-0.00	-0.02	0.02
(6) D(Litigation)	0.02	-0.05	0.04	-0.04	0.04		**0.10**	-0.04	0.02	-0.07	0.03	**0.15**	0.04	0.03	0.06	0.06	0.03
(7) Legal Fees	0.05	**0.11**	**0.10**	**0.10**	0.05	0.07		**0.48**	0.07	0.05	-0.02	**-0.15**	**-0.25**	**0.15**	**0.16**	**-0.11**	**0.15**
(8) Size	-0.00	**0.13**	0.06	**0.24**	-0.02	-0.04	**0.16**		**0.11**	**0.40**	-0.04	**-0.56**	**-0.35**	**0.33**	**0.25**	**-0.20**	0.03
(9) Leverage	0.00	-0.02	-0.02	0.00	-0.01	0.05	-0.02	0.03		**-0.18**	**0.11**	**-0.14**	-0.05	0.03	**-0.10**	0.09	0.05
(10) ROA	**-0.12**	**0.21**	0.07	0.08	0.04	-0.03	**0.17**	0.06	**-0.51**		**0.38**	**-0.38**	-0.00	**0.41**	**0.14**	0.04	**-0.13**
(11) Sales	**-0.11**	**0.20**	0.01	-0.03	**-0.11**	0.00	-0.07	-0.09	-0.00	**0.12**		**-0.14**	**0.15**	**0.21**	0.03	**0.22**	-0.00
(12) D(VC)	**0.22**	**-0.26**	-0.09	**-0.14**	-0.04	**0.15**	**-0.21**	**-0.13**	-0.03	**-0.23**	**-0.12**		**0.26**	**-0.23**	0.06	**0.26**	0.07
(13) Overhang %	-0.07	-0.05	-0.00	**-0.18**	-0.06	0.02	**-0.27**	**-0.17**	-0.02	0.01	**0.16**	**0.21**		**-0.14**	-0.01	**0.15**	**-0.15**
(14) Secondary %	**-0.11**	**0.21**	0.09	0.03	0.01	-0.01	**0.15**	**0.13**	-0.04	**0.18**	0.08	**-0.26**	**-0.11**		**0.17**	-0.02	0.04
(15) UW Rank	**0.13**	0.06	0.02	0.07	-0.02	0.04	**0.21**	0.06	**-0.23**	**0.32**	-0.03	0.08	-0.04	0.11		0.00	0.05
(16) Age	-0.04	-0.03	-0.06	0.06	0.05	0.03	-0.06	0.02	0.08	-0.02	0.08	0.08	0.08	0.01	0.01		-0.00
(17) S-1 Rev Crowdsourcing	**0.36**	**0.16**	**0.20**	**0.50**	-0.06	0.03	**0.18**	0.09	0.04	0.01	-0.06	-0.01	**-0.15**	0.00	**0.11**	0.01	

This table presents univariate correlations between our variables of disclosure crowdsourcing, our dependent variables of interest and controls. We present Pearson correlations below the diagonal and Spearman above the diagonal. Bolded numbers indicate the correlation is significant at the 5% level. All variables are defined in Appendix A

Table 4 –Disclosure Crowdsourcing and the SEC Review Process

	Dependent Variables:					
	Phrase Count		*Conv Count*		*Time*	
	(1)		(2)		(3)	
S-1 Prep Crowdsourcing	0.02	(0.56)	-0.07**	(-1.98)	-0.12	(-1.53)
Legal Fees	0.07*	(1.81)	0.04	(1.37)	0.09	(1.24)
D(Multi Law Firms)	-0.06	(-0.75)	-0.01	(-0.21)	-0.26*	(-1.82)
Size	0.13***	(4.18)	0.00	(0.13)	0.02	(0.37)
Leverage	0.02	(0.28)	-0.06	(-1.20)	-0.17	(-1.40)
ROA	-0.28***	(-2.71)	-0.15*	(-1.94)	-0.32*	(-1.79)
Sales	0.23***	(2.70)	0.00	(0.01)	0.26*	(1.79)
D(VC)	0.13	(1.10)	0.06	(0.68)	0.02	(0.11)
Overhang %	-0.00	(-1.53)	-0.00	(-0.29)	-0.01***	(-2.78)
Sec %	0.00	(0.54)	0.00	(0.65)	-0.00	(-0.35)
UW Rank	-0.06*	(-1.68)	-0.01	(-0.19)	-0.05	(-0.75)
Age	-0.00	(-0.31)	0.00	(0.59)	0.02**	(2.29)
Law Firm FE	Yes		Yes		Yes	
Year FE	Yes		Yes		Yes	
Industry FE	Yes		Yes		Yes	
Adjusted R^2	0.466		0.143		0.137	
N. observations	408		408		408	

This table presents the results of an OLS regression where the dependent variable is *Phrase Count*, *Conv Count*, or *Time*. Appendix A contains detailed variable definitions. *, **, and *** represent significance at 10%, 5%, and 1%, respectively (two-tailed). T-statistics are presented to the right of the coefficient estimates in parentheses.

Table 5 –Disclosure Crowdsourcing and the Speed of post-IPO Price Formation

	Dependent Variable: *Rank IPT20*			
	(1)		(2)	
S-1 Prep Crowdsourcing	28.98***	(2.78)	29.45***	(2.82)
Legal Fees			-3.75	(-0.40)
D(Multi Law Firms)			57.30***	(3.00)
Size			2.73	(0.38)
Leverage			0.81	(0.05)
ROA			33.04	(1.35)
Sales			-33.90*	(-1.71)
D(VC)			-34.89	(-1.30)
Overhang %			-0.30	(-0.79)
Sec %			0.29	(0.38)
UW Rank			3.00	(0.34)
Age			1.16	(0.86)
Law Firm FE	Yes		Yes	
Year FE	Yes		Yes	
Industry FE	Yes		Yes	
Adjusted R^2	0.098		0.128	
N. observations	408		408	

This table presents the results of an OLS regression where the dependent variable is the rank of 20 day intra-period timeliness measure and the main independent variable of interest is S-1 Prep Crowdsourcing. Appendix A contains detailed variable definitions. *, **, and *** represent significance at 10%, 5%, and 1%, respectively (two-tailed). T-statistics are presented to the right of the coefficient estimates in parentheses.

Table 6 –Disclosure Crowdsourcing and the Post-IPO Shareholder Litigation

	Dependent Variable: *D(Litigation)*			
	(1)		(2)	
S-1 Prep Crowdsourcing	-0.01	(-1.10)	-0.02*	(-1.89)
Legal Fees			0.02	(0.95)
D(Multi Law Firms)			-0.11***	(-2.82)
Size			-0.00	(-0.08)
Leverage			0.02	(0.49)
ROA			-0.01	(-0.18)
Sales			0.02	(0.36)
D(VC)			0.15***	(3.05)
Overhang %			-0.00	(-0.77)
Sec %			0.00	(0.36)
UW Rank			0.01	(0.92)
Age			0.00	(0.17)
Year FE	Yes		Yes	
Industry FE	Yes		Yes	
Adjusted R^2	-0.005		0.025	
N. observations	354		354	

This table presents the results of a linear probability model where the dependent variable is an indicator variable equal to one if the IPO firm has a lawsuit within 3 years of the IPO and zero otherwise. The sample suffers some attrition because there are no lawsuits for IPO years 2003, 2004, 2006, 2008, 216, or 2017 and thus the IPO year fixed effect perfectly predicts the dependent variable in those cases. Appendix A contains detailed variable definitions. *, **, and *** represent significance at 10%, 5%, and 1%, respectively (two-tailed). T-statistics are presented to the right of the coefficient estimates in parentheses.

Table 7 – Analysis with S-1 Review Crowdsourcing

	Dependent Variables:									
	Phrase Count		*Conv Count*		*Time*		*Rank IPT20*		*D(Litigation)*	
	(1)		(2)		(3)		(4)		(5)	
S-1 Prep CS	-0.07	(-1.60)	-0.12***	(-3.90)	-0.33***	(-5.14)	28.34***	(2.62)	-0.03*	(-1.84)
S-1 Review CS	0.14***	(8.67)	0.09***	(7.25)	0.32***	(13.06)	1.66	(0.40)	0.00	(0.47)
Legal Fees	0.06*	(1.65)	0.03	(1.18)	0.06	(1.00)	-3.90	(-0.41)	0.02	(0.93)
D(Multi Law Firms)	-0.02	(-0.26)	0.01	(0.25)	-0.16	(-1.37)	57.81***	(3.02)	-0.11***	(-2.81)
Size	0.13***	(4.81)	0.01	(0.29)	0.03	(0.74)	2.79	(0.38)	-0.00	(-0.09)
Leverage	0.01	(0.21)	-0.07	(-1.38)	-0.19*	(-1.89)	0.73	(0.04)	0.02	(0.47)
ROA	-0.21**	(-2.22)	-0.10	(-1.43)	-0.15	(-1.04)	33.92	(1.38)	-0.01	(-0.15)
Sales	0.22***	(2.91)	-0.00	(-0.05)	0.25**	(2.10)	-33.97*	(-1.71)	0.02	(0.36)
D(VC)	0.16	(1.58)	0.08	(1.03)	0.11	(0.69)	-34.44	(-1.29)	0.15***	(3.06)
Overhang %	-0.00	(-0.93)	0.00	(0.33)	-0.01**	(-2.26)	-0.28	(-0.75)	-0.00	(-0.74)
Sec %	0.00	(0.72)	0.00	(0.80)	-0.00	(-0.24)	0.30	(0.39)	0.00	(0.34)
UW Rank	-0.07**	(-2.18)	-0.01	(-0.47)	-0.07	(-1.39)	2.87	(0.32)	0.01	(0.91)
Age	-0.00	(-0.90)	0.00	(0.16)	0.02*	(1.96)	1.12	(0.83)	0.00	(0.16)
Law Firm FE	Yes		Yes		Yes		Yes		No	
Year FE	Yes		Yes		Yes		Yes		Yes	
Industry FE	Yes		Yes		Yes		Yes		Yes	
Adjusted R^2	0.564		0.259		0.430		0.125		0.022	
N. observations	408		408		408		408		354	

This table presents the results of an OLS regression where the dependent variable is *Phrase Count, Conv Count*, *Time*, *Rank IPT20*, or *D(Litigation)*. In these models, we have included the variable *S-1 Review Crowdsourcing* as a control variable. Appendix A contains detailed variable definitions. *, **, and *** represent significance at 10%, 5%, and 1%, respectively (two-tailed). T-statistics are presented to the right of the coefficient estimates in parentheses.

Table 8 – Path analysis of direct and indirect effects of disclosure crowdsourcing on speed of price formation

	Path = Phrase Count			Path = Time			Path = S-1 Review CS			Path = Conv Count		
	Coef.	t-Stat		Coef.	t-Stat		Coef.	t-Stat		Coef.	t-Stat	
Direct Path												
p[S-1 Prep CS, Rank IPT20]	22.44	2.00	**	22.44	2.00	**	22.44	2.00	**	22.44	2.00	**
Mediated Path												
I. p[S-1 Prep CS, PATH]	0.02	0.56		-0.12	-1.53		0.67	4.77	***	-0.07	-1.98	**
II. P[PATH, Rank IPT20]	-0.59	-0.04		-27.91	-2.22	**	8.27	1.58		27.68	1.07	
Indirect Effect (I × II)	-0.01	0.00		3.29	0.01		5.54	0.13		-1.80	0.00	
Control Variables Direct Path												
p[Legal Fees, Rank IPT20]	-3.19	-0.34										
p[D(Multi Law Firms, Rank IPT20]	53.01	2.76	***									
p[Size, Rank IPT20]	3.60	0.48										
p[Leverage, Rank IPT20]	-2.61	-0.16										
p[ROA, Rank IPT20]	32.39	1.31										
p[Sales, Rank IPT20]	-26.79	-1.32										
p[D(VC), Rank IPT20]	-33.51	-1.25										
p[Overhang %, Rank IPT20]	-0.44	-1.14										
p[Sec %, Rank IPT20]	0.22	0.29										
p[UW Rank, Rank IPT20]	1.10	0.12										
p[Age, Rank IPT20]	1.55	1.13										
N. observations	408											
Fixed Effects	Law Firm, Industry, and Year											
Adjusted Goodness of Fit	0.84											

This table presents selected results of a structural equation model of the direct effect of disclosure crowdsourcing in the preparation period on tthe speed of price formation as well as the indirect effects of disclosure crowdsourcing in the preparation period through the number of issues identified by the SEC in their response to the initial S-1 (*Phrase Count*), the length of time of the SEC review process (*Time*), the number of S-1 amendments (*Conv Count*) and the S-1 download activity in the review period (*S-1 Review Crowdsourcing*). We present the unstandardized path coefficients above with standard errors derived from the full matrix of the structural equation. The significance of the indirect effect (i.e., the product coefficient) is assessed using the Sobel (1982) test. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 9 – Path analysis of direct and indirect effects of disclosure crowdsourcing on post-IPO litigation

	Path = Phrase Count			Path = Time			Path = S-1 Review CS			Path = Conv Count		
	Coef.	t-Stat		Coef.	t-Stat		Coef.	t-Stat		Coef.	t-Stat	
Direct Path												
p[S-1 Prep CS, D(Litigation)]	-0.03	-1.82	*	-0.03	-1.82	*	-0.03	-1.82	*	-0.03	-1.82	**
Mediated Path												
I. p[S-1 Prep CS, PATH]	0.00	-0.10		-0.02	-0.44		0.97	11.89	***	-0.01	-0.54	
II. P[PATH, D(Litigation)]	-0.04	-1.26		-0.01	-0.41		0.01	0.65		0.06	1.20	
Indirect Effect (I × II)	0.00	0.10		0.00	0.23		0.01	0.09		0.00	-0.21	
Control Variables Direct Path												
p[Legal Fees, D(Litigation)]	0.017	0.91										
p[D(Multi Law Firms, D(Litigation)]	-0.117	-2.94	***									
p[Size, D(Litigation)]	0.003	0.23										
p[Leverage, D(Litigation)]	0.018	0.54										
p[ROA, D(Litigation)]	-0.008	-0.17										
p[Sales, D(Litigation)]	0.029	0.63										
p[D(VC D(Litigation)]	0.152	3.10	***									
p[Overhang %, D(Litigation)]	-0.001	-0.91										
p[Sec %, D(Litigation)]	0.000	0.22										
p[UW Rank, D(Litigation)]	0.010	0.64										
p[Age, D(Litigation)]	0.000	0.15										
N. observations	354											
Fixed Effects	Industry and Year											
Adjusted Goodness of Fit	0.53											

This table presents selected results of a structural equation model of the direct effect of disclosure crowdsourcing in the preparation period on tthe speed of price formation as well as the indirect effects of disclosure crowdsourcing in the preparation period through the number of issues identified by the SEC in their response to the initial S-1 (*Phrase Count*), the length of time of the SEC review process (*Time*), the number of S-1 amendments (*Conv Count*) and the S-1 download activity in the review period (*S-1 Review Crowdsourcing*). We present the unstandardized path coefficients above with standard errors derived from the full matrix of the structural equation. The significance of the indirect effect (i.e., the product coefficient) is assessed using the Sobel (1982) test. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.