Journal of Accounting and Economics 62 (2016) 228–233

Contents lists available at ScienceDirect

Journal of Accounting and Economics

journal homepage: www.elsevier.com/locate/jae

Discussion of using unstructured and qualitative disclosures to explain accruals

Gerard Hoberg

The University of Southern California Marshall School of Business, United States

ARTICLE INFO

Available online 15 September 2016

Keywords:
Accruals
Disclosure
Textual analysis
Support vector regression

1. Introduction

There is a long tradition of research explaining accruals. One of the most well-known articles, Jones (1991), has been cited 843 times according to the ISI Web of Knowledge, and 5874 times according to Google Scholar at the time of this writing. Much of the focus has been on understanding how accruals arise, and in some cases, linking accruals to manipulative behavior such as earnings management. Despite extensive progress, the ability to explain accruals is far from complete (see Ball and Shivakumar, 2006 for example). It is thus easy to motivate the question addressed by Frankel et al. (2016) (henceforth FJL) as being important.

FJL propose that text in verbal disclosures such as management's discussion and analysis (henceforth MD&A) and conference calls can explain accruals. Because many studies find these disclosures to be informative, the potential of this agenda is substantial. Although FJL do not provide an improved model of discretionary accruals, an implication of FJL is the possibility of doing so in the future. Progress in this area could have important ramifications not only for accounting scholars interested in signal extraction, but also for regulators and theories predicting how accruals arise.

FJL also contribute to the growing literature using computational linguistics in accounting. A common objective of studies in this area is to use text, a resource that many still view as untapped, to explain or predict numerical variables such as accruals or stock returns.[1] Studies using computational linguistics to inform business outcomes can be further categorized into three groups. Some articles use these tools (1) to create a measure of an important variable that cannot be measured using standard constructs. Others (2) attempt to extract a maximally informative signal from a corpus regarding a specific outcome variable, and others (3) focus on interpretable content and test a specific theory.

FJL is in the second category, and the authors highlight that MD&A and conference calls can strongly explain accruals. Further binning the paper in this category, FJL state on page one that " (Friedman, 1953) and (Bloomfield et al., 2016) argue for testing predictions based on theoretical constructs, which we do not do." A concern is that such an atheoretical approach can not only fail to inform the validity of theories, but it can also fail to produce meaningful conclusions about accruals themselves. This is the well-known "black box critique". For example, consider the following rhetorical question: *Is an accrual really "explained" if it is absorbed by a variable we cannot interpret*?

[1] Examples of other studies that address related issues are Kothari et al. (2009) and Brown and Jennifer (2011).

http://dx.doi.org/10.1016/j.jacceco.2016.07.002

An example in the third category is Ball et al. (2016) (BHM), who also examine the ability of MD&A to explain accruals. BHM do not attempt to maximize signal strength, but instead consider interpreting content to test hypotheses relating to a specific theory: business change as a driver of accruals (see Owens et al., 2013). To do so, BHM use Latent Dirichlet Allocation (LDA), which was discovered by Blei et al. (2003). Unlike SVR, LDA does not allow a specific outcome variable to be given as input. Some aspects of my discussion will draw upon comparisons between FJL and BHM.

After summarizing relevant institutional details regarding how MD&A is created, my discussion will focus on the following critical items: (1) reverse causality stemming from the experiment's design, (2) overfitting concerns relating to SVR, and (3) the tendency of SVR to anchor on common words and those with multiple interpretations, which hinders the interpretation of its results.

I also note FJL's positive contributions: (1) FJL establish that there is a strong signal in MD&A and conference call text that can predict accruals, and (2) the paper shows that MD&A disclosures are roughly as informative as are conference calls. This latter result is perhaps one of the more important findings of FJL. It suggests that the regulatory framework established by the Securities and Exchange Commission, despite concerns by many that it primarily encourages boilerplate content, does in fact create informative disclosure. This conclusion in FJL echoes similar findings in Hanley and Hoberg (2010) who study IPO prospectuses.

2. Theoretical and institutional framework

Before interpreting the output of an algorithm like SVR, it is important to consider how MD&A is generated. Doing so can either (A) validate why SVR might indeed be appropriate or (B) indicate why SVR might be problematic, as it may indicate a potential mechanistic effect or other potential biases.

Practical and regulatory requirements indicate that MD&A is generated according to the flow of time depicted in Fig. 1. As is the case for financial statements, MD&A text is meant to summarize business activities and results in a specific fiscal year. Hence, the fiscal year must first be realized and its state observed (left most side of the figure), before either the financial statements or the MD&A can be created (right side of the figure). Fig. 1 defines the fiscal year's realization as a high dimensional state variable of the firm, which would include quantitative realizations such as revenues, cash balances, earnings and accruals. However, the state variable also includes qualitative realizations such as growth options, product launches, and future investment plans from the given fiscal year.

After the fiscal year state is realized, the second event in the flow of time in Fig. 1 is the creation of the financial statements. These statements summarize the quantitative aspects of the state variable mentioned above, but they cannot summarize the qualitative aspects. Hence Fig. 1 shows that only a fraction of the information in the complete state variable can flow to the financial statements.

Critically, the third event in time is the writing of the firm's MD&A. This is the primary item studied in FJL. It is important to note that it is difficult to write MD&A until both (1) the fiscal year is complete and (2) the financial statements are complete. This is because the goal of the MD&A is to explain and interpret the financial statements for shareholders, and also to provide a summary of the non-quantitative aspects of the firm's performance.

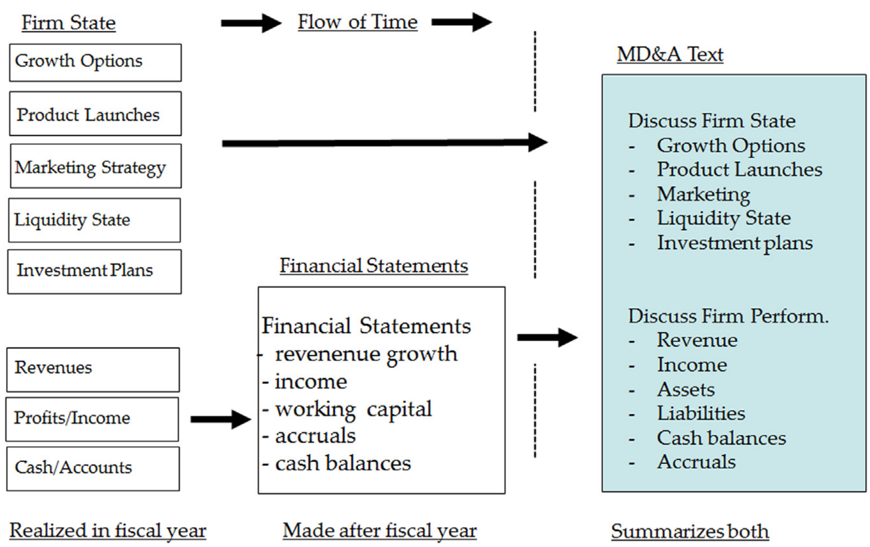

Fig. 1. *Notes*: The Figure displays the timeline of events, and the informational state space, resulting in the creation of a firm's MD&A in a given year.