Chapter 8
A Review of Textual Analysis in Economics and Finance

Carissa L. Tudor
Board of Governors of the Federal Reserve System, USA

Clara Vega
Board of Governors of the Federal Reserve System, USA

ABSTRACT

This chapter provides an overview of studies in finance and economics that use automated textual analysis algorithms to analyze the informational content of a wide variety of texts, including journalist’s coverage of news events, management-issued statements, and Internet stock message boards. In these studies, researchers quantify qualitative information with one or more of the following textual tone variables: textual negativity, positivity, and uncertainty. The studies show that textual negativity and positivity conveyed by managers and journalists helps predict future firm level and aggregate economic activity. Textual negativity and positivity, in turn, affect asset prices, although the information is sometimes incorporated with some delay. Textual uncertainty of management-issued information is associated with future cash flow volatility and asset price volatility. In contrast, the textual tone of stock market message board postings is, on average, not very informative in explaining asset prices. The use of automated textual analysis algorithms in finance and economics is a relatively new phenomenon and research in this area is expected to continue to grow.

INTRODUCTION

Research in economics and finance has greatly benefited from the development of automated textual analysis algorithms and programs that implement them, such as DICTION. In particular, this technology has increased the number of information events and the types of information events that are available to researchers, who, in turn, use them to forecast future economic activity and estimate the impact public information has on asset prices. In this review, we discuss selected papers that analyze the informational content of a wide variety of texts. Li (2010) provides a
comprehensive review of the corporate disclosure literature that uses automated textual analysis algorithms. Demers and Yu (2013) provide a discussion of uncertainty in the accounting and finance literature. We intend for the review to be readily accessible to a broad audience from varied research fields. The studies we review primarily contribute to two research areas in finance and economics: (a) asset price response to public information and (b) forecasts of firm’s earnings and aggregate economic activity.

Prior to the development of automated textual analysis, researchers used a small number of pre-scheduled public announcements that are easily quantified (e.g. corporate earnings, industrial production, GDP) to forecast future economic activity and estimate the impact public information has on asset prices. Thanks to this development, researchers are now able to analyze a larger number of information events and different types of information events, such as unscheduled public announcements (e.g. the FDA approval of a new drug, natural disaster news), text that complements pre-scheduled announcements (e.g. manager’s explanation of corporate earnings, journalists’ and analysts’ coverage of corporate earnings announcements), and news emanating from different sources (e.g. journalists, managers, stock market message board postings, government agencies). Importantly, research shows that by taking into account some of these new sources of information, economists are able to improve their forecasts of economic activity and that some of these new sources of information have an economically and statistically significant impact on asset prices.

In this review we discuss eight papers, which we’ve grouped based on the type of text they analyze. In Table 1, we provide a brief summary of these studies. Antweiler and Frank (2004), Das, Matinez-Jerez, and Tufano (2005), and Das and Chen (2007) analyze the textual tone of internet stock message boards. Engelberg (2008) and Tetlock, Saar-Tsechansky, and Macskassy (2008) analyze the textual tone of firm-specific Dow Jones

Table 1. Summary of the eight papers reviewed

<table>
<thead>
<tr>
<th>Authors</th>
<th>Text Analyzed</th>
<th>Linguistic Analysis Method</th>
<th>Asset Pricing Finding</th>
<th>Fundamental’s Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antweiler and Frank (2004)</td>
<td>Internet stock message boards</td>
<td>Naive Bayes algorithm</td>
<td>No significant effect</td>
<td>NA</td>
</tr>
<tr>
<td>Das, Martinez-Jerez, and Tufano (2005)</td>
<td>Internet stock message boards</td>
<td>Combination of methods</td>
<td>No significant effect</td>
<td>NA</td>
</tr>
<tr>
<td>Das and Chen (2007)</td>
<td>Internet stock message boards</td>
<td>Combination of methods</td>
<td>No significant effect</td>
<td>NA</td>
</tr>
<tr>
<td>Engelberg (2008)</td>
<td>Firm-specific media coverage surrounding earnings announcements</td>
<td>General Inquirer and type dependency parsing</td>
<td>Negative words affect announcement period returns and 80-day post-announcement period returns.</td>
<td>NA</td>
</tr>
<tr>
<td>Tetlock et al. (2008)</td>
<td>Firm-specific media coverage</td>
<td>DICTION</td>
<td>Negative words affect announcement period returns</td>
<td>Negative words are related to firm’s earnings. Change in net optimism (negative minus positive words) is related to firm’s earnings.</td>
</tr>
<tr>
<td>Davis and Piger (2012)</td>
<td>Management-issued text accompanying earnings announcements</td>
<td>DICTION, General Inquirer, Loughran and McDonald (2011)</td>
<td>Change in net optimism (negative minus positive words) affect announcement period returns. The effect of net optimism on asset prices depends on announcement and firm characteristics. Textual uncertainty affects announcement period and post-announcement period volatility.</td>
<td>Change in net optimism (negative minus positive words) and textual uncertainty are related to firm’s fundamentals, earnings and volatility of those earnings.</td>
</tr>
<tr>
<td>Demers and Vega (2012)</td>
<td>Management-issued text accompanying earnings announcements</td>
<td>Ravenpack</td>
<td>Change in net optimism (negative minus positive words) predicts future GDP</td>
<td>Change in net optimism (negative minus positive words) and textual uncertainty are related to firm’s fundamentals, earnings and volatility of those earnings.</td>
</tr>
</tbody>
</table>
 Related Content

A Linguistic Exploration of Indigenous Languages Adverts: A Critical Discourse Approach
[www.irma-international.org/article/a-linguistic-exploration-of-indigenous-languages-adverts/232230/](www.irma-international.org/article/a-linguistic-exploration-of-indigenous-languages-adverts/232230/)

Quantifying the Lemma Massa as a Proper-Name in the John and Massa Tales: Using Quantitative and Qualitative Pragmatic Methods to Analyze Common Nouns Used as Proper Names

K-20 Education and Globalization
[www.irma-international.org/chapter/k-20-education-and-globalization/97459/](www.irma-international.org/chapter/k-20-education-and-globalization/97459/)

Intercultural Competence for Teachers of Young ELLs
Ellen Yeh and Guofang Wan (2018). *Handbook of Research on Pedagogies and Cultural Considerations for Young English Language Learners* (pp. 192-216).
[www.irma-international.org/chapter/intercultural-competence-for-teachers-of-young-ells/190980/](www.irma-international.org/chapter/intercultural-competence-for-teachers-of-young-ells/190980/)

Factors Related to EFL/ESL Readers' Reading Strategy Use: A Literature Review
[www.irma-international.org/article/factors-related-to-eflesl-readers-reading-strategy-use/222825/](www.irma-international.org/article/factors-related-to-eflesl-readers-reading-strategy-use/222825/)