Enhancing Substantive Analytical Procedures with Third-Party Generated Information from Social Media

"Investors, and others, are accessing and analyzing massive amounts of information from sources, like social media, unimaginable just a few years ago. This new data may be empowering investors to make smarter investment decisions"

Kara Stein – SEC Commissioner 2015

Objectives

- **Do Twitter proxies of consumer intention to buy and consumer satisfaction enhance the effectiveness of substantive analytical procedures for the revenue account?**

- Investigate whether information generated by third-parties on social media can

  1) Improve the accuracy of substantive analytical procedures
  2) Improve the error detection ability of substantive analytical procedures

Motivation

- Research suggests that social media information contains incremental information about firms’ stock prices, and sales performance (e.g. Bollen, Mao, Zheng 2011; Tang 2017)

- Inspection findings indicate that accounting firms failed to develop precise expectations (PCAOB 2007; PCAOB 2016a)

- Social media consumer postings about firms’ products and brands offer an independent benchmark, difficult to tamper with, easily accessible, and timely

- More research is needed to examine the usefulness of new nonfinancial sources of potential audit evidence (Yoon 2016)

Research Design (1)

Sample – 24 industries

- Likefolio, https://home.likefolio.com/, and Compustat

Sample Selection - Firm-Quarter Observations 2012-2017

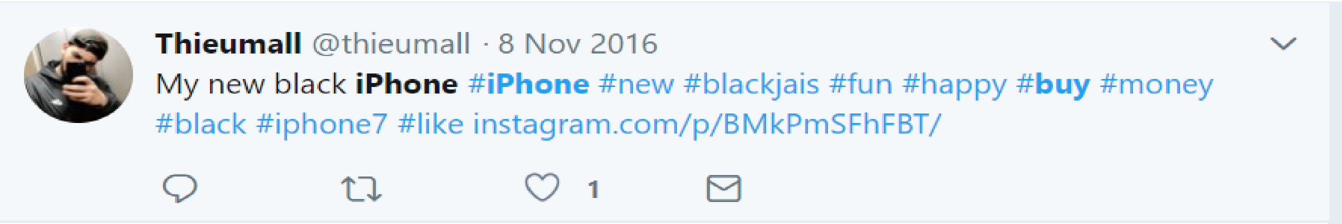
	Firms	Firm-Quarter Observations
Firms that are publicly listed and have third-party generated Twitter information	194	4,656
Less: Firms with missing financial information or zero values	(9)	(216)
Less: Firms with missing information from Twitter for either Purchase Interest or Se	(15)	(360)
Less: Firms without four quarters of data	(73)	(1,752)
Less: Firms in the Financial Services Industry	(9)	(216)
Total	88	2,112

- Quarterly financial information is interpolated into monthly observations and matched with Twitter data

Research Design (2)

Twitter Measures

- Likefolio, https://home.likefolio.com/ ,provided consumer intent and satisfaction for products and brands of 194 B2C
  - Mapping of brands and products to the company
  - Consumer Intention to Buy



TCI: total count of tweets related to the firm's product or brand past/future intent to buy

  - Consumer Sentiment

TCS: ratio of positive tweets to total (positive and negative) tweets

Research Design (3)

Models

(1)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \epsilon$
(5)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 TCI_{it} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 TCI_{it} + \epsilon$
(6)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 TCS_{it} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 TCS_{it} + \epsilon$
(2)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 GDP_{t-12} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 GDP_{t-1} + \epsilon$
(7)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 TCI_{it} + \beta_3 GDP_{t-12} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 TCI_{it} + \beta_3 GDP_{t-1} + \epsilon$
(8)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 TCS_{it} + \beta_3 GDP_{t-12} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 TCS_{it} + \beta_3 GDP_{t-1} + \epsilon$
(3)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 AR_{it} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 AR_{it} + \epsilon$
(9)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 AR_{it} + \beta_3 TCI_{it} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 AR_{it} + \beta_3 TCI_{it} + \epsilon$
(10)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 AR_{it} + \beta_3 TCS_{it} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 AR_{it} + \beta_3 TCS_{it} + \epsilon$
(4)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 AR_{it} + \beta_3 GDP_{t-12} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 AR_{it} + \beta_3 GDP_{t-1} + \epsilon$
(11)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 AR_{it} + \beta_3 TCI_{it} + \beta_3 GDP_{t-12} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 AR_{it} + \beta_3 TCI_{it} + \beta_3 GDP_{t-1} + \epsilon$
(12)	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-12} + \beta_2 AR_{it} + \beta_3 TCS_{it} + \beta_3 GDP_{t-12} + \epsilon$	$Sales_{it} = \beta_0 + \beta_1 Sales_{it-1} + \beta_2 AR_{it} + \beta_3 TCS_{it} + \beta_3 GDP_{t-1} + \epsilon$

Results (1)

Prediction Performance

	TCI				TCS			
	(1) vs. (5)	(2) vs. (7)	(3) vs. (9)	(4) vs. (11)	(1) vs. (6)	(2) vs. (8)	(3) vs. (10)	(4) vs. (12)
Traditional - SAP	16 of 24	16 of 24	16 of 24	14 of 24	15 of 24	14 of 24	12 of 24	15 of 24
Proposed - SAP	19 of 24	21 of 24	18 of 24	22 of 24	14 of 24	20 of 24	14 of 24	22 of 24

Results (2)

Error Detection Performance – alpha .33

	False Positive							
	TCI				TCS			
	(1) vs. (5)	(2) vs. (7)	(3) vs. (9)	(4) vs. (11)	(1) vs. (6)	(2) vs. (8)	(3) vs. (10)	(4) vs. (12)
Traditional - SAP	13 of 24	11 of 24	11 of 24	11 of 24	9 of 24	9 of 24	12 of 24	9 of 24
Proposed - SAP	17 of 24	15 of 24	13 of 24	15 of 24	16 of 24	12 of 24	16 of 24	13 of 24

Results (3)

Error Detection Performance – alpha .33

	False Negative							
	TCI				TCS			
	(1) vs. (5)	(2) vs. (7)	(3) vs. (9)	(4) vs. (11)	(1) vs. (6)	(2) vs. (8)	(3) vs. (10)	(4) vs. (12)
Traditional - SAP	8 of 24	4 of 24	7 of 24	7 of 24	7 of 24	8 of 24	9 of 24	12 of 24
Proposed - SAP	6 of 24	10 of 24	11 of 24	10 of 24	7 of 24	10 of 24	10 of 24	8 of 24

Conclusion

- Provided insights on the usefulness of external nonfinancial information produced by social media platforms
- Used third-party generated Tweets of firms’ products and brands to examine power of analytical procedures