How Have Stock Markets Responded to 35 Years of Analyst Reports? Evidence from Machine Learning and Textual Analysis*

Mustafa Gultekin† gultekin@unc.edu

Abena Owusu§ owusua@rpi.edu

Thomas Shohfi‡§ shohft@rpi.edu

Majeed Simaan¶ msimaan@stevens.edu

December 29, 2019

Abstract

Using machine learning (ML) and contrasting with simple textual sentiment score and principal components analysis (PCA) methods, we examine the time series of content within over 700,000 sell-side analyst research reports from 1983 to 2017. We find that analyst reports have significantly changed across a variety of dimensions including length and content of four existing and two new dictionaries related to valuation and alternative metrics. We find that the naive net tone of reports only explains contemporaneous, not future, equity returns. On the other hand, we find that ML methods provide substantially different results from naive sentiment and PCA approaches on determining the impact of analyst reports on financial markets. We also examine the ability of reports to predict changes in firm short interest and volatility. Overall, we find that sell-side analyst reports have stronger impact on smaller cap stocks.

Keywords: Analyst Reports, Machine Learning, Sell-Side Analysts, Textual Analysis

JEL Codes: C19, C45, C55, C63, G10, G20, G24, M41

*We thank participants at the 2019 Eastern Finance Association Annual Meeting and R/Finance 2019. We also thank seminar participants at Rensselaer Polytechnic Institute and Stevens Institute of Technology. We thank the Donald Shohfi Financial Research Fund for computing support. We thank Yuvraj Chopra and Avnish Grover for research assistance. All errors are our own.

†Kenan-Flagler Business School, University of North Carolina, 300 Kenan Center Drive, Chapel Hill, NC 27599, USA.

‡Corresponding author. Tele: (518) 276-6582, Fax: (518) 276-8661, Web: shohfi.com

§Lally School of Management, Rensselaer Polytechnic Institute, 110 8th Street, Pittsburgh Building, Troy, NY 12180, USA.

¶School of Business, Stevens Institute of Technology, 1 Castle Point Terrace, Hoboken, NJ 07030, USA.

How Have Stock Markets Responded to 35 Years of Analyst Reports? Evidence from Machine Learning and Textual Analysis

Abstract

Using machine learning (ML) and contrasting with simple textual sentiment score and principal components analysis (PCA) methods, we examine the time series of content within over 700,000 sell-side analyst research reports from 1983 to 2017. We find that analyst reports have significantly changed across a variety of dimensions including length and content of four existing and two new dictionaries related to valuation and alternative metrics. We find that the naive net tone of reports only explains contemporaneous, not future, equity returns. On the other hand, we find that ML methods provide substantially different results from naive sentiment and PCA approaches on determining the impact of analyst reports on financial markets. We also examine the ability of reports to predict changes in firm short interest and volatility. Overall, we find that sell-side analyst reports have stronger impact on smaller cap stocks.

Keywords: Analyst Reports, Machine Learning, Sell-Side Analysts, Textual Analysis

JEL Codes: C19, C45, C55, C63, G10, G20, G24, M41