Cluster Analysis of Risk Factors

Michael Seeligson* Michael Seeligson†

September 16, 2017

Abstract

We study the risk factors included in the annual reports of corporate public filings. Using several different linguistic algorithms, we identify prototypical examples of each type of risk factor. We demonstrate the evolution of these risk factors over time and the introduction of new risk factor concepts. The methodology and results of this work operationalize the written language included in annual reports and provides data that can supplement the existing quantitative analysis in the finance, accounting, and economics.

*The University of Texas at Dallas, Naveen Jindal School of Management, 800 W. Campbell Rd. (GC10), Richardson, TX 75080, USA; E: michael.seeligson@utdallas.edu; P: +1-972-883-4464.

†The University of Texas at Dallas, Naveen Jindal School of Management, 800 W. Campbell Rd. (GC10), Richardson, TX 75080, USA; E: michael.seeligson@utdallas.edu; P: +1-972-883-4464.

Introduction

1 The Model

2 Data and Empirical Methodology

3 Results and Empirical Analysis

4 Conclusion

Appendix A The LDA Model

Appendix B The LSA Model

Appendix C Dictionary of Legal Tuples

References