1 Applications of news analytics in finance: A review

Leela Mitra and Gautam Mitra

1.1 Introduction 1
1.2 News data 4
  1.2.1 Data sources 4
  1.2.2 Pre-analysis of news data 6
1.3 Turning qualitative text into quantified metrics and time-series 10
1.4 Models and applications 17
  1.4.1 Information flow and computational architecture 17
  1.4.2 Trading and fund management 18
  1.4.3 Monitoring risk and risk control 22
  1.4.4 Desirable industry applications 23
1.5 Summary and discussions 24
1.A Appendix: Structure and content of news data 25
  1.A.1 Details of Thomson Reuters News Analytics equity coverage
       and available data 25
  1.A.2 Details of RavenPack News Analytics—Dow Jones Edition:
       Equity coverage and available data 30
1.B References 36
PART I QUANTIFYING NEWS: ALTERNATIVE METRICS

2 News analytics: Framework, techniques, and metrics
Sanjiv R. Das

2.1 Prologue 43
2.2 Framework 44
2.3 Algorithms 46
2.3.1 Crawlers and scrapers 46
2.3.2 Text pre-processing 50
2.3.3 Bayes Classifier 50
2.3.4 Support vector machines 52
2.3.5 Word count classifiers 54
2.3.6 Vector distance classifier 54
2.3.7 Discriminant-based classifier 55
2.3.8 Adjective–adverb classifier 57
2.3.9 Scoring optimism and pessimism 57
2.3.10 Voting among classifiers 58
2.3.11 Ambiguity filters 58
2.3.12 Network analytics 59
2.3.13 Centrality 59
2.3.14 Communities 61
2.4 Metrics 62
2.4.1 Confusion matrix 62
2.4.2 Accuracy 62
2.4.3 False positives 63
2.4.4 Sentiment error 63
2.4.5 Disagreement 63
2.4.6 Correlations 64
2.4.7 Aggregation performance 64
2.4.8 Phase lag metrics 67
2.4.9 Economic significance 67
2.5 Discussion 68
2.6 References 69

3 Managing real-time risks and returns: The Thomson Reuters NewsScope Event Indices
Alexander D. Healy and Andrew W. Lo

3.1 Introduction 73
3.2 Literature review 74
3.3 Data 75
3.3.1 News data 75
3.3.2 Foreign exchange data 76
3.4 A framework for real-time news analytics 77
3.4.1 Assigning scores to news 78
3.4.2 A natural extension to alerts 78
3.4.3 Creating keyword and topic code lists 79
3.4.4 Algorithmic considerations 79
3.5 Validating Event Indices 82
3.5.1 Event analysis 82
3.5.2 Examples of event studies 83
3.5.3 Testing for a change in mean 85
3.5.4 Levene’s Test for equality of variance 88
3.5.5 The $\chi^2$ test for goodness of fit 89
3.6 News indices and FX implied volatility 90
3.6.1 Data pre-processing 89
3.6.2 Implied volatility events 92
3.7 Event study analysis through September 2008 92
3.8 Conclusion 92
3.A Appendix 100
3.A.1 Properties of foreign exchange quote data 100
3.A.2 Properties of Thomson Reuters NewsScope Data 102
3.A.3 Monte Carlo null distributions of the $t$-statistic 102
3.B References 108

4 Measuring the value of media sentiment: A pragmatic view 109
Marion Munz
4.1 Introduction 109
4.2 The value of news for the US stock market 110
4.3 News moves markets 110
4.4 News moves stock prices 111
4.5 News vs. noise 111
4.6 Regulated vs. unregulated news 112
4.6.1 Regulated news 112
4.6.2 Unregulated news 113
4.7 The news component of the stock price 113
4.8 Materiality is near 114
4.9 Size does matter 115
4.10 Corporate senior management under the gun 115
4.11 A case for regulated financial news media 116
4.12 Wall Street analysts may create “material” news 116
4.13 Traders may create news 117
4.14 Earnings news releases 117
4.15 News sentiment used for trading or investing decisions 117
4.16 News sentiment systems 118
4.17 Backtesting news sentiment systems 119
4.18 The value of media sentiment 120
4.19 Media sentiment in action 121
4.20 Conclusion 128

5 How news events impact market sentiment 129
Peter Ager Hafez
5.1 Introduction 129
5.2 Market-level sentiment 131
5.2.1 Data and news analytics 131
5.2.2 Market-level index calculation 132
5.2.3 Strategy and empirical results 133
5.3 Industry-level sentiment 135
  5.3.1 Data and news analytics 137
  5.3.2 Industry-level index calculation 137
  5.3.3 Strategy and empirical results 139
  5.3.4 A directional industry strategy 140
5.4 Conclusion 143
5.A Market-level sentiment data 143
  5.A.1 CRS: Company Relevance Score 143
  5.A.2 ESS: Event Sentiment Score 144
  5.A.3 ENS: Event Novelty Score 144
5.B Industry-level sentiment data 144
  5.B.1 Company Relevance Score 144
  5.B.2 WLE: Word and phrase detection 144
  5.B.3 PCM: Projections, corporate news 144
  5.B.4 ECM: Editorials, commentary news 145
  5.B.5 RCM: Reports, corporate action news 145
  5.B.6 VCM: Merger, acquisitions, and takeover news 145
5.C References 145

PART II NEWS AND ABNORMAL RETURNS 147

6 Relating news analytics to stock returns 149
  David Leinweber and Jacob Sisk
  6.1 Introduction 149
  6.2 Previous work 150
    6.2.1 Behavioral basis 150
    6.2.2 Risk management and news 151
    6.2.3 Broad long-period analysis of the relation between news and stock returns 151
  6.3 News data structure and statistics 153
    6.3.1 Sample news data 153
    6.3.2 Descriptive news statistics and trends 153
  6.4 Improving news analytics with aggregation 153
    6.4.1 Event studies 153
    6.4.2 News analytic parameters for these studies 155
    6.4.3 Adjusting aggregate event parameters and thresholds, and segmentation by sector 157
    6.4.4 Adjusting sentiment thresholds 157
  6.5 Refining filters using interactive exploratory data analysis and visualization 158
  6.6 Information efficiency and market capitalization 162
  6.7 US portfolio simulation using news analytic signals 163
    6.7.1 Investment hypothesis 163
    6.7.2 Portfolio construction 164
6.7.3 Performance 165
6.7.4 Monthly performance 165
6.7.5 Portfolio characteristics 166
6.7.6 Return distribution 166
6.7.7 Portfolio beta and market correlation 166
6.8 Discussion of RNSE and portfolio construction 168
6.9 Summary and areas for additional research 170
6.9.1 Directions for future research. Is this just for quants? 170
6.10 Acknowledgments 171
6.11 References 171

7 All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors 173
Brad M. Barber and Terrance Odean
7.1 Related research 177
7.2 Data 179
7.3 Sort methodology 181
7.3.1 Volume sorts 181
7.3.2 Returns sorts 182
7.3.3 News sorts 183
7.4 Results 183
7.4.1 Volume sorts 183
7.4.2 Returns sorts 186
7.4.3 News sorts 190
7.4.4 Volume, returns, and news sorts 190
7.4.5 Size partitions 193
7.4.6 Earnings and dividend announcements 197
7.5 Short-sale constraints 197
7.6 Asset pricing: Theory and evidence 201
7.7 Conclusion 206
7.8 Acknowledgments 208
7.9 References 209

8 The impact of news flow on asset returns: An empirical study 211
Andy Moniz, Gurvinder Brar, Christian Davies, and Adam Strudwick
8.1 Background and literature review 211
8.1.2 Guided tour 213
8.2 Aspects of news flow datasets 213
8.2.1 Timeliness of news 213
8.2.2 Relevance of news 214
8.2.3 Classification of news 214
8.2.4 Independence of news 215
8.2.5 Informational content of news 216
8.3 Understanding news flow datasets 217
8.4 Does news flow matter? 219
8.5 News flow and analyst revisions 221
8.6 Designing a trading strategy 224
8.6.1 Turning a dataset into a trading signal 224
8.6.2 How to define the event? 224
8.6.3 What is the informational content of the event? 225
8.6.4 What is the holding period? 225
8.7 Summary and discussions 227
8.8 References 228

9 Sentiment reversals as buy signals
John Kittrell
9.1 Introduction 231
9.2 The quantification of sentiment 233
9.3 Sentiment reversal universes 235
9.4 Monte Carlo–style simulations 239
9.5 Conclusion 243
9.6 Acknowledgments 243
9.7 References 244

PART III NEWS AND RISK

10 Using news as a state variable in assessment of financial market risk
Dan diBartolomeo
10.1 Introduction 247
10.2 The role of news 248
10.3 A state-variable approach to risk assessment 250
10.4 A Bayesian framework for news inclusion 252
10.5 Conclusions 253
10.6 References 254

11 Volatility asymmetry, news, and private investors
Michal Dzielinski, Marc Oliver Rieger, and Tõnn Talpsepp
11.1 Introduction 255
11.2 What causes volatility asymmetry?
11.2.1 Measuring volatility asymmetry 256
11.2.2 Volatility asymmetry comparison 257
11.2.3 Market-wide causes for volatility asymmetry 258
11.2.4 Volatility asymmetry, news, and individual investors 259
11.3 Who makes markets volatile?
11.3.1 Google and volatility 261
11.3.2 Who’s in the market when it becomes volatile? 264
11.4 Conclusions 268
11.5 Acknowledgments 269
11.6 References 269

12 Firm-specific news arrival and the volatility of intraday stock index and futures returns
Petko S. Kalev and Huu Nhan Duong
12.1 Introduction 271
13 Equity portfolio risk estimation using market information and sentiment

Leela Mitra, Gautam Mitra, and Dan diBartolomeo

13.1 Introduction and background
13.2 Model description
13.3 Updating model volatility using quantified news
13.4 Computational experiments
   13.4.1 Study I
   13.4.2 Study II
13.5 Discussion and conclusions
13.6 Acknowledgements
13.A Sentiment analytics overview
   13.A.1 Tagging process
   13.A.2 Sentiment classifiers
   13.A.3 Score calculation
   13.A.4 Summary of classifiers and scores
13.B References

PART IV INDUSTRY INSIGHTS, TECHNOLOGY, PRODUCTS AND SERVICE PROVIDERS

14 Incorporating news into algorithmic trading strategies: Increasing the signal-to-noise ratio

Richard Brown
—So, how can one incorporate news into algorithmic strategies to improve trading performance?
—So, how does one increase the signal-to-noise ratio, ensuring protection from unforeseen exposures without an excessive number of halts or items to review?
—Sounds logical, right? So how exactly can this be done?
—So what about offensive strategies? How can one generate alpha using news?

15 Are you still trading without news?
Armando Gonzalez
—The underpinnings of news analytics
—Quantcentration and news
—Detecting news events automatically
—Finding “liquidity” in the news
16 News analytics in a risk management framework for asset managers
Dan diBartolomeo

17 NORM—towards a new financial paradigm: Behavioural finance with news-optimized risk management
Mark Vreijling and Thomas Dohmen
17.1 Introduction 319
17.2 The problem of incomplete information in market risk assessment 319
17.3 Refining VaR and ES calculation using semantic news analysis 320
17.4 The implementation of semantic news analysis 320
17.5 NORM goals 321
17.6 NORM uses semantic news analysis technology 321
17.7 Conclusion: NORM contribution to risk assessment 322

18 Question and answers with Lexalytics
Jeff Catlin

19 Directory of news analytics service providers
Event Zero 328
InfoNgen 330
Kapow Technologies 332
Northfield Information Services, Inc. 334
OptiRisk Systems 336
RavenPack 338
SemLab BV 340
The Chartered Institute for Securities & Investment 342
Thomson Reuters 344

Index 347