Understanding the quality of the narratives in corporate filings

Feng Li
University of Michigan

LSE/LUMS/MBS Conference
London, June 24, 2013
Roadmap

• Why studying narrative information (using computer programs)

• Framework

• Current research

• Future opportunities
Why narratives?
Analysis of narratives is nothing new

- Bible gospels authorship
- Suicide notes study by psychologists
- Was the rise of fascism tied to the contents in radio broadcasting?
- Spam filtering
- Machine translation
- Bioinformatics
- Searching for terrorists in online forums
- Customer feedback research
Important information source

- 207 pages
- 30 pages of tables
- The rest is narrative disclosures
  ~ Chairman’s letter to stockholders
  ~ MD&A
  ~ Notes to financial statements
  ~ ...

Typical 10-K: 300+ numbers, 30,000+ words
Help us understand quantitative data and firm disclosure behavior

• Data generating function
  ~ Notes to the financial statements
  ~ E.g., sales revenue increases, but revenue recognition method changed

• Can be forward-looking compared with many quantitative disclosures

• Provide a richer environment to test disclosure theories

“We have incurred significant losses since our inception, and we expect to continue to incur net losses for the foreseeable future.”
Understanding managers’ cognitive processes

- How do you measure cognitive/behavioral characteristics using archival data?
- Attributions: concepts, attitudes, beliefs, intentions, emotions, mental states and cognitive processes.
- Social relationships: authority, power.
Example: Manager heuristics (self-attribution bias)

American International Group, Inc. 2006 Annual Report

• “Solid execution of our strategies and the absence of significant catastrophes contributed to our outstanding results in 2006. Around the world and across all of our business segments we are capitalizing on growth opportunities, using our business diversity and matrix management structure to respond quickly to customer needs.”
Example: Manager heuristics (self-attribution bias)

American International Group, Inc. 2008 Annual Report

• “AIG reported that the continued severe credit market deterioration, particularly in mortgage-backed securities, and charges related to ongoing restructuring activities, contributed to a record net loss for the fourth quarter of $61.7 billion, or $22.95 per diluted share, compared to a 2007 fourth quarter net loss of $5.3 billion, or $2.08 per diluted share.”
Human vs. machine

• Human coding
  ~ Very costly (time and money)

• Recent technology has made quantitative measurement of content easier
  ~ Digitalization of the record
  ~ Computing power to analyze and record language
Framework
Essence: data reduction

A Miracle Occurs

Numbers

Ford 2000 10-K (113 pages)

“Tone” = -0.62

2001 poor performance
Two dimensions

Characteristics of the narratives

Amount
Transparency
Optimism
Cognitive traits
Similarity
...

...
Two dimensions

Characteristics of the narratives

- Amount
- Transparency
- Optimism
- Cognitive traits
- Similarity
- ...

Economic hypotheses

Determinants

Consequences
GE’s “greatometer” (Scott Davis, Morgan Stanley's lead GE analyst)

- In 2002 Q3 call, Messrs. Immelt and Sherin said "great" more than 20 times.
- 2005 Q2, 70 times (GE shares rose 37%)
- 2006 Q3, 37 times (GE stock fell 10%)

• But … 1 number might > 1000 words
Narrative information as contextual variable

• Combine narratives with quantitative information

• GE’s “greatometer” may help us understand its earnings quality
  ~ When GE CEO uses “great” more often, its earnings may have higher quality
Current research
(Cole and Jones 2005; Li 2011)
What has been done?

Characteristics of the narratives:
- Amount
- Transparency
- Optimism
- Cognitive traits
- Similarity
- …

Economic hypotheses:
- Determinants
- Consequences

Number of words (Li 2008)

Readability (Li 2008, Loughran and McDonald 2010, Miller 2010)

Li 2010, Davis et al. (2011), Rogers et al. (2012)
What has been done?

Characteristics of the narratives:
- Amount
- Transparency
- Optimism
- Cognitive traits
- Similarity
- ...

Economic hypotheses:
- Determinants
  - Strategic disclosure and impression management: Li (2008) and Huang et al. (2013)
- Consequences
What has been done?

Characteristics of the narratives
- Amount
- Transparency
- Optimism
- Cognitive traits
- Similarity

Economic hypotheses
- Stock market impact (Davis et al. 2009, Li 2009, Feldman et al. 2010)
- Litigation (Rogers et al. 2012)
- Impact on financial analysts’ information processing efficiency (Lehavy et al. 2012)
Current methodology: often preliminary

- Bag of words representation

- A sentence or a document is represented as an (unordered) collection of word.
  ~ disregarding grammar and even word order.

- “Tom ate the wolf” = “The wolf ate Tom”

- Frequency count of specific words based on dictionaries
Future opportunities
What can be done?

Characteristics of the narratives

- Amount
- Transparency
- Optimism
- Cognitive traits
- Similarity

Economic hypotheses

Determinants

Consequences

Manager over-confidence, self-attribution bias, short-termism, power, narcissism
What can be done?

Economic hypotheses
Determinants
Consequences

Characteristics of the narratives

Amount
Transparency
Optimism
Cognitive traits
Similarity

Change specification!!!
Example

• Time-series comparison
  ~ BP 2012 annual report compared with its 2011 report

• Cross-sectional comparison
  ~ BP annual report compared with that of Exxon Mobil (e.g., they have different shale gas reserve booking rules)
Cosine similarity

- Distance between vectors \( d_1 \) and \( d_2 \) captured by the cosine of the angle \( \theta \) between them.
Cosine similarity

\[ \text{sim}(d_j, d_k) = \frac{\vec{d}_j \cdot \vec{d}_k}{\|\vec{d}_j\| \|\vec{d}_k\|} = \frac{\sum_{i=1}^{n} w_{i,j} w_{i,k}}{\sqrt{\sum_{i=1}^{n} w_{i,j}^2} \sqrt{\sum_{i=1}^{n} w_{i,k}^2}} \]

- Cosine of angle between two vectors
- The denominator measures the lengths of the vectors.
What can be done?

Characteristics of the narratives

- Amount
- Transparency
- Optimism
- Cognitive traits
- Similarity
- ...

Economic hypotheses

Determinants
- Cost of capital
- Fraud detection
- Bankruptcy prediction
- Social responsibility commitment
- Strategy

Consequences
What can be done? Fundamental analysis

• Can we mimic Warren Buffett using computer programs?

“Other guys read Playboy, I read annual reports.”
Fundamental analysis: promising direction

<table>
<thead>
<tr>
<th>Consolidated Statements of Income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>(In millions, except per share amounts)</td>
</tr>
<tr>
<td>REVENUES</td>
</tr>
<tr>
<td>OPERATING EXPENSES:</td>
</tr>
<tr>
<td>Salaries and employee benefits</td>
</tr>
<tr>
<td>Purchased transportation</td>
</tr>
<tr>
<td>Rentals and landing fees</td>
</tr>
<tr>
<td>Depreciation and amortization</td>
</tr>
<tr>
<td>Fuel</td>
</tr>
<tr>
<td>Maintenance and repairs</td>
</tr>
<tr>
<td>Impairment and other charges</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>OPERATING INCOME</td>
</tr>
<tr>
<td>OPERATING INCOME inadequate provision</td>
</tr>
</tbody>
</table>

Search for all the sentences that have “sales” or “revenues” as NP, and then extract the VP, ADJP, and ADVP in these sentences for further analysis (e.g., “increase”)

MICHIGAN ROSS SCHOOL OF BUSINESS
What can be done? more structured approach in terms of algorithms

- Par-of-speech (POS) tagging
  - Rule-Based tagging (Voutilainen 1995)
  - Stochastic (e.g., Hidden Markov Model) tagging (Brants 2000)
  - Transformation-based tagging (Brill 1995)
Stanford NLP parser used in Chen and Li (2013)

http://nlp.stanford.edu/software/lex-parser.shtml
Chen and Li (2013)

• “PCFG” (Probabilistic context-free grammars)
  - Direct object (verb or object)
    - “Estimate the receivables”
  - Passive nominal subject (verb or object)
    - “Receivables are estimated as”
  - Adjective modifier
    - “Likely loss”
  - Noun compound subjects
    - “Estimation value”
  - Quantifier phrase modifier
    - “Is approximately $100 million”
Conclusion

• This is an important area with significant research potential!

• E-mail me at feng@umich.edu if you are interested.

• Thank you!