THE DETERMINANTS OF VOLUNTARY DISCLOSURE IN XBRL FINANCIAL STATEMENTS

XBRL Europe Days
June 2nd, 2016 – Frankfurt

Pierre Teller (University of Nice, France)
Dominique Dufour (University of Nice, France)
Philippe Luu (University of Nice, France)
Eric Séverin (University of Lille 1, France)
abstract

- USA: Financial reports are now sent to the SEC in XBRL (eXtensible Business Reporting Language) files
- Most of these files use the USGAAP taxonomy (standard set of tags, created by the FASB)
- But the filer can create custom tags (extensions) if deemed necessary
- Object of this study: we searched the determinants of extension creation (seen as a matter of voluntary disclosure)
Summary

1. XBRL (definition, recent evolutions)
2. XBRL taxonomies and voluntary disclosure
3. Our model and data collected
4. Results
5. Conclusion and discussion
XBRL

• XBRL is one of the many ways to improve data management (for financial data)

• Non exhaustive list of some common problems:
  – Too much data, and data quantity goes up exponentially
  – Reporting frequency & information needs are getting bigger
  – Data often poorly contextualized
  – IT tools still not efficient to deal with the **semantic** of data (digital data is still human readable, not machine readable)

• => growing need of adapted IT tools: IT is the cause AND the solution to this “data deluge”
XBRL & knowledge mgt

• XBRL is one of these IT tools, a standardized way to create & collect financial data

• An answer to the major issues of knowledge management:
  1. Make it possible to process more data within less time (easy): OK
  2. Allows processing of data of all kinds (harder): OK
  3. Facilitate or automate the analysis of data (much harder): OK???
XBRL: technical aspects 1/2

- XBRL is created upon XML, standard markup meta-language normalized by W3C
- Standardized representation: pieces of information stored in shared taxonomies (dictionaries of tags)
- Very rigorous formalism (essential to create machine readable data): tree structure with nested tags, the last opened tag is closed first
XBRL: technical aspects 2/2

- Can store various data types
- Opening tags can be used to contextualize a piece of data
- XBRL formulas can be used to check validity
Recent evolutions

- First, adoption on a voluntary basis...
- In the USA, XBRL filing became mandatory for all listed companies (between 2010 and 2012)
- Files publicly available on the “Edgar Cloud”
- In Europe:
  - Support of the European Commission
  - Used in different countries (UK, Spain, Belgium, Netherlands, etc.) for different purpose
Taxonomies / instance documents

- Facts are stored in an **instance document**
- **Taxonomy** = dictionary of tags
- Several reference taxonomies (IFRS, USGAAP, etc.), national or industry specific, etc.
- Can use several taxonomies in the same instance document
- The filer can use custom tags, creating *de facto* a **custom taxonomy**
is extension creation an issue?

- There are pros and cons
  - Comparability Vs Relevance
- Many discussions and debate
- Our goal is not to “decide” if extensions are good or bad
- Not even to find the “best proportion” of extensions in an XBRL file
- It is just to investigate the factors encouraging a filer to create extensions
Research question

• eXtensible → the filer can create custom tags (extensions)
• Proportion of extensions can be really different (ranging from 0% to 81% in our sample, 16% on average)
• This additional work is a voluntary choice of the filer → we considered that it is a matter of voluntary disclosure
• We used the definition given by Pourtier (2004): “we considered that anything that has not been either standardized nor made mandatory is therefore voluntary”
• Goal of this work: identify the determinants of voluntary disclosure in XBRL filings
MODEL & DATA
XBRL and the agency theory

• In this conceptual framework, voluntary disclosure is a mean to reduce agency costs

• Our hypothesis:
  – H1: extensions (%) are positively correlated with the size of the filer
  – H2: extensions (%) are positively correlated with debt ratio
  – H3: extensions (%) are positively correlated with ownership diffusion
Control variables

• Taken from both agency and signal theory
• Second model (robustness tests) uses the following variables
  – Liquidity: results from a good management, it is therefore an incentive for broader communication
  – Profitability: a more profitable firm tends to communicate more information to its environment
  – Auditors reputation: the reputation and expertise of the auditors encourages voluntary disclosure
## Proxys

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Name</th>
<th>Proxy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leverage</td>
<td>Lev</td>
<td>Ratio of Debt over total assets (Depoers, 2000)</td>
</tr>
<tr>
<td>Firm Size</td>
<td>Size</td>
<td>Logarithm of market capitalization (Petersen and Plenborg, 2006)</td>
</tr>
<tr>
<td>Ownership Dispersion</td>
<td>OD</td>
<td>Logarithm of the number of shares owned by shareholders (Lan et al., 2013)</td>
</tr>
<tr>
<td>Liquidity</td>
<td>Liq</td>
<td>Ratio of Short-term over Long-term assets (Lan et al., 2013)</td>
</tr>
<tr>
<td>Profitability</td>
<td>Prof</td>
<td>Ratio of Operating Income over Assets (Petersen and Plenborg, 2006)</td>
</tr>
<tr>
<td>Auditor Size</td>
<td>Big4</td>
<td>Dummy variable (1 for Big Four, 0 Otherwise) (Depoers, 2000)</td>
</tr>
</tbody>
</table>
Sample & methodology

• XBRL filings from the EDGAR platform, in 2012: 7817 files
• Finding the tickers of the filers, from the CIK code of the SEC: 5904 tickers found
• Searching financial data in the database Thomson One, removing incomplete data: 2924 observations remaining

• To this aim, we created 2 computer programs:
  – One to count the number of tags of a specific taxonomy in an XBRL file (in Java)
  – One to get the Thomson Ticker from the CIK code of the SEC (Python)
Models

[1] PCT\_EXT = \beta_0 + \beta_1 \text{LEV} + \beta_2 \text{SIZE} + \beta_3 \text{OD} + \varepsilon

[2] PCT\_EXT\_AJU = \beta_0 + \beta_1 \text{LEV} + \beta_2 \text{SIZE} + \beta_3 \text{OD} + \varepsilon

[3] PCT\_EXT\_AJU = \beta_0 + \beta_1 \text{LEV} + \beta_2 \text{SIZE} + \beta_3 \text{OD} + \beta_4 \text{PROF} + \beta_5 \text{LIQ} + \beta_6 \text{BIG4} + \varepsilon

- Model 2: percentage of sector-adjusted extensions
- Model 3: addition of the control variables
### Variables exogènes

<table>
<thead>
<tr>
<th>Variables exogènes</th>
<th>Signe attendu</th>
<th>Modèle 1</th>
<th>Modèle 2</th>
<th>Modèle 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$Y = \text{PCT_EXT}$</td>
<td>$Y = \text{PCT_EXT_AJU}$</td>
<td>$Y = \text{PCT_EXT_AJU}$</td>
</tr>
<tr>
<td>LEV</td>
<td>+</td>
<td>0,000021</td>
<td>0,000023</td>
<td>0,000005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0,000028)</td>
<td>(0,000028)</td>
<td>(0,000031)</td>
</tr>
<tr>
<td>SIZE</td>
<td>+</td>
<td>0,024100 ***</td>
<td>0,022820 ***</td>
<td>0,022570 ***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0,000817)</td>
<td>(0,000808)</td>
<td>(0,000923)</td>
</tr>
<tr>
<td>OD</td>
<td>+</td>
<td>0,017570 ***</td>
<td>0,015920 ***</td>
<td>0,015930 ***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0,00132)</td>
<td>(0,00130)</td>
<td>(0,00133)</td>
</tr>
<tr>
<td>PROF</td>
<td>+/-</td>
<td>0,000020</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0,000014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIQ</td>
<td>+/-</td>
<td></td>
<td>-0,000908 ***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0,000297)</td>
<td></td>
</tr>
<tr>
<td>BIG4</td>
<td>+/-</td>
<td>0,001290</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0,00469)</td>
<td></td>
</tr>
</tbody>
</table>

| N                  | 2924          | 2924      | 2924      |
| Adjusted R²        | 0,2287        | 0,2145    | 0,2159    |
| F statistic        | 289,96        | 265,76    | 135,16    |
| p-Value            | <0,0001       | <0,0001   | <0,0001   |

* Significant at .10  
** Significant at .05  
*** Significant at .01
results

• H1 and H3 are confirmed
  – Larger firms tend to have more extensive reporting expectations
  – Ownership dispersion also make firm more likely to voluntary disclose

• But H2 is not!
  – Voluntary disclosure should be a mean to reduce the monitoring costs resulting from a High debt ratio, but this relationship is not statistically significant
Extensions & limits

• Possible extensions:
  – Include several years in our sample
  – Evolution of the extension proportion over the years
  – Distinguish extensions between financial statements
categories (income statement, assets, equity, etc.)

• Limitations:
  – 2012 data → was XBRL reporting “mature” enough?
  – we do not address the relevance of extensions, we
    suppose that extensions are always created to improve
    financial statements

    – is it possible to enhance relevance and comparability at the
      same time?
THANK YOU FOR YOUR ATTENTION