Summary of the Paper

Motivation

“The world’s most valuable resource is no longer oil, but data” (The Economist 2017)

• Public and policy debate show increasing interest in the power of data accumulating firms like Google, Facebook, and Amazon and its implications for consumers
• Almost no empirical research on data firms

Research questions

1. How to identify data-rich firms in a systematic approach?
2. Is it more probable for data firms to engage in M&A?

Methods

• Textual analysis
• (Multinomial) logistic regressions

Results

The proposed measure reflects expected characteristics of data intensive firms.

• Higher data intensity of firms corresponds to a higher probability of being either an acquirer or target in an M&A transaction.
• Attention by competition authorities for data intensive acquirers is lower if the target is small, but higher if the target is public.

Data Economy and Mergers & Acquisitions

Daniela S. Schoch
schoch@dwl.lmu.de

Data firm measure

General idea and procedure

Firms that share specific characteristics with known data firms are defined as data intensive firms.

1. Find firms that are recognized as data firms in antitrust cases by FTC/EU Commission.

• Characteristic: privacy concerns due to higher data concentration within combined firm after M&A/Joint Venture.
• Found antitrust cases of data firms:
  - Google/DoubleClick acquisition 2008
  - CoreLogic/DataQuick acquisition 2013
  - Facebook/WhatsApp acquisition 2014
  - Microsoft/LinkedIn acquisition 2016
  - Sandefi/Google/DMI joint venture 2016

2. Measure similarity of business descriptions and risk factors to the firms found in step I

Measurement

DataIntensity: standardized sum of mean cosine similarities to the five firms of each: 1. 10-K business descriptions (Item 1)
2. 10-K risk factors (Item 1A)
3. Data firm related keywords in 10-K business descriptions (Item 1)
4. Data firm related keywords in 10-K risk factors (Item 1A)

Used keywords

For business descriptions, e.g., advertising, user, data, network, server

For risk factors, e.g., data protection, laws regarding privacy, security breach

Cosine similarity (Hobing & Phillips 2016)

a) Construct list of nouns that appear in less than 25% of all 10-K business descriptions (risk factors) in one fiscal year
b) For each firm, a dummy variable Pi, indicates whether a noun of the list constructed in a) occurs in its business description or description factors (1) or not (0)
c) Cosine similarity: similarityij = Pi · Pj · (Pj)0.5 · (Pi)0.5

Data firm dummy: constructed, since an inflated data intensity measure is probable due to discussion of data relevant issues independent of actual business activity.

Procedure

1. Manually classify 400 (1%) randomly selected 10-K’s into data and non-data firms

21 data firms and 379 non-data firms (45 data firms and 355 non-data firms)

2. ‘K-nearest neighbour’ algorithm sorts rest of the sample into binary data firm measure

5.85% data firms (2,260 firm filings) published in

French 49 Industries (>500 firms)

Results

Descriptive statistics

Firms

Data firm dummy

Table

Do known data firms have high scores?

• Highest score in the sample: ‘Yahoo’
• Among top 20 data firms are: ‘Yahoo’, ‘Zynga’
• U.S. firms in Deutshe Bank Cloud & Big Data Index: 2% firm-year observations in 10th percentile

Is it more likely for data firms to invest in M&A?

• Regulatory attention, multiples, and returns

• Multinomial logit regression: no change in results for the data firm variables when controlling for firm age, life cycle stage, and competition.

Conclusions

• The data firm measure based on textual analysis reflects expected data firm characteristics: data firms are younger, have higher market to book ratios and known data firms score high.
• Data intensive firms are more likely to invest in M&A.
• Data intensive acquirers get less attention by competition authorities if they buy small firms. However, if they buy large firms, data intensive acquirers are more likely to require approval by competition authorities.
• Deal value to sales multiples and combined announcement returns are not different for data firms compared to non-data firms.