Data Economy and Mergers & Acquisitions

Daniela S. Schoch
schoch@bwl.lmu.de

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.

Building blocks of data firms/markets

Typical characteristics

- Often platform based (two-/many-sided markets)
  → Use data to better match different sides of the market

- High economies of scale and network effects
  → Can increase entry barriers and simplify monopolization

- Sometimes markets even have "winner-take-all" characteristics
  → Natural monopolies as a probable outcome of competition

Motives for investing in M&A

- Particularly high incentives due to data as valuable resource/asset and data market characteristics
  → Combine complementary resources and exploit economies of scale
  → Enhance market power to intensify network effects

- Difficulties of competition authorities to evaluate such data firm M&A

- Typical consolidation activity in young industries

Data

- Annual reports (10-K filings) published in 2006-2018 for fiscal year end dates between 01 July 2006 – 30 June 2018 (SEC Edgar files from Loughran and McDonald (2019))
- M&A deals (Thomson SDC)
- Accounting data (Compustat)
- Stock data (CRSP)

Data firm measure

General idea and procedure

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

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

Data Firm Def.

→ 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
- Sanofi/Google/DMI joint venture 2016

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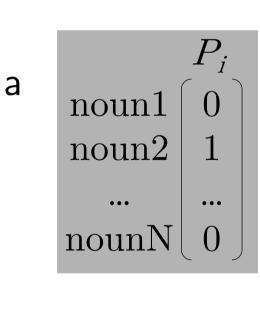
Measurement

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** (Hoberg & 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 vector P_i indicates whether a noun of the list constructed in a) occurs in its business description or risk factors (1) or not (0)
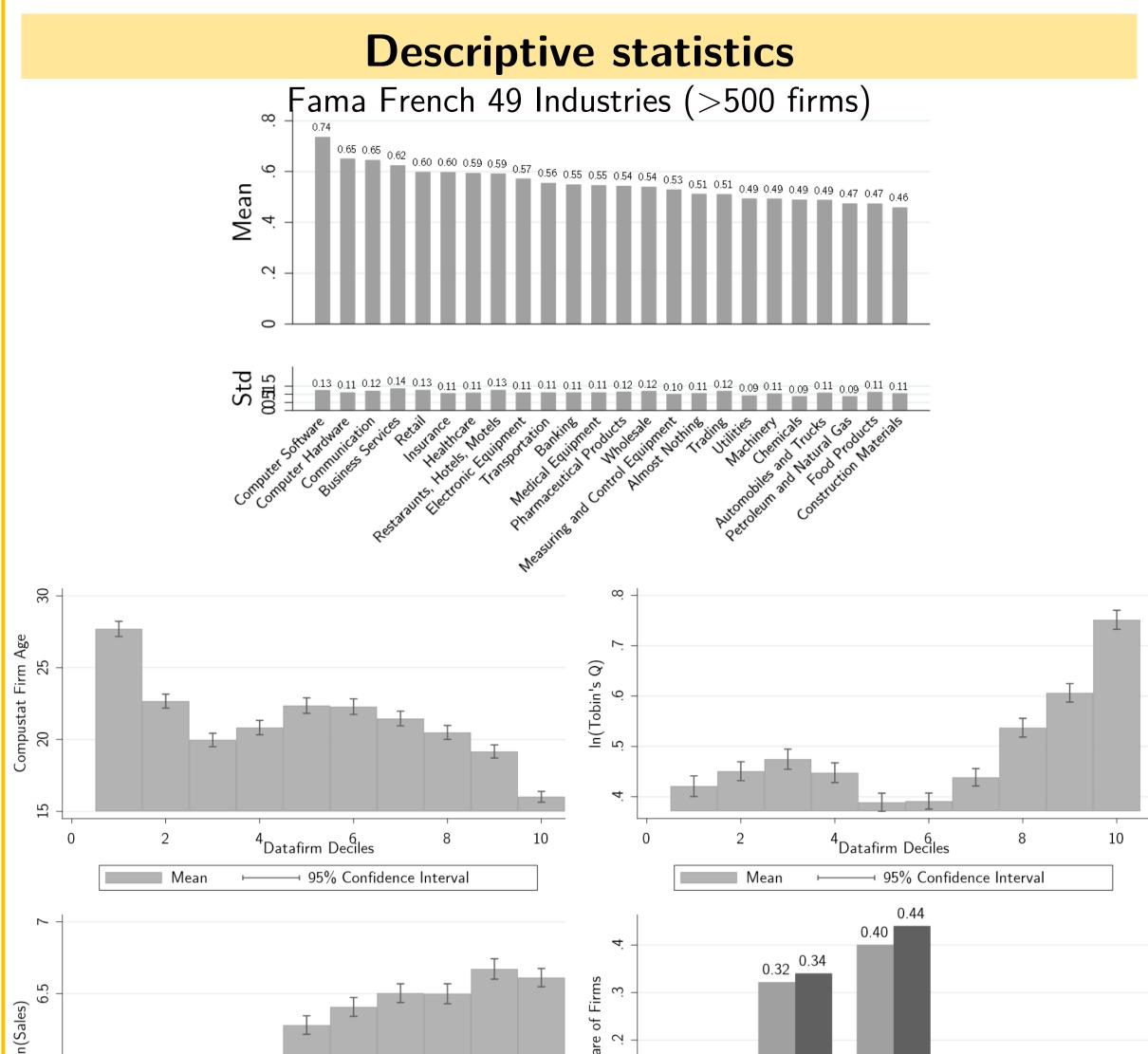
$$P_i \begin{matrix} \text{noun1} \\ \text{noun2} \\ ... \\ \text{nounN} \end{matrix} \begin{bmatrix} 0 \\ 1 \\ ... \\ 0 \end{bmatrix}$$

c) Cosine similarity:
$$similarity_{ij} = P_i' P_j (P_i' P_i)^{-0.5} (P_j' P_j)^{-0.5}$$

DatafirmDummy: 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 (more loose specification including also 'data-analytics or -storage firms': 45 data firms and 355 non-data firms)

2. '*k*-nearest neighbor' algorithm sorts rest of the sample into binary data firm measure
   → 5.85% data firms (2,260 firm-fiscal year observations)

Results

Descriptive statistics

Fama French 49 Industries (>500 firms)

Do known data firms have high scores?

- Highest score in the sample: Adobe
- Among top 20 data firms are: YAHOO!, tripadvisor, zynga
- f G LinkedIn : 8th decile in '06-07, 9th in '08-09, else 10th
- amazon: 10th decile, besides '08 and '16, ebay: 10th decile
- U.S. firms in *Deutsche Bank Cloud & Big Data Index*: 2% firm-fiscal year obs in 8th, 10% in 9th, 88% in 10th decile

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

DV: TransactionParty	Acquirer			Target		
DataIntensity	0.164***			0.036**		
	[5.290]			[2.910]		
DatafirmDummy StrictDef		-0.000			0.002	
		[-0.035]			[0.327]	
DatafirmDummy LooseDef			0.018*			0.011**
			[2.009]			[2.871]
lnMarketToBook	0.046***	0.048***	0.047***	-0.019***	-0.019***	-0.019***
	[6.973]	[7.223]	[7.134]	[-7.523]	[-7.292]	[-7.438]
BenchmarkDum	yes	yes	yes	yes	yes	yes
Constant	yes	yes	yes	yes	yes	yes
Year/Industry FE	yes	yes	yes	yes	yes	yes
# Observations	36,461	36,461	36,461	36,461	36,461	36,461
# Firms	5,352	5,352	5,352	5,352	5,352	5,352

[Z scores]; ***p<0.001, **p<0.01, *p<0.05, +p<0.1.

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

Regulatory attention, multiples, and returns

Dependent Variable:	Competition Authority Approval Required					
Datafirm Variable	Intensity	StrictDef	LooseDef	Intensity	StrictDef	LooseDef
Datafirm	-0.041	-0.005	-0.006	-0.167***	-0.031	-0.033
	[-1.537]	[-0.477]	[-0.626]	[-3.614]	[-1.040]	[-1.553]
PublicTargetDummy	0.059**	0.135***	0.133***	0.056***	0.088***	0.087***
	[3.029]	[27.872]	[26.864]	[2.845]	[20.243]	[19.422]
DatafirmxPublicTarget	0.132***	0.025	0.028*	0.054	0.008	0.011
	[4.037]	[1.541]	[2.296]	[1.621]	[0.505]	[0.924]
isDealValue				0.034	0.131***	0.129***
				[1.218]	[16.797]	[16.228]
DatafirmxisDealValue				0.170***	0.034	0.035+
				[3.627]	[1.234]	[1.758]
AcquirorlnSales	0.010***	0.009***	0.009***	0.014***	0.014***	0.014***
	[7.336]	[7.586]	[7.641]	[11.306]	[11.621]	[11.670]
BenchmarkDummies	yes	yes	yes	yes	yes	yes
Constant, Year/Industry FE	yes	yes	yes	yes	yes	yes
# Observations	18,407	18,407	18,407	18,407	18,407	18,407

[Z scores]; ***p<0.001, **p<0.01, *p<0.05, +p<0.1.

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.