

Large-scale Open Corporate Data Collection and Analysis as an Enabler of Corporate Social Responsibility Research

Vasiliki Gkatziaki¹, Symeon Papadopoulos¹, Sotiris Diplaris¹ and Ioannis Kompatsiaris¹

¹ Information Technologies Institute (ITI), CERTH-ITI, Thessaloniki Greece

**4th International Conference on Internet Science,
22-24 November 2017, Thessaloniki, Greece**



“

*“A world financial registry would deal
a fatal blow to financial secrecy.”*

*Gabriel Zucman - assistant professor of economics at
University of California*

Introduction

The last decades, **citizens** and **transparency initiatives** put increasing pressure on **governments**, **organizations** and **companies** to be **more transparent**.



WikiRate*

- ⦿ A platform for collecting and analyzing information about companies **Environmental, Social and Governance Performance**.

Securities and Exchange Commission (SEC)

- ⦿ Requires US public companies to **file statements** regarding their **financial performance**.

* <http://wikirate.org/>

Main Research Question

How can open financial data be leveraged to extract useful inferences regarding the social performance of companies?



Related Work

◎ Numerous studies have been conducted researching **different aspects** of **Corporate Social Responsibility (CSR)**:

Does the **country of origin** or the type of **industry influence** the **disclosure** of a company's CSR?
(Wanderley et al.)

Does **CSR engagement** **effect** the value of a company that operates on a **controversial industry**?
(Cai et al.)

In what ways **sustainable ratings** **affect consumers**?
(Parguel et al.)

Related Work

Corporate Social Performance - Corporate Financial Performance

Do **strict pollution controls** translate to **less profit?**

(Bragton and Marlin)

Which are the most **effective CSR aspects** over the **financial performance** of a firm?

(Margolis et al.)

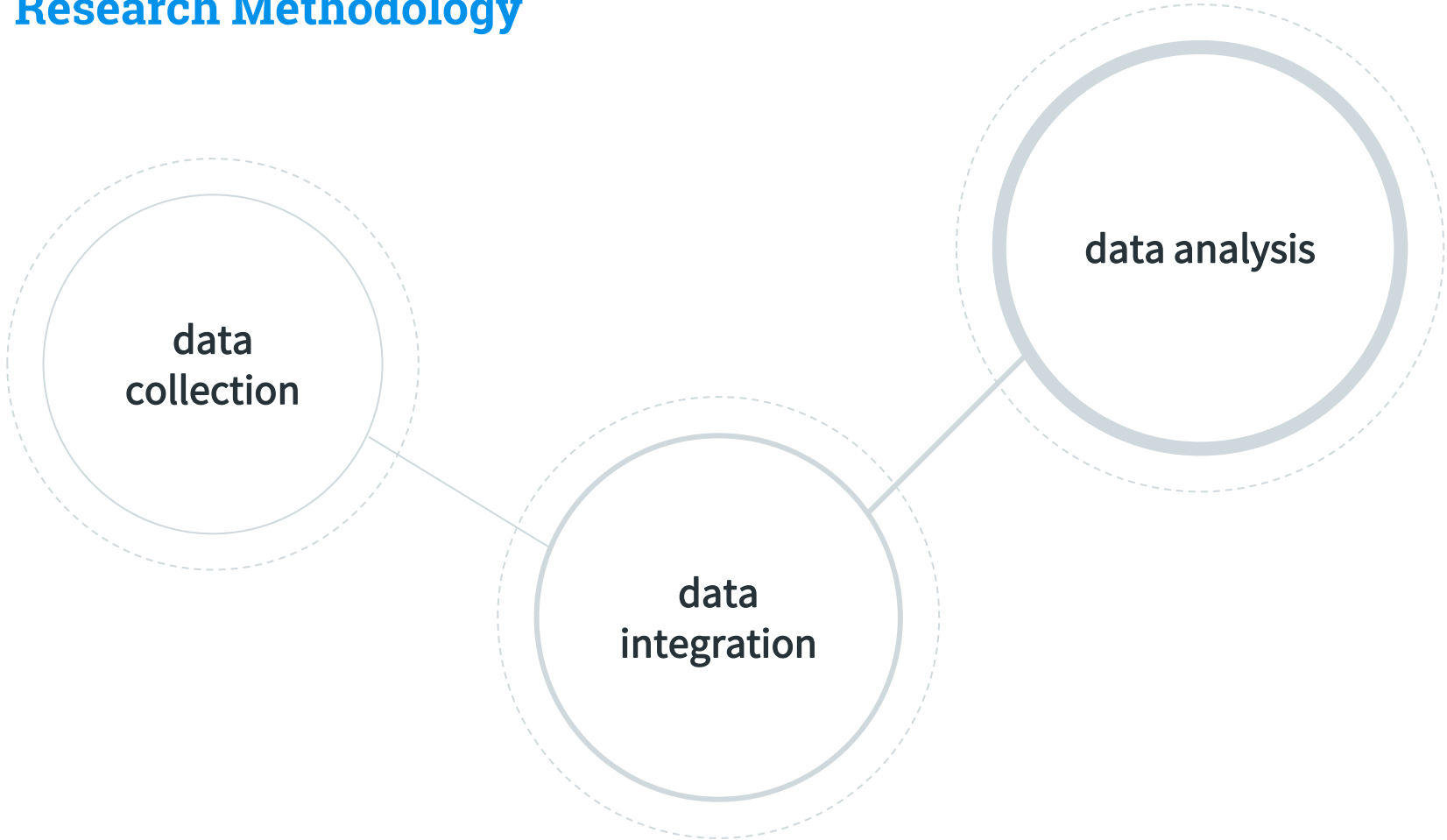
Can **stock market reaction** to problematic behavior of companies **function as an instrument of social control?**

(Bromiley and Marcus)

In what ways **CSP affect** a firm's ability to **access capital?**

(Cheng et al.)

Research Methodology





DATA COLLECTION & EXTRACTION FROM SEC

- *Conflict Minerals Report*
 - companies are obliged to disclose in these documents their use of conflict minerals considered necessary to production.
- *Form 10-K*
 - these documents make available audited financial statements regarding the annual financial performance of the company



50,219

US companies

3,442


Conflict Minerals Reports

162,349

Form 10-K filings

39,029

Form 10-K filings in XBRL
format



A background of a network graph with nodes and edges in light gray.

529,000*

financial facts were extracted related to
different aspects of financial performance of
companies

*all the extracted data are available through a RESTful API:

http://easie.itι.gr/sec_dataset/

Descriptive Statistics



COMPANIES DEMOGRAPHICS

16% located in **California**

16% in **New York**

8% in **Texas**

52% incorporated in **Delaware**

22% incorporated in **Nevada**

Delaware is considered as a domestic tax haven and its paradigm has been extensively studied (Dyreng et al., Wayne et al.)

COMPANIES INDUSTRIES

43% belong to Finance,
Insurance and Real Estate
Industry

20% to Manufacturing

12% to Services

7% to Transportation,
Communications, Electric, Gas
and Sanitary Services

REPORT Conflict Minerals

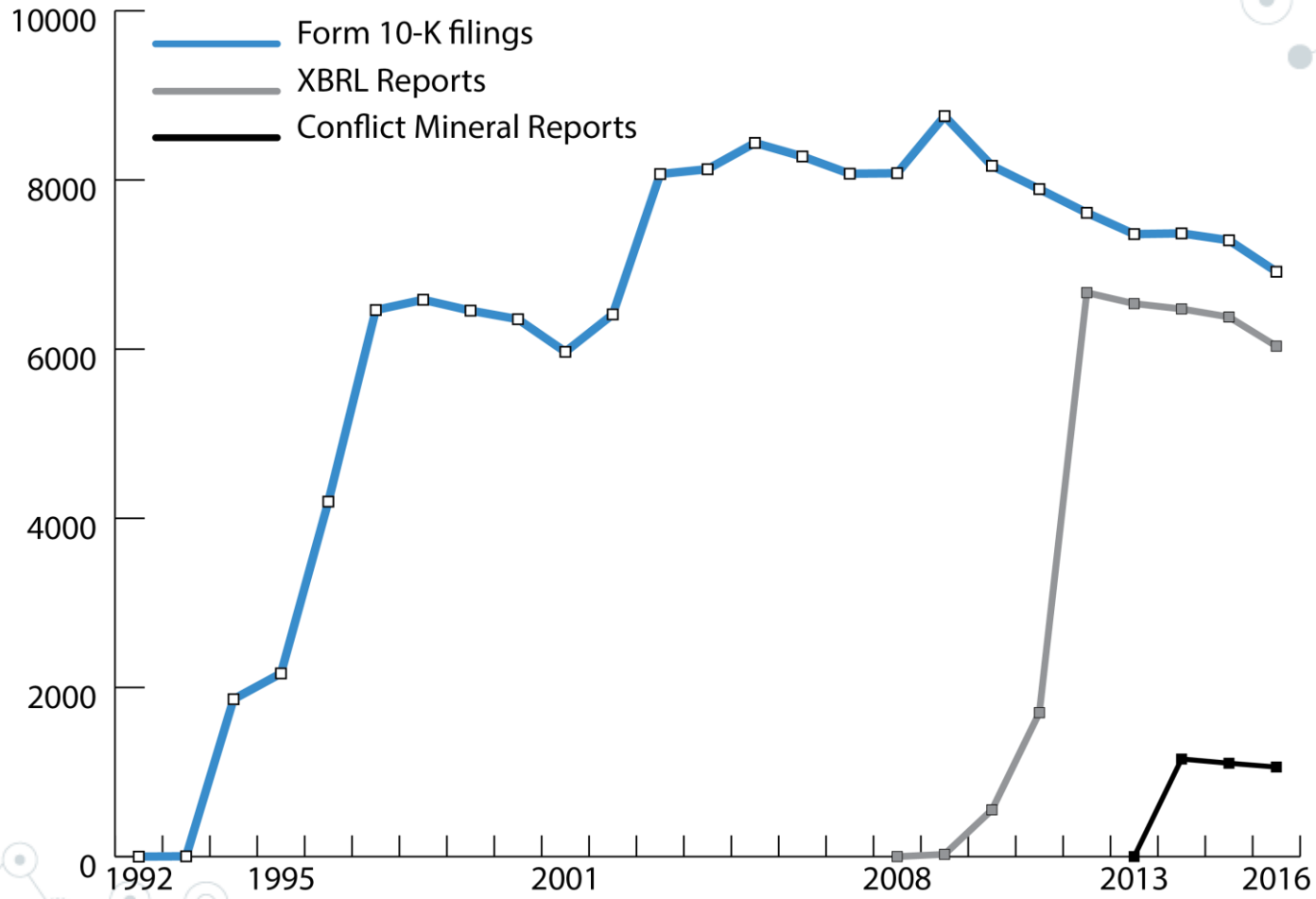
76% belong to Manufacturing

7% to Retail Trade

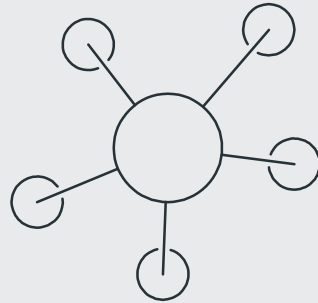
7% to Services



Collected Reports per Year



Statistical Inferences on CSR Performance



Datasets

Financia Facts Dataset

25,500 financial facts
about 7,700 US companies

Newsweek Green Rankings Dataset

Newsweek Green Scores (2016)
about 465 US companies
combined with several financial
statems from 10-K filings.

In total, 4 hypotheses were studied.

Hypothesis 1

There is no correlation between the reported undistributed earnings of foreign subsidiaries, the profits and the R&D expenses

- ◎ Kolmogorov-Smirnov and Shapiro-Wilk tests for normality showed that the undistributed earnings of foreign subsidiaries, profits and R&D expenses are significantly non-normal with $p < .001$
- ◎ Spearman's rho correlation coefficient showed a significant positive association between all three variables and the rejection of the null hypothesis 1.
- ◎ The correlation of R&D expenses appears to be stronger to undistributed earnings of foreign subsidiaries ($r_s = 0.582$) than to profits ($r_s = 0.135$) in terms of coefficient value.

Top 10 companies ranked by the amount of profits in 2016

In many cases the undistributed earnings of foreign subsidiaries **exceed** the amount of profits.

Company	Industry	State of Inc.	Profits (in \$M)	UEFS (in \$M)
Apple Inc.	Manufacturing	California	\$45,687	\$109,800
JP Morgan Chase & Co.	FIRE	Delaware	\$24,733	\$38,400
Berkshire Hathaway Inc.	FIRE	Delaware	\$24,427	\$12,400
Wells Fargo & Co.	FIRE	Delaware	\$22,045	\$2,400
Alphabet Inc.	Services	Delaware	\$19,478	\$60,700
Bank of America Corp.	FIRE	Delaware	\$17,906	\$17,800
Microsoft Corp.	Services	Washington	\$16,798	\$124,000
Walmart Stores Inc.	Retail Trade	Delaware	\$15,080	\$26,100
Citigroup Inc.	FIRE	Delaware	\$14,975	\$47,000
Gilead Sciences Inc.	Manufacturing	Delaware	\$13,488	\$37,600

Hypothesis 2 (I)

There is no difference between the means of the reported undistributed earnings of foreign subsidiaries between the years 2013, 2014, 2015 and 2016.

- ⊙ Our data **failed** the assumptions of **normality** for each category of the independent variable (year) as well as the **homogeneity of variances** assumption.
- ⊙ The **non-parametric test** of **Kruskal-Wallis** was used to detect differences between the groups.
- ⊙ The **undistributed earnings of foreign subsidiaries** were significantly different over the years, $H(3) = 14.867$; $p < .01$ and that led to the **rejection of the null hypothesis 2**.

Hypothesis 2 (II)

There is no difference between the means of the reported undistributed earnings of foreign subsidiaries between the years 2013, 2014, 2015 and 2016.

© **Mann-Whitney tests** were used to reveal **between which years** there was a significant difference in the undistributed earnings of foreign subsidiaries.

© **Significant differences** in the undistributed earnings of foreign subsidiaries were detected only **regarding years 2013 to 2015** ($U = 372062.5$; $r = -.067$) **and 2016** ($U = 361298.0$; $r = -.086$)

© **Jonckheere's test** revealed a **significant trend** in the data: the median of the undistributed earnings of foreign subsidiaries are **significantly increased over the years** ($J = 2621883.0$; $z = 3.790$; $r = .063$).

Hypothesis 3

There is no correlation between the Newsweek green score with profits, Good Will and R&D expenses.

- ◎ Spearman's correlation coefficient was used to discover correlations and the strength of the existing relationships between them.
- ◎ There was a **strong positive correlation** between all variables and the highest coefficient was occurred in the case of profits with $r_s = .401$
- ◎ Consequently, we could claim that **big companies**, in terms of financial performance, **tend to adopt more sustainable practices.**

Hypothesis 4

There is no difference in the environmental performance of companies reporting conflict minerals with those that do not.

- © Mann-Whitney's non-parametric test was performed since our data violated the normality assumption.
- © Results indicated that the environmental performance of companies that report conflict minerals is significantly higher than those companies that do not ($U = 20358$; $r = -0.15$; $p < .01$).



Conclusions

we managed:

- ◎ to **collect** numerous financial and conflict minerals reports
- ◎ to **extract** a plethora of financial facts
- ◎ to **make data available** to third parties.

we examined:

- ◎ how financial **open data** can be **harnessed** to extract useful inferences regarding specific CSR aspects.
- ◎ two datasets and tested four hypotheses.

future work:

- ◎ could focus on researching more **complex** hypotheses.
- 

Thanks!

Any Questions?

You can find me at:
vasgat@iti.gr



References

- Wanderley, L.S.O., Lucian, R., Farache, F., de Sousa Filho, J.M.: Csr information disclosure on the web: a context-based approach analysing the influence of country of origin and industry sector. *Journal of Business Ethics* 82(2), 369-378 (2008)
- Cai, Y., Jo, H., Pan, C.: Doing well while doing bad? csr in controversial industry sectors. *Journal of Business Ethics* 108(4), 467-480 (2012)
- Bauman, C.W., Skitka, L.J.: Corporate social responsibility as a source of employee satisfaction. *Research in Organizational Behavior* 32, 63-86 (2012)
- Namkung, Y., Jang, S.: Are consumers willing to pay more for green practices at restaurants? *Journal of Hospitality & Tourism Research* 41(3), 329-356 (2017)
- Parguel, B., Benoit-Moreau, F., Larceneux, F.: How sustainability ratings might deter greenwashing: A closer look at ethical corporate communication. *Journal of business ethics* 102(1), 15 (2011)
- Bragdon, J., Marlin, J.: Is pollution profitable. *Risk management* 19(4), 9-18 (1972)
- Marcus, A.: The deterrent to dubious corporate behavior: Profitability, probability and safety recalls. *Strategic management journal* 10(3), 233-250 (1989)
- Margolis, J.D., Elfenbein, H.A., Walsh, J.P.: Does it pay to be good? a meta-analysis and redirection of research on the relationship between corporate social and financial performance. *Ann Arbor* 1001, 48109-1234 (2007)
- Cheng, B., Ioannou, I., Serafeim, G.: Corporate social responsibility and access to finance. *Strategic Management Journal* 35(1), 1-23 (2014)
- Eckbo, B.E., Nygaard, K., Thorburn, K.S.: Does gender-balancing the board reduce firm value? (2016)
- Dyreng, S.D., Lindsey, B.P., Thornock, J.R.: Exploring the role delaware plays as a domestic tax haven. *Journal of Financial Economics* 108(3), 751-772 (2013)
- Wayne, L.: How delaware thrives as a corporate tax haven. *The New York Times* 30 (2012)