



UNIVERSITY OF  
CAMBRIDGE  
Department of Sociology



This project is co-funded by the  
Seventh Framework Programme for  
Research and Technological  
Development of the European Union



CORRUPTION  
RESEARCH CENTER  
BUDAPEST

# Exploring government contracting data to hold governments accountable in the Big Data Era

Mihály Fazekas

University of Cambridge, [mf436@cam.ac.uk](mailto:mf436@cam.ac.uk)

Data for Policy Conference, Cambridge, UK, 17/6/2015

# The challenge of public procurement

9.3 trillion USD per year\*

20-25% lost to corruption\*\*

Should we be worried?

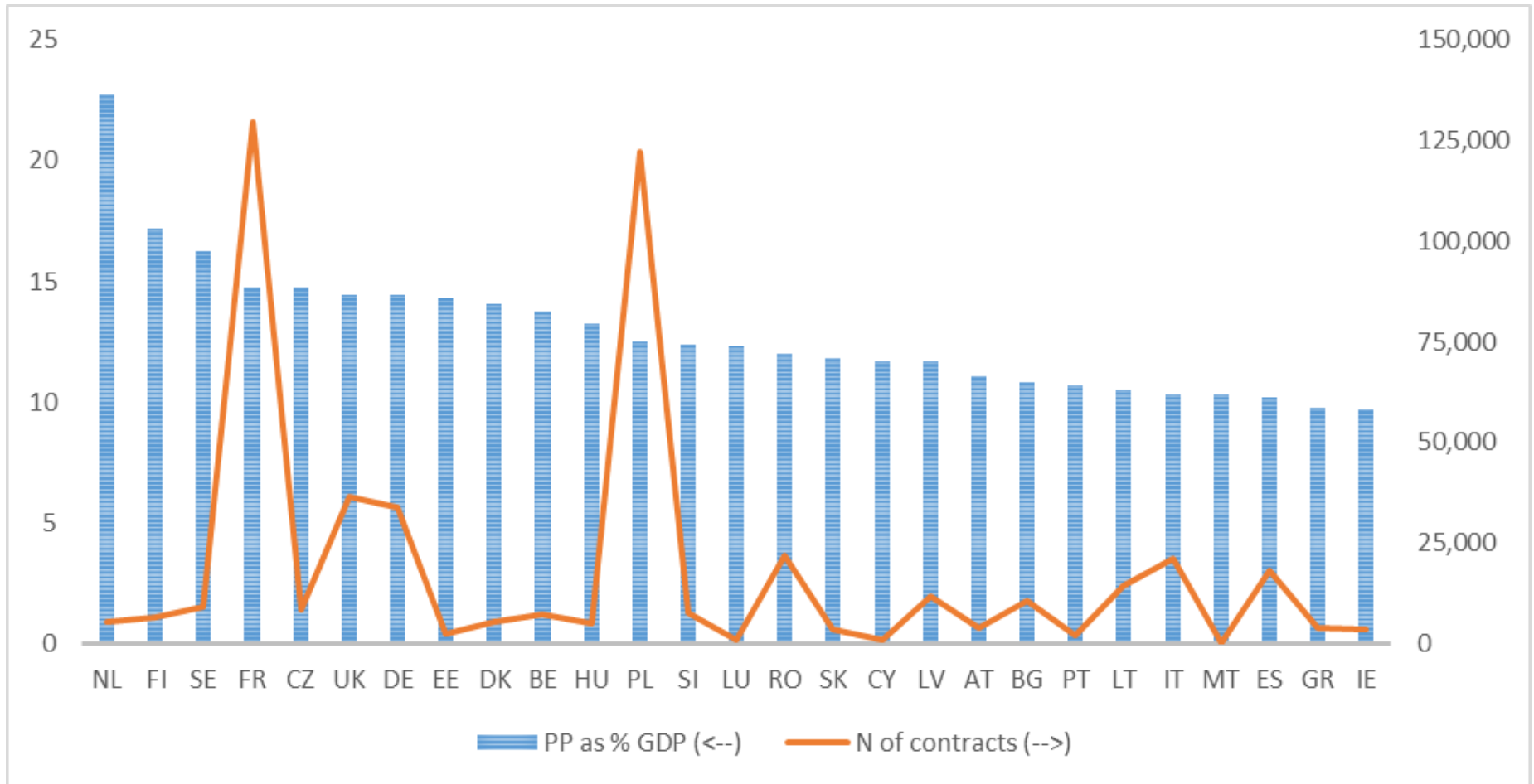
\* <http://www.open-contracting.org/data-standard-announcement>

\*\* OECD (2013), Implementing the OECD Principles for Integrity in Public Procurement. Progress since 2008. OECD, Paris

# Overview

- Big Data in government contracting?
- State of data and data collection innovations
- Indicators
- 3 examples
- How you can get involved?

# Why we need big data in PP?



2012, EU, TED

2025. 11. 26.

# State of data I

# DISASTER

One  
example  
disaster



Browser tabs: Istvan J, Facebook, Orbán, Beérke, wg11.p, cw910, teol.lut, volunt, Új lap, SZ SZTAKI, prograi, Közbes.

Address bar: kozbeszerzes.hu/adatbazis/mutat/hirdetmeny/portal\_33172\_2010/

Navigation: KÖZBESZERZÉSI HATÓSÁG, ENGLISH, DEUTSCH, HATÓSÁG, JOGORVOSLAT, JOGI HÁTTÉR, TEVÉKENYSÉGEK, KAPCSOLATOK

Buttons: Belépés / Regisztráció, Jelszó-emeleketető

Cím: Cím

|                                                 |                                                                                       |
|-------------------------------------------------|---------------------------------------------------------------------------------------|
| Közbeszerzési Értesítő száma:                   | 2010/142                                                                              |
| Beszermzés tárgya:                              | Árubeszerzés                                                                          |
| Hirdetmény tárgya:                              | tájékoztató az eljárás eredményéről (1-es minta) KÉ                                   |
| Eljárás fajtája:                                | Nyílt                                                                                 |
| Közzététel dátuma:                              | 2010.12.01.                                                                           |
| Iktatószám:                                     | 33172/2010                                                                            |
| CPV Kód:                                        | 33696500-0                                                                            |
| Ajánlatkérő:                                    | Vas Megyei Markusovszky Kórház Nonprofit Zrt.                                         |
| Teljesítés helye:                               | Vas Megyei Markusovszky Kórház Nonprofit Zrt. 9700 Szombathely, Markusovszky L. u. 5. |
| Ajánlattételi/részvételi jelentkezési határidő: |                                                                                       |
| Nyertes ajánlattevő:                            | Csertex Kft.; Diagnosticum Zrt.; Laborexper Kutató, Szolgáltató és Kereskedelmi Kft.  |
| Ajánlatkérő típusa:                             | Egyéb                                                                                 |
| Ajánlatkérő fő tevékenységi köre:               |                                                                                       |
| Letöltés:                                       | <a href="#">Hirdetmény letöltése PDF formátumban</a>                                  |

1. melléklet az 5/2009. (III.31.) IRM rendelethez  
KÖZBESZERZÉSI ÉRTESÍTŐ  
A Közbeszerzések Tanácsának Hivatalos Lapja 1024 Budapest, Margit krt. 85.  
Fax: 06 1 336 7751; 06 1 336 7757  
E-mail: hirdetmeny@kozbeszerzesek-tanacsa.hu On-line értesítés: <http://www.kozbeszerzes.hu>  
tájékoztató  
az eljárás eredményéről  
A Közbeszerzések Tanácsa (Szerkesztőbizottsága) tölti ki  
A hirdetmény kézhezvételének dátuma \_\_\_\_\_  
KÉ nyilvántartási szám \_\_\_\_\_  
I. SZAKASZ: AJÁNLATKÉRŐ  
1.1) Név, cím és kapcsolattartási pont(ok)  
Hivatalos név: Vas Megyei Markusovszky Kórház Nonprofit Zrt.  
Postai cím: Markusovszky L. u. 5.  
Város/Község: Szombathely  
Postai irányítószám: 9700  
Ország: Magyarország  
Kapcsolattartási pont(ok): Dr. Lakner László főigazgató főorvos  
Címzett: Dr. Lakner László főigazgató főorvos  
Telefon: 06-94-311-542  
E-mail: titkarsag@markusovszky.hu  
Fax: 06-94-337-872

Downloads: wg4.pdf, g8.pdf, szakiskola\_2014\_tanul....zip

Összes letöltés megjelenítése...

Taskbar: Windows, Explorer, Chrome, Word, Excel, PowerPoint, HU, 11:20 AM 3/20/2014

2025. 11. 26.

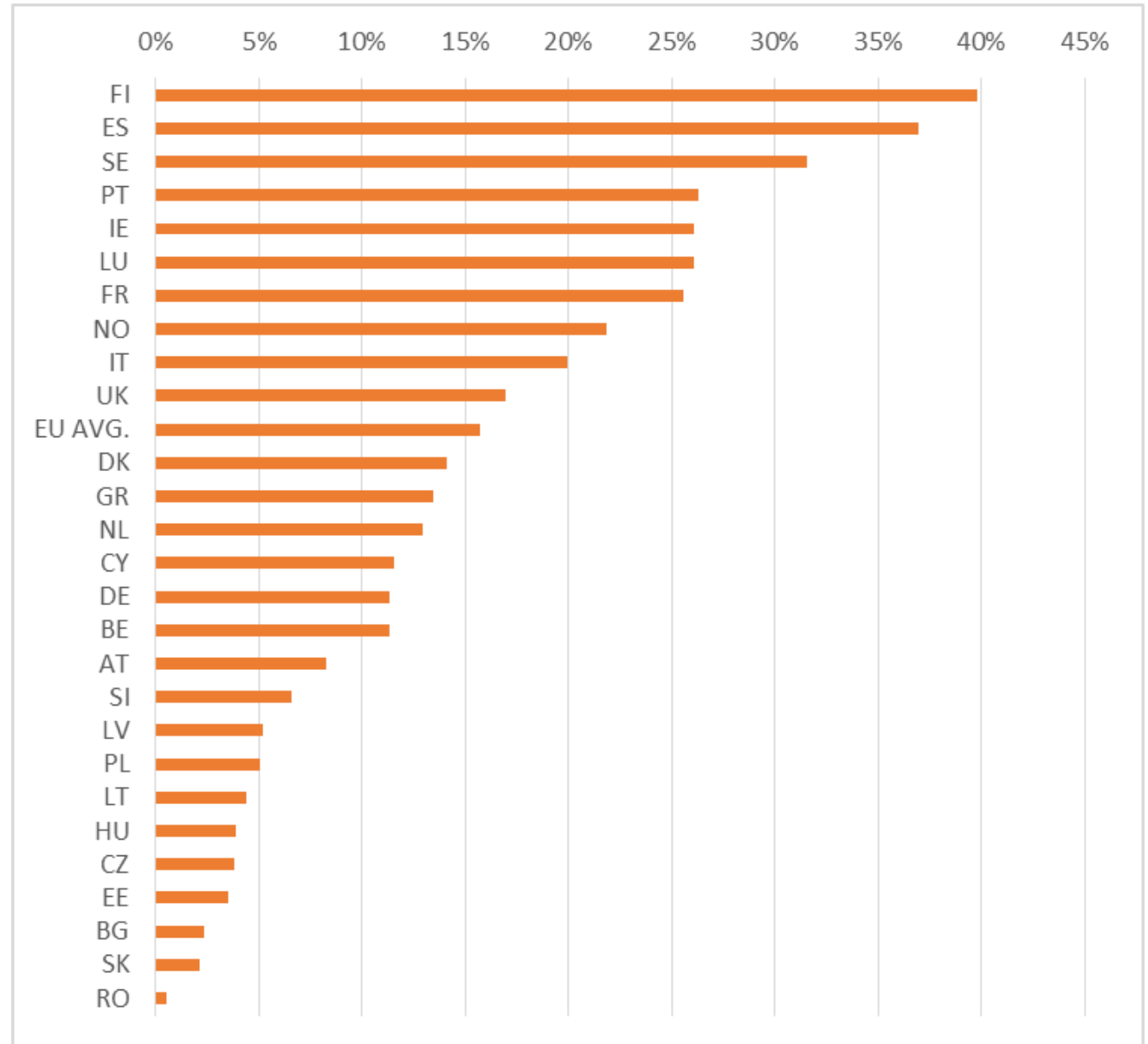
# Average admin error

TED:2009-13

13 mandatory items:

- Organisation name, address
- Contract values
- Subcontract
- Dates

2025. 11. 26.



# State of data II

- Patchwork of data sources
  - Formats: xml,html,pdf, etc
  - Contents
  - Quality
  - Linking
- Administrative data provision is not ready for Big Data
  - Public initiatives: OCDS
  - Civil society/research community: DIGIWHIST



# DIGIWHIST

The Digital Whistleblower. Fiscal Transparency, Risk Assessment and Impact of Good Governance Policies Assessed

- Goals
  - Advancing anticorruption, transparency, and spending efficiency in public procurement
  - Open data and indicators for 35 European countries: EU, EEA, Caucasus
  - Enabling losers of corruption to mobilize
  - Helping audit bodies fighting corruption, fraud, and collusion
- Scope
  - March 2015 – February 2018
  - 3 million eur (Horizon2020 funding)
  - Consortium: Cambridge, Hertie, Open Knowledge Foundation DE, CRCB, Datlab, Transcrime

# DIGIWHIST: Data

- Transparency and procurement legislation
- Procurement data
- Company information
- Public organisation information
- Asset declarations

# DIGIWHIST: Indicators+utilization

- Indicators:
  - Corruption/favouritism
  - Transparency
  - Administrative quality
- Utilization
  - Web portals, mobile apps
  - Whistleblower reporting: Big Data+user info
  - Risk assessment software for public servants

# Indicator building

- New approach to corruption in PP
  - harnessing BIG DATA,
  - built on thorough understanding of context, and
  - ,open-ended'
- Indicator characteristics:
  - Specific
  - Real-time
  - 'Objective'/hard
  - Micro-level
  - Aggregatable + comparative



# 3 hands-on examples

For

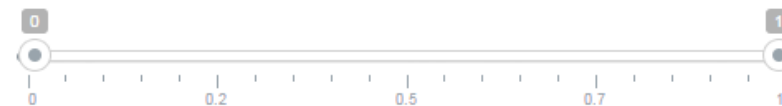
1. The activist
2. The policy maker
3. The economist

# Big Data for accountability

- [www.tendertracking.eu](http://www.tendertracking.eu)
- Goal: Holding governments accountable
- All regulated contract since 2005 in Hungary
- Structured micro data+aggregates
- Indicators making sense of Big Data
- Target groups:
  - Citizens, journalists
  - Policy makers
  - Researchers

**Winner tenderer**

- ☒ Partial or full winner name
- ☐ Groups of winners

**Winner's name****Tender caller's name****Tender title****Product/Service****Type of procurement****CRI****From CRI****To CRI****CRI****Date****From Date****To Date**

Hint: Double click on the year to navigate easily between years

**€ Price****From Price****To Price**

Tenders filtered out from the results - Last updated at Sat 13 Jun 2015

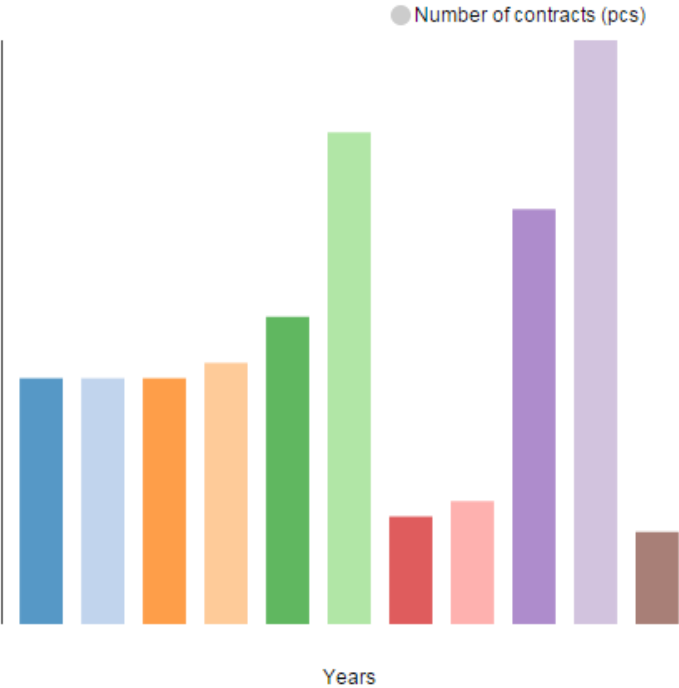
Tenders with missing data - Last updated at Sat 13 Jun 2015

Tájékoztató a keresőfelületen meg nem jelenített szerződésekről.

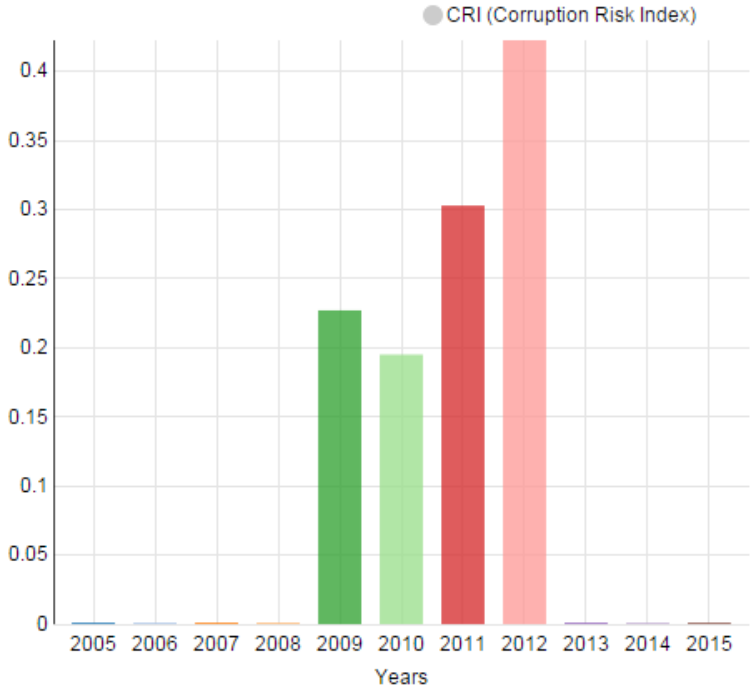
Search   
203 results



Number of contracts



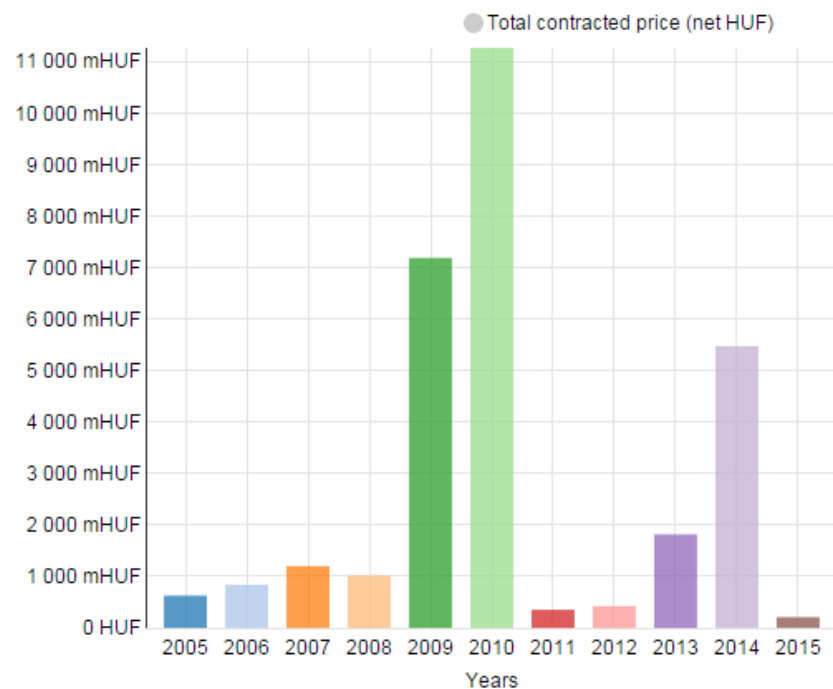
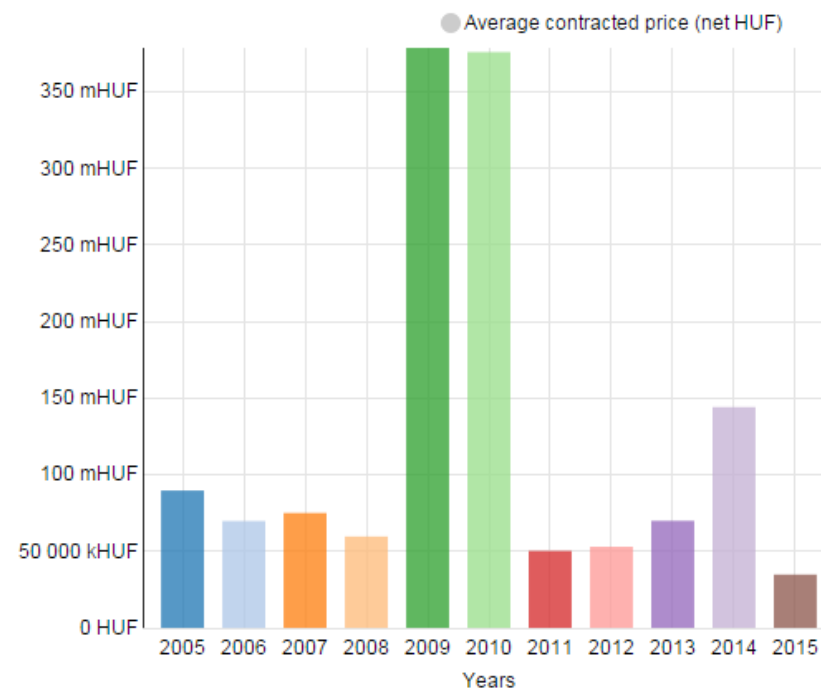
CRI



Copyright: CRCB, 2014

The website and the database it is built upon has been financed by [EU FP7 Framework Programme](#), [ANTICORRP Programme](#); tender #290529; beneficiary: [Corvinus University of Budapest](#) ). This project couldn't have been done without the invaluable help of the volunteers of [CRCB](#)



**Total contracted price****Average contracted price**

Map



● Tender callers' offices    ● Winning tenderer companies' offices

Copyright: CRCB, 2014

The website and the database it is built upon has been financed by [EU FP7 Framework Programme, ANTICORRP Programme; tender #290529](#); beneficiary: [Corvinus University of Budapest](#) ). This project couldn't have been done without the invaluable help of the volunteers



CORRUPTION  
RESEARCH CENTER  
BUDAPEST

## Descriptive statistics

|              | Count | Minumum       | Maximum           | Average         | Deviation       | Median         | Mode          |
|--------------|-------|---------------|-------------------|-----------------|-----------------|----------------|---------------|
| <b>Price</b> | 186   | 1 168 640 HUF | 2 948 000 000 HUF | 163 629 274 HUF | 460 357 839 HUF | 30 425 000 HUF | 2 500 000 HUF |
| <b>CRI</b>   | 67    | 0.06          | 0.62              | 0.24            | 0.14            | 0.22           | 0.07          |
| <b>N.B.*</b> | 141   | 1.00          | 11.00             | 3.04            | 1.62            | 3.00           | 3.00          |

\*Number of bidders

203 results (1. page out of 9)

| Subject                                                                                                                                                                                                        | Winner                                                                                                  | Tender caller                               | Product/Service                         | Date       | Price          |                   |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|---------------------------------------------|-----------------------------------------|------------|----------------|-------------------|
| tájékoztató az eljárás eredményéről-iskolainformációs szoftver                                                                                                                                                 | FOK Fővárosi Oktatástechnológiai Központ Szolgáltató Közhasznú Nonprofit Korlátolt Felelősségű Társaság | Nyíregyháza Megyei Jogú Város Önkormányzata | Database software development services. | 2011-03-11 | 11 107,00 eFt  | <a href="#">→</a> |
| tájékoztató az eljárás eredményéről-könyvvizsgálói feladatok ellátása                                                                                                                                          | VERSATILE AUDIT Könyvvizsgáló és Adószakértő Korlátolt Felelősségű Társaság                             | Nyíregyháza Megyei Jogú Város Önkormányzata | Accounting and auditing services.       | 2011-02-18 | 14 100,00 eFt  | <a href="#">→</a> |
| tájékoztató az eljárás eredményéről- 500 millió forint összegű folyószámlahitelkeret biztosítása                                                                                                               | OTP Bank Nyrt.                                                                                          | Nyíregyháza Megyei Jogú Város Önkormányzata |                                         | 2011-01-17 | 40 000,00 eFt  | <a href="#">→</a> |
| „A Petőfi Sándor Általános Iskola infrastrukturális fejlesztése, bővítése” című projekt keretén belül a Nyíregyháza, Alma u. 70., meglévő iskola épületének bővítése, felújítása, akadálymentesítése, valamint | RÓZSAÉP KFT                                                                                             | Nyíregyháza Megyei Jogú Város Önkormányzata | Construction work.                      | 2011-07-20 | 137 607,73 eFt | <a href="#">→</a> |

Tájékoztató az eljárás eredményéről -

Copyright: CRCB, 2014

The website and the database it is built upon has been financed by EU FP7 Framework Programme, ANTICORRP Programme; tender #290529; beneficiary: Corvinus University of Budapest ). This project couldn't have been done without the invaluable help of the volunteers of CRCB



CORRUPTION RESEARCH CENTER BUDAPEST

tájékoztató az eljárás eredményéről-könyvvizsgálói feladatok ellátása

Information

|                     |                                                                                                                                                     |
|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| Tender caller:      | Nyíregyháza Megyei Jogú Város Önkormányzata                                                                                                         |
| Winner:             | VERSATILE AUDIT Könyvvizsgáló és Adószakértő Korlátolt Felelősségű Társaság                                                                         |
| Product/Service:    | Számviteli és könyvvizsgálói szolgáltatások.                                                                                                        |
| Date of Publishing: | 2011-02-18                                                                                                                                          |
| Contracted Price:   | 14 100 000 HUF                                                                                                                                      |
| Number of bidders:  | 3 pcs                                                                                                                                               |
| Show procurement:   | <a href="http://kozbeszerzes.hu/adatbazis/mutat/hirdetmeny/portal_0846_2011">http://kozbeszerzes.hu/adatbazis/mutat/hirdetmeny/portal_0846_2011</a> |

CRI



|  |           |   |           |
|--|-----------|---|-----------|
|  | 0         | - | 0.1274738 |
|  | 0.1274738 | - | 0.2086671 |
|  | 0.2086671 | - | 0.272782  |
|  | 0.272782  | - | 0.3828584 |
|  | 0.3828584 | - | 1         |

Map

Pictures

Product/Service: Számviteli és könyvvizsgáló szolgáltatások

tendertracking.eu

Date of Publishing: 2011-02-18

Top results

Search

About us



Contracted Price: 14 100 000 HUF

Number of bidders: 3 pcs

Show procurement: [http://kozbeszerzes.hu/adatbazis/mutat/hirdetmeny/portal\\_0846\\_2011](http://kozbeszerzes.hu/adatbazis/mutat/hirdetmeny/portal_0846_2011)

|  |           |   |           |
|--|-----------|---|-----------|
|  | 0         | - | 0.1274738 |
|  | 0.1274738 | - | 0.2086671 |
|  | 0.2086671 | - | 0.272782  |
|  | 0.272782  | - | 0.3828584 |
|  | 0.3828584 | - | 1         |



Map



Pictures

You must sign in to upload images!



Comments

You must sign in to comment

Copyright: CRCB, 2014

The website and the database it is built upon has been financed by EU FP7 Framework Programme, ANTICORRP Programme; tender #290529; beneficiary: Corvinus University of Budapest ). This project couldn't have been done without the invaluable help of the volunteers



CORRUPTION  
RESEARCH CENTER  
BUDAPEST

2025. 11. 26.



# Discussion points

Which features would you like to see on such a portal?



# Big Data for austerity

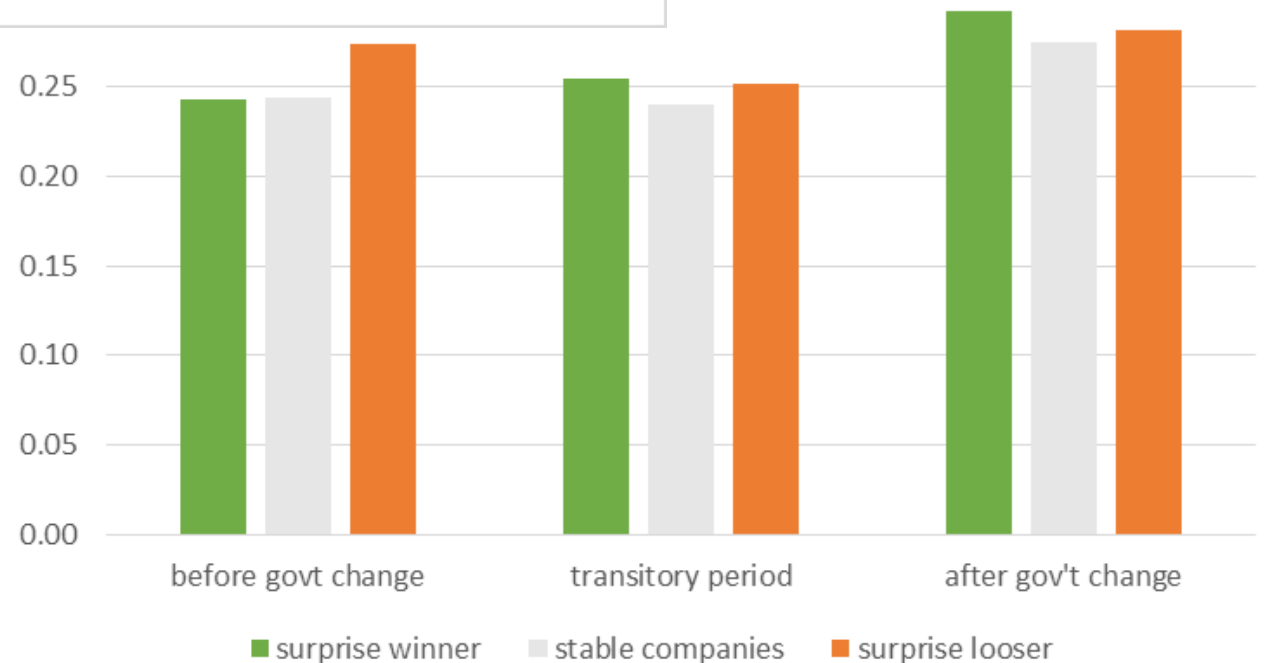
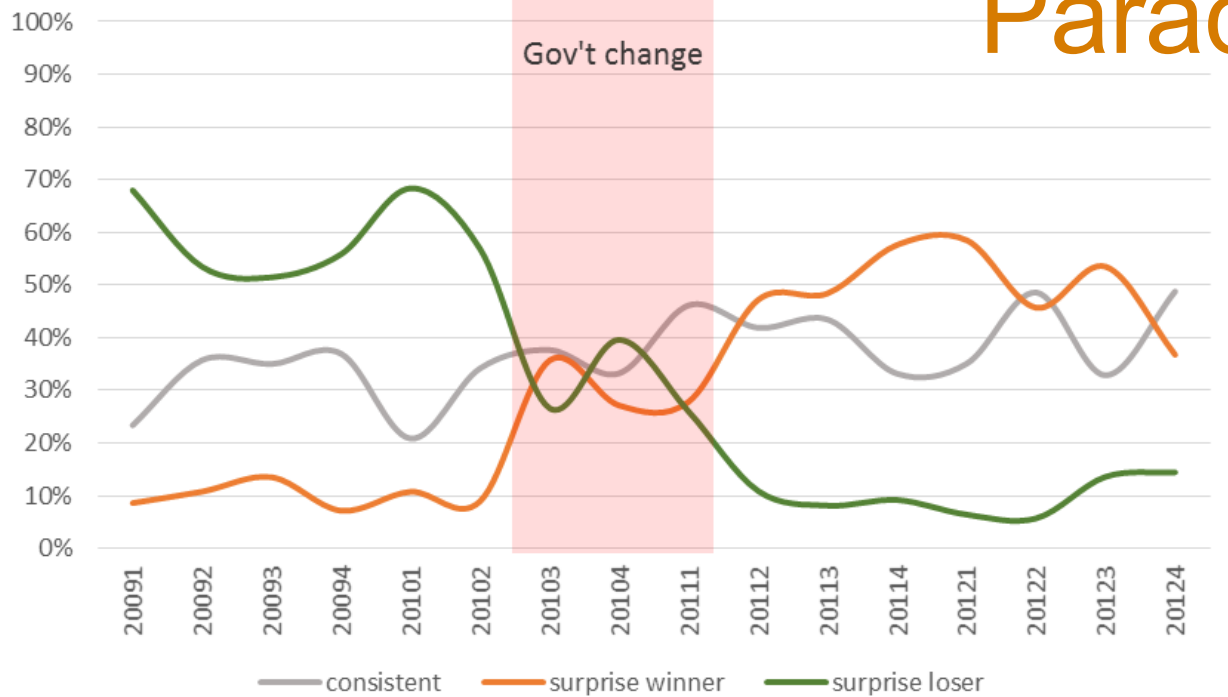
- <http://www.crcb.eu/?p=867>
- Goal: identifying companies with preferential treatment
- Data: public procurement+company data: UK
- Target group
  - Policy makers (those not benefitting of course😊)
  - Losing bidders
  - Journalists

# Political Influence Indicator (PII)

- political influence on companies' market success
- Change of government as a natural experiment
  - Company performance before-after gov't change
  - Evidence of favouritism: tendering red flags

# Paradigmatic case: Hungary

Government  
change + CRI  
2009-2012



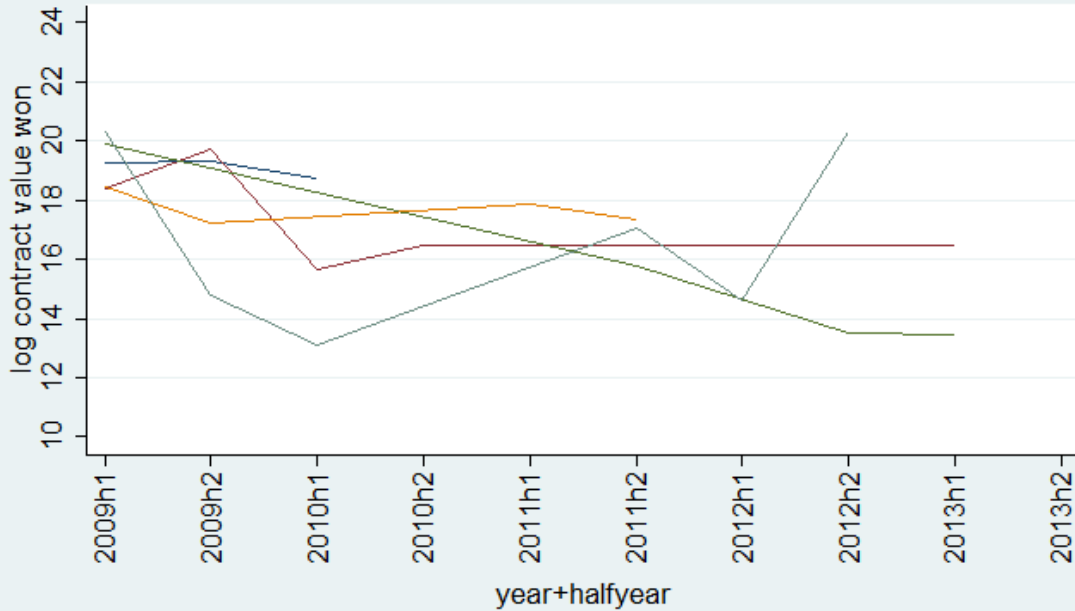
2025. 11. 26.

# How does the UK score?

- Central government only
- 2009-2010-2013
- Company panel data
- TED data
  - large contracts
  - Imperfect company identification
- Download link: <http://www.crcb.eu/?p=867>
- Let's explore the data together!

# SOME suspicion

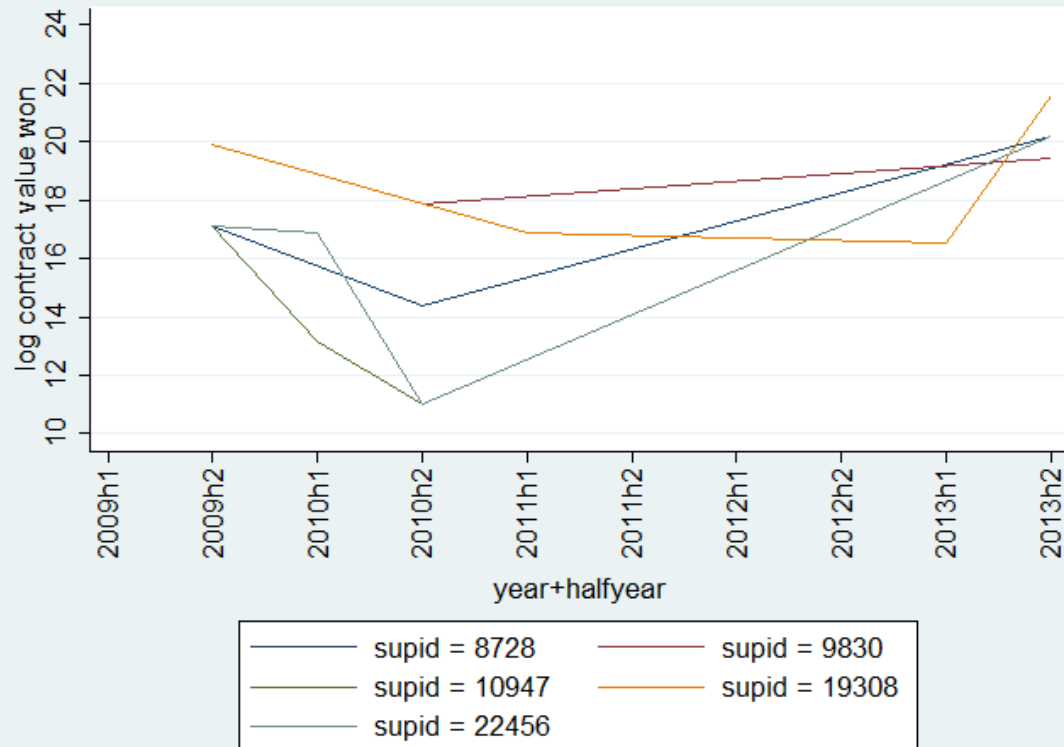
- ◀ 2009h1 largest 5
- ▼ 2013h2 largest 5



xtline sum\_ls\_org if best==1,  
ylabel(10(2)24) xlabel(#10,  
angle(vertical)) ytitle("log contract  
value won") overlay

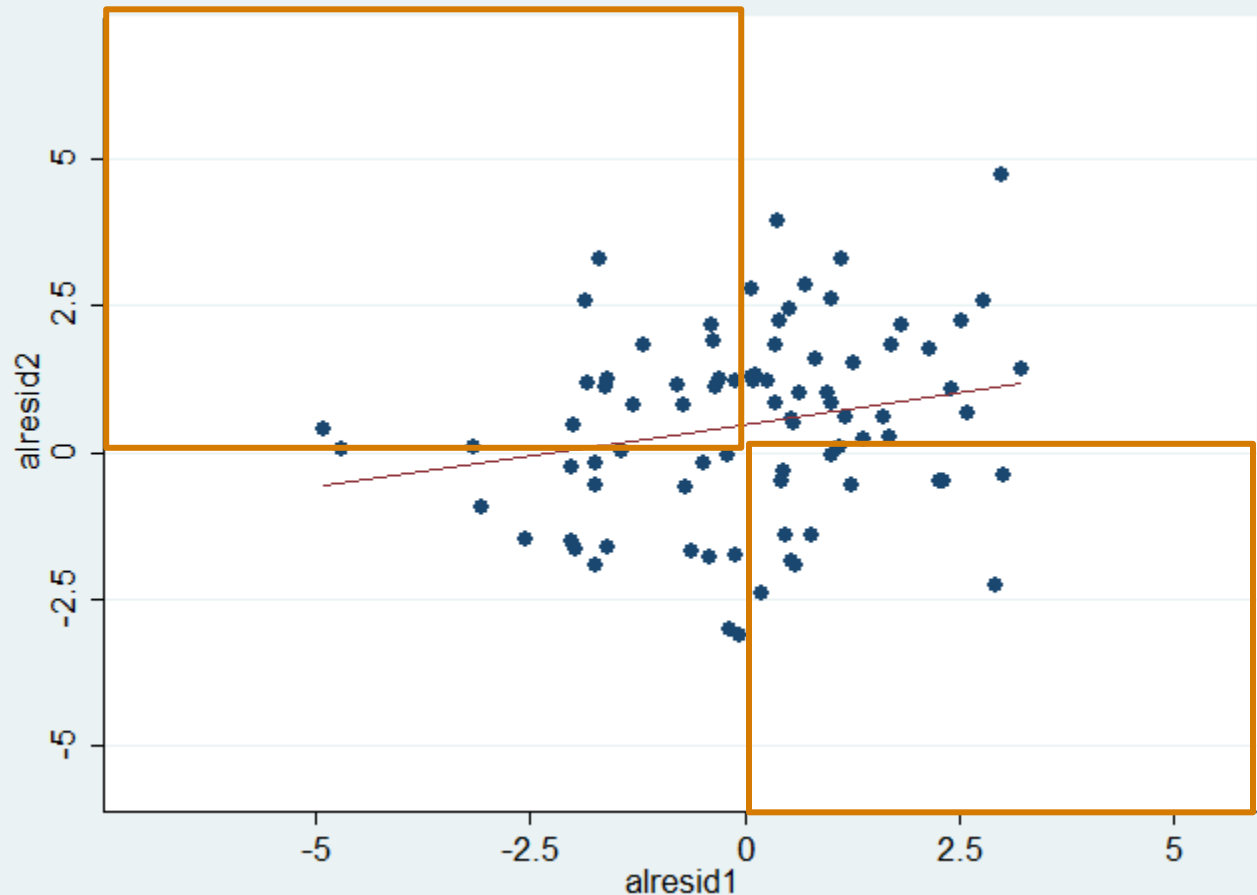
xtline sum\_ls\_org if best==2,  
ylabel(10(2)24) xlabel(#10,  
angle(vertical)) ytitle("log contract  
value won") overlay

2025. 11. 26.



# Suspicious error term patterns

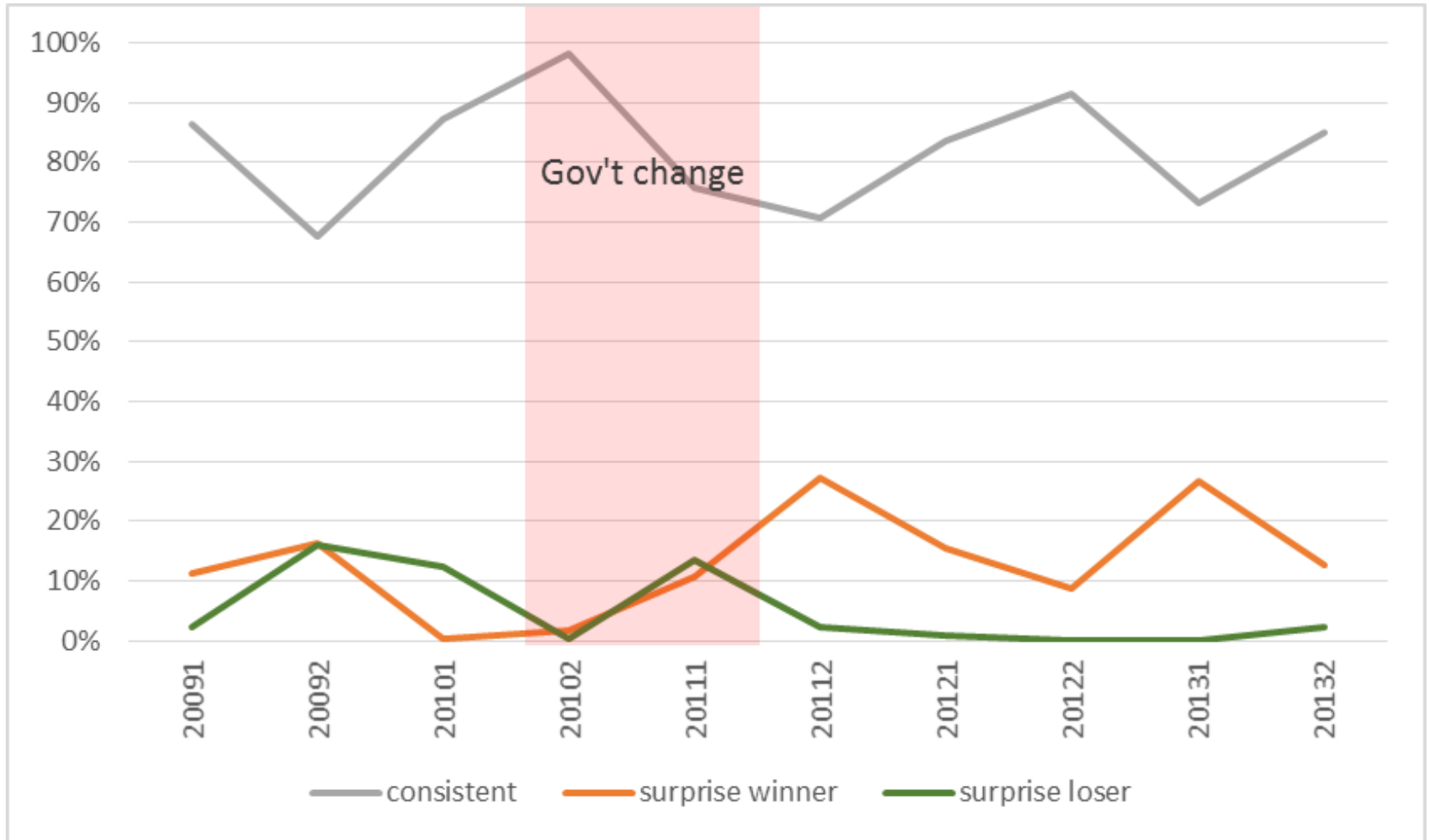
Systematic deviations from the regression line



```
twoway (scatter alresid2  
alresid1)(lfit alresid2  
alresid1), xlabel(-5(2.5)5)  
ylabel(-5(2.5)5) legend(off)  
ytile("alresid2")
```

2025. 11. 26.

# UK central admin: PII



# Discussion points

- Any comments on the methodology?
- How to use the results if nothing is apparently illegal?
- Someone wants to write an article about these companies😊?



# Big Data for competition

- Bidder data, Hungary, 2007&2009
- Goal: identifying suspicious bidding patterns
- Target group:
  - Competition authorities
  - Losing bidders

# Network data set-up

- Co-bidding network
  - Construction work for pipelines, communication and power lines, for highways, roads, airfields and railways
  - 3 regional and 1 national markets in 2007&2009
- Nods: bidding firms (winners and losers)
  - size=number of tenders won
  - colour=network position (green=cut-point)
- Edges: bidding on the same tender
  - width=number of times companies co-bid

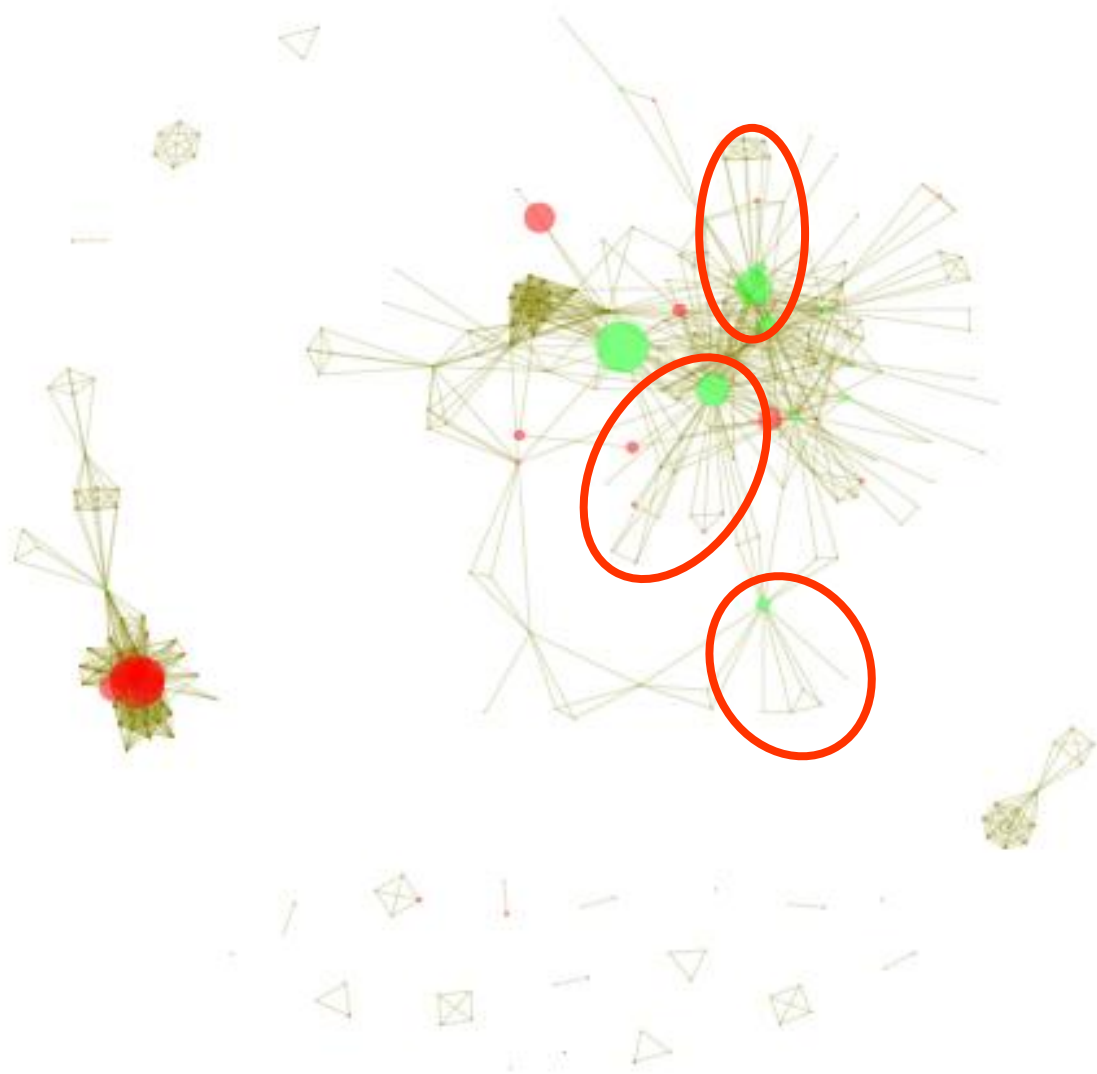
# Co-bidding patterns: benchmark

- 2007
- Dense networks
- Few cutpoints
- Cutpoints **don't benefit** from position



# Co-bidding patterns: suspicion

- 2009
- Dense networks
- Many cutpoints
- **Cutpoints seem to benefit from position**



# Discussion points

- Any comments on the methodology?
- How to mainstream such Big Data analysis in government work?
- How to bring suspicious cases to court?

# Getting involved

- Email me: [mf436@cam.ac.uk](mailto:mf436@cam.ac.uk)
- Visit: [digiwhist.eu](http://digiwhist.eu)
- As a User
  - Hunt for suspicious transactions
  - Download the data for analysis
- As a Contributor
  - Let us know what you need
  - Tell us about your story
  - Link your data

# Background material

SOON COMING UP: [digiwhist.eu](http://digiwhist.eu)

Fazekas, M. and Tóth, I. J. (2014). *From corruption to state capture: A new analytical framework with empirical applications from Hungary*. CRC-WP/2014:01, Budapest: Corruption Research Centre.

Czibik, Ágnes; Fazekas, Mihály; Tóth, Bence; and Tóth, István János (2014), *Toolkit for detecting collusive bidding in public procurement. With examples from Hungary*. Corruption Research Center Budapest, CRCB-WP/2014:02.

Fazekas, M., Chvalkovská, J., Skuhrovec, J., Tóth, I. J., & King, L. P. (2014). *Are EU funds a corruption risk? The impact of EU funds on grand corruption in Central and Eastern Europe*. In A. Mungiu-Pippidi (Ed.), *The Anticorruption Frontline. The ANTICORRP Project*, vol. 2. (pp. 68–89). Berlin: Barbara Budrich Publishers.

Fazekas, M., Tóth, I. J. (2014), *Three indicators of institutionalised grand corruption using administrative data*. U4 Anti-Corruption Resource Centre, Bergen, U4 Brief 2014:9.

Fazekas, M., Tóth, I. J., & King, L. P. (2013). *Anatomy of grand corruption: A composite corruption risk index based on objective data*. CRC-WP/2013:02, Budapest: Corruption Research Centre.

Fazekas, M., Tóth, I. J., & King, L. P. (2013). *Corruption manual for beginners: Inventory of elementary “corruption techniques” in public procurement using the case of Hungary*. CRC-WP/2013:01, Corruption Research Centre, Budapest.

Fazekas, M., Tóth, I. J., & King, L. P. (2013). *Hidden Depths. The Case of Hungary*. In A. Mungiu-Pippidi (Ed.), *Controlling Corruption in Europe vol. 1* (pp. 74–82). Berlin: Barbara Budrich Publishers.

# THANKS A LOT!