CORRUPTION RESEARCH CENTER BUDAPEST

ANTICORRP

UNIVERSITY OF CAMBRIDGE
Department of Sociology

Second-generation indicators of High-Level Corruption using Government Contracting Data: Examples from Eastern Europe

Mihály Fazekas* - István János Tóth+

*: University of Cambridge and Corruption Research Center Budapest, mihaly.fazekas@crcb.eu, corresponding author

+: Hungarian Academy of Sciences and Corruption Research Center Budapest

University of Oxford, Kellogg College, 20 June 2014, Oxford, UK

Corruption Research Center Budapest (CRCB)

- Interdisciplinary research team:
  - Ágnes Czibik (economist)
  - Mihály Fazekas (economist, political scientist)
  - Gyula Fóra (economics, IT)
  - Bence Tóth (economist)
  - István János Tóth (economist, sociologist)
- Experts
  - Zoltán Kelemen (lawyer)
  - Jenő Gyenese (IT)
  - Zoltán Nagy (compeittion ecomomics)
  - Sándor Rácz (computer programmer)
  - Zoltán Siposs (journalist)
  - Tamás Uhrin (IT)
- Aims
  - Help citizens to reach reliable data about public spending and government activities / effectiveness
  - quantitative analysis of corruption, effectiveness of government and state capacity

Corruption Research Center Budapest (CRCB)

- Major ongoing projects
  - Public procurement data collection and analysis of corruption risks across Europe
  - Public procurement cartels in Hungary
  - Quality of legislation in the EU
  - Transparency of local government in Hungary
  - Integrity in state owned enterprises in Hungary

Overview

- Measurement approach
- Definition of corruption
- Data
- Indicators
- CEE applications
- Options for UK applications

Starting point

- Available indicators are either biased or too idiosyncratic
  - Perception-based survey instruments measure PERCEPTIONS
  - Experience-based survey instruments suffer from conformity bias and lack of access
  - Audits and case studies lack scope and representativeness

→ Need for new indicators

The CRCB measurement approach

- 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

Why public procurement?

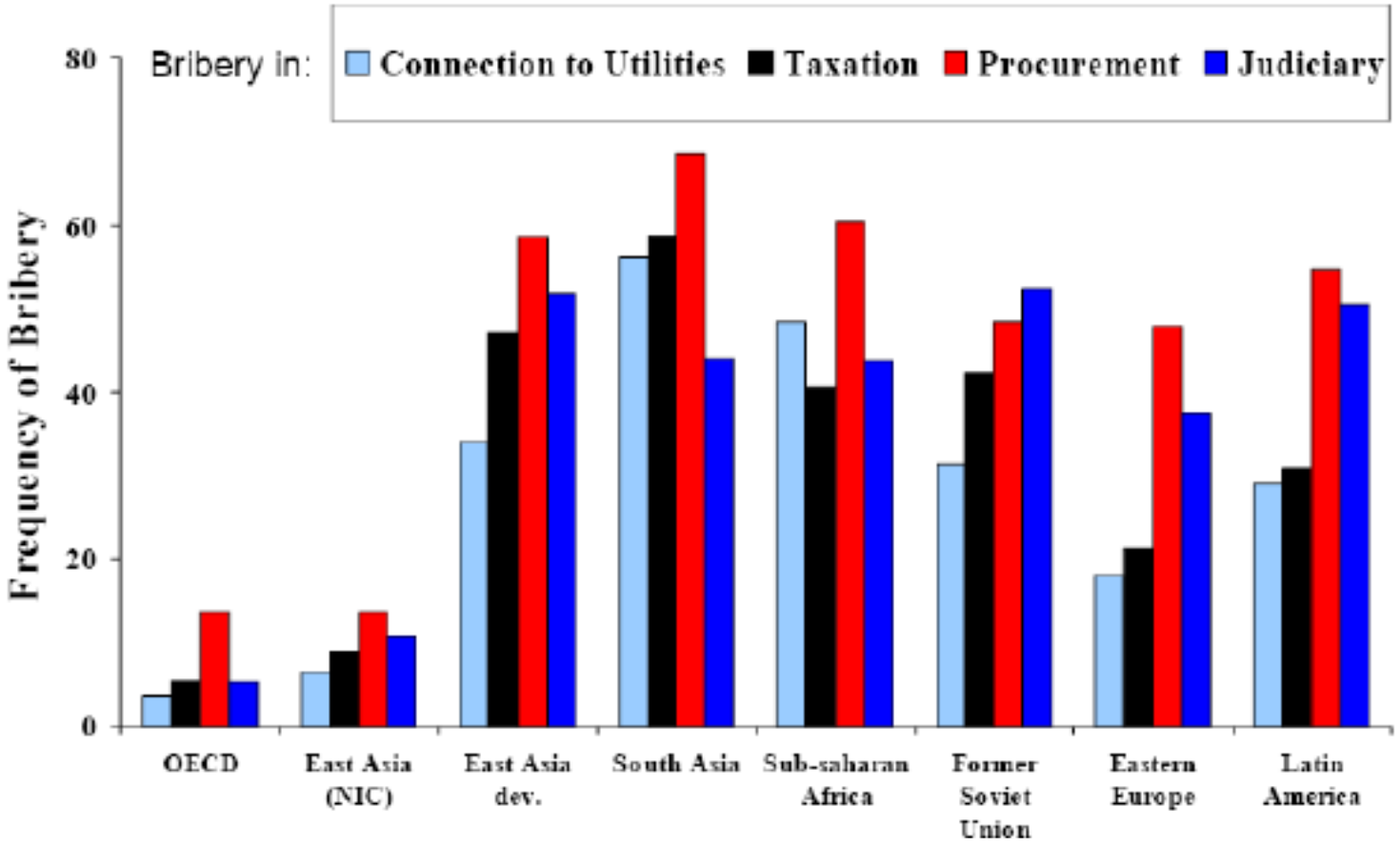
1. A lot of money involved

2. Crucial role in development (e.g. capital accumulation)

3. Indicates the broader quality of institutions

Why public procurement?

4. Very corrupt



CORRUPTION RESEARCH CENTER BUDAPEST

Definition of instutionalised grand corruption

- Specific definition (just like measurement)
- Institutionalised grand corruption in public procurement

> *institutionalised grand corruption in public procurement refers to the regular particularistic allocation and performance of public procurement contracts by bending universalistic rules and principles of good public procurement in order to benefit a group of individuals while denying access to all others.*

Definition in detail

- What it is NOT:
  - Not necessarily bribery
  - Not only „abuse of public office for private gain”
- What it IS:
  - Corruption=particularism and restricted access
  - Institutionalised=recurrent, stable, systemic
  - Grand=high-level politics and business
- Sources:

  - Mungiu-Pippidi, A. (2006). Corruption: Diagnosis and Treatment. *Journal of Democracy*, *17*(3), 86–99.
  - Rothstein, B., & Teorell, J. (2008). What Is Quality of Government? A Theory of Impartial Government Institutions. *Governance*, *21*(2), 165–190.
  - North, D. C., Wallis, J. J., & Weingast, B. R. (2009). *Violence and Social Orders. A Conceptual Framework for Interpreting Recorded Human History*. Cambridge, UK: Cambridge University Press.
  - Kaufmann, D., & Vicente, P. C. (2005). *Legal Corruption*. World Bank
  - Lambsdorff, J. G. (2007). The Institutional Economics of Corruption and Reform. Theory, Evidence and Policy, Cambridge University Press, Cambridge

The CRCB data template

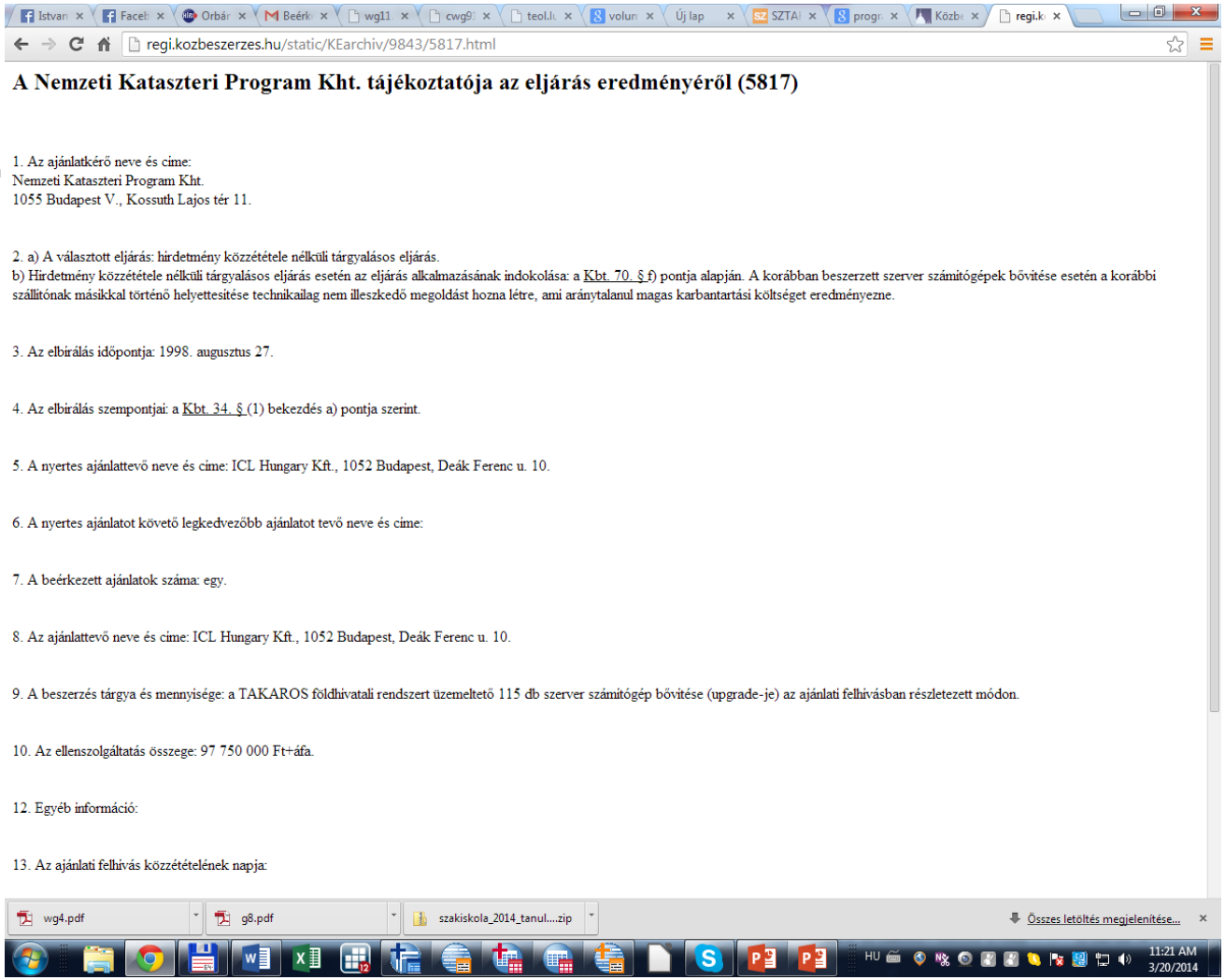
- Public procurement data
- Company financial and registry data
- Company ownership and management data
- Political officeholder data
- Treasury accounts of public organisations

Feasibility across the globe

- Transition economies:
  - HU, CZ, SK: already done
  - Romania, Croatia, Slovenia: ongoing work
- Developed/emerging economies
  - Italy: ongoing work
  - EU, US
  - Russia, Chile, Brazil
- Developing countries
  - World Bank data
  - Development agencies' procurement announcements: e.g. http://www.devbusiness.com/Default.aspx
  - National portals: Georgia: http://tendermonitor.ge/en

Data extraction & database building

From this…



regi.kozbeszerzes.hu/static/KEarchiv/9843/5817.html

A Nemzeti Kataszteri Program Kht. tájékoztatója az eljárás eredményéről (5817)

1. Az ajánlatkérő neve és címe:
Nemzeti Kataszteri Program Kht.
1055 Budapest V., Kossuth Lajos tér 11.

2. a) A választott eljárás: hirdetmény közzététele nélküli tárgyalásos eljárás.
b) Hirdetmény közzététele nélküli tárgyalásos eljárás esetén az eljárás alkalmazásának indokolása: a Kbt. 70. § f) pontja alapján. A korábban beszerzett szerver számítógépek bővítése esetén a korábbi szállítónak másikkal történő helyettesítése technikailag nem illeszkedő megoldást hozna létre, ami aránytalanul magas karbantartási költséget eredményezne.

3. Az elbírálás időpontja: 1998. augusztus 27.

4. Az elbírálás szempontjai: a Kbt. 34. § (1) bekezdés a) pontja szerint.

5. A nyertes ajánlattevő neve és címe: ICL Hungary Kft., 1052 Budapest, Deák Ferenc u. 10.

6. A nyertes ajánlatot követő legkedvezőbb ajánlatot tevő neve és címe:

7. A beérkezett ajánlatok száma: egy.

8. Az ajánlattevő neve és címe: ICL Hungary Kft., 1052 Budapest, Deák Ferenc u. 10.

9. A beszerzés tárgya és mennyisége: a TAKAROS földhivatali rendszert üzemeltető 115 db szerver számítógép bővítése (upgrade-je) az ajánlati felhívásban részletezett módon.

10. Az ellenszolgáltatás összege: 97 750 000 Ft+áfa.

12. Egyéb információ:

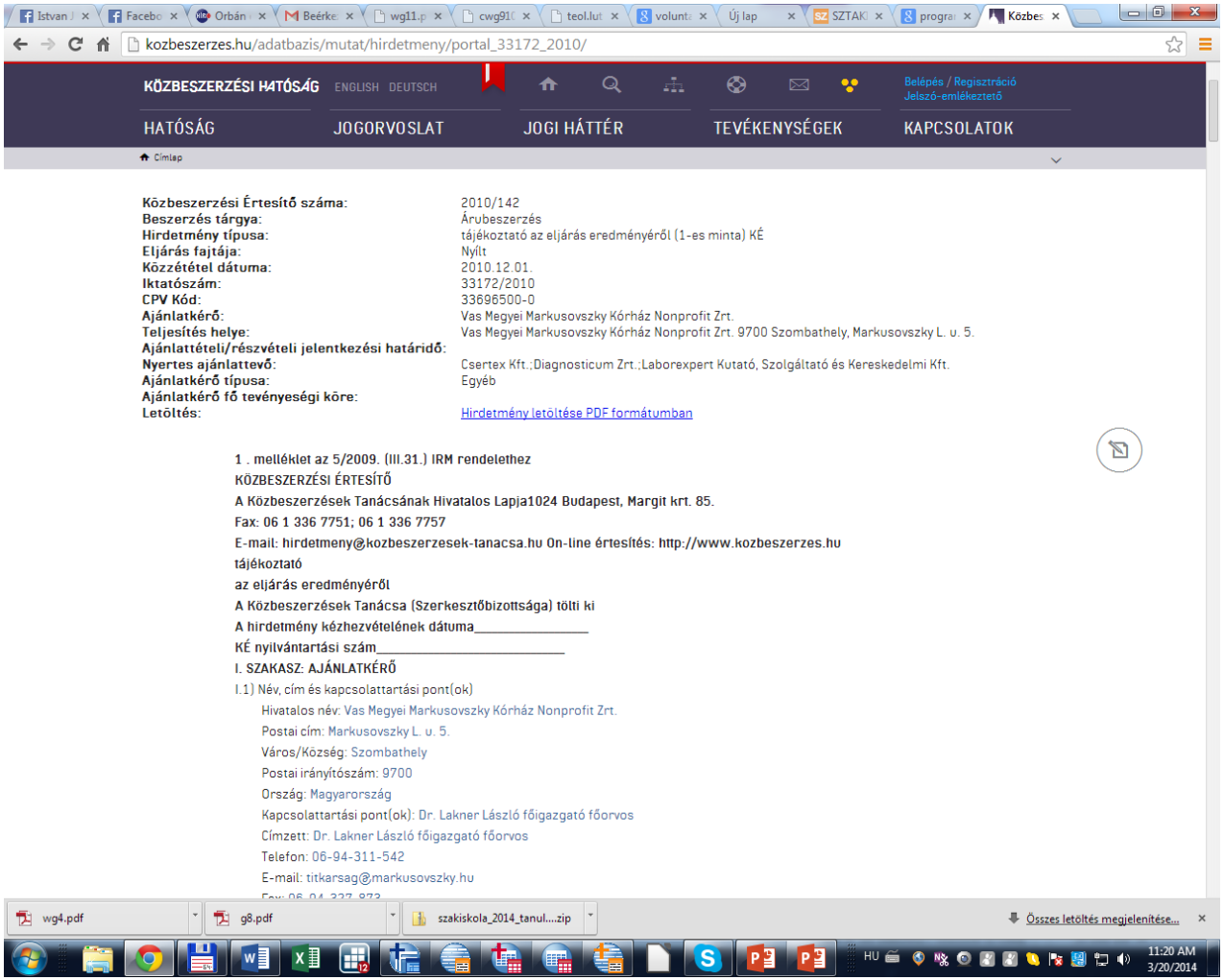
13. Az ajánlati felhívás közzétételének napja:

wg4.pdf g8.pdf szakiskola_2014_tanul....zip Összes letöltés megjelenítése...

11:21 AM 3/20/2014

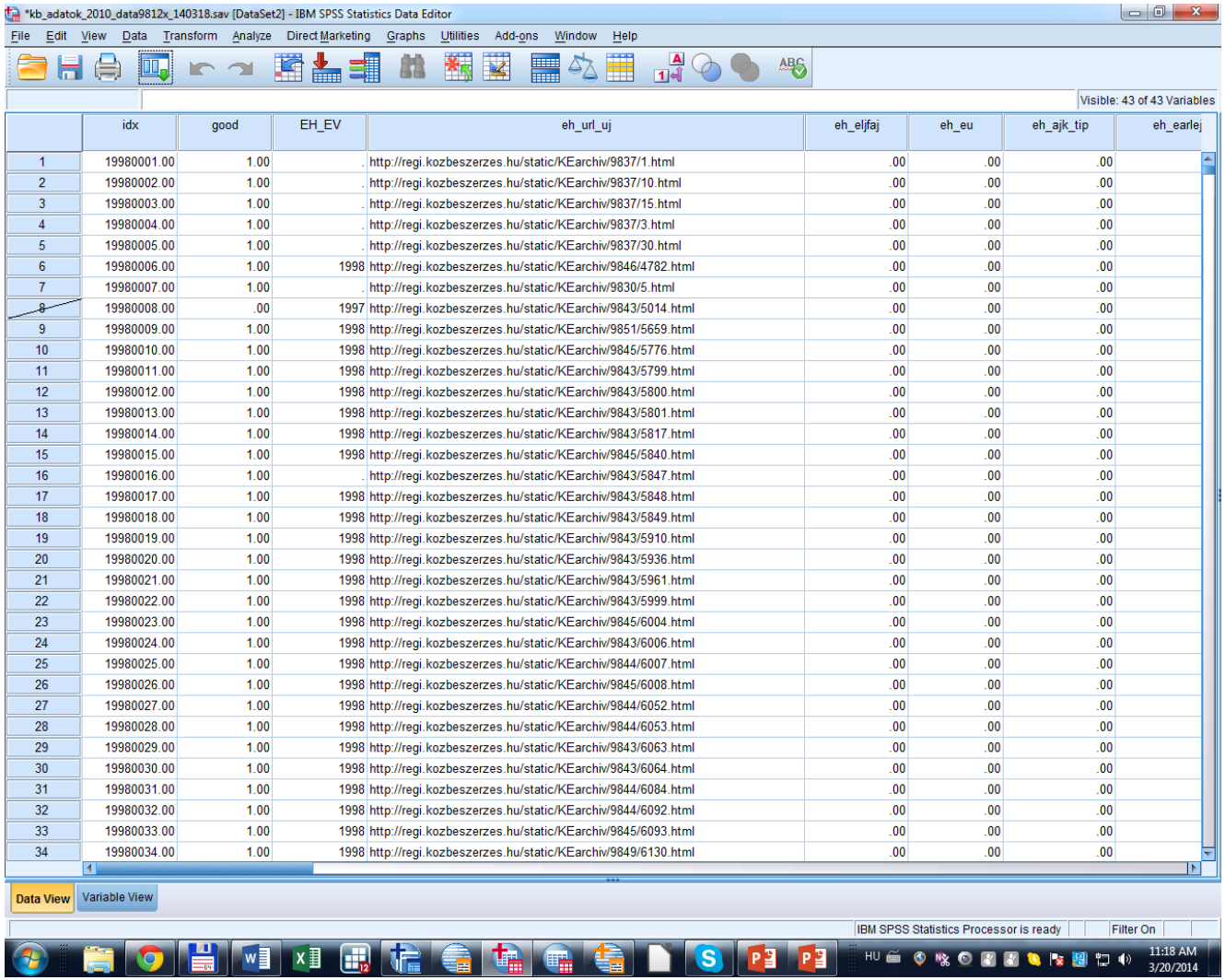
Data extraction & database building

…or from this…



Data extraction & database building

…to this



	idx	good	EH_EV	eh_url_uj	eh_eljfaj	eh_eu	eh_ajk_tip	eh_earlej
1	19980001.00	1.00	.	http://regi.kozbeszerzes.hu/static/KEarchiv/9837/1.html	.00	.00	.00	
2	19980002.00	1.00	.	http://regi.kozbeszerzes.hu/static/KEarchiv/9837/10.html	.00	.00	.00	
3	19980003.00	1.00	.	http://regi.kozbeszerzes.hu/static/KEarchiv/9837/15.html	.00	.00	.00	
4	19980004.00	1.00	.	http://regi.kozbeszerzes.hu/static/KEarchiv/9837/3.html	.00	.00	.00	
5	19980005.00	1.00	.	http://regi.kozbeszerzes.hu/static/KEarchiv/9837/30.html	.00	.00	.00	
6	19980006.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9846/4782.html	.00	.00	.00	
7	19980007.00	1.00	.	http://regi.kozbeszerzes.hu/static/KEarchiv/9830/5.html	.00	.00	.00	
8	19980008.00	.00	1997	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5014.html	.00	.00	.00	
9	19980009.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9851/5659.html	.00	.00	.00	
10	19980010.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9845/5776.html	.00	.00	.00	
11	19980011.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5799.html	.00	.00	.00	
12	19980012.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5800.html	.00	.00	.00	
13	19980013.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5801.html	.00	.00	.00	
14	19980014.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5817.html	.00	.00	.00	
15	19980015.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9845/5840.html	.00	.00	.00	
16	19980016.00	1.00	.	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5847.html	.00	.00	.00	
17	19980017.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5848.html	.00	.00	.00	
18	19980018.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5849.html	.00	.00	.00	
19	19980019.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5910.html	.00	.00	.00	
20	19980020.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5936.html	.00	.00	.00	
21	19980021.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5961.html	.00	.00	.00	
22	19980022.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/5999.html	.00	.00	.00	
23	19980023.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9845/6004.html	.00	.00	.00	
24	19980024.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/6006.html	.00	.00	.00	
25	19980025.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9844/6007.html	.00	.00	.00	
26	19980026.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9845/6008.html	.00	.00	.00	
27	19980027.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9844/6052.html	.00	.00	.00	
28	19980028.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9844/6053.html	.00	.00	.00	
29	19980029.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/6063.html	.00	.00	.00	
30	19980030.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9843/6064.html	.00	.00	.00	
31	19980031.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9844/6084.html	.00	.00	.00	
32	19980032.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9844/6092.html	.00	.00	.00	
33	19980033.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9845/6093.html	.00	.00	.00	
34	19980034.00	1.00	1998	http://regi.kozbeszerzes.hu/static/KEarchiv/9849/6130.html	.00	.00	.00	

Data extraction & database building

From non-stuctured/semi-structured (text, html, pdf) data to a structured database

1. Database definition (sql)
2. Data mining / text mining (Phyton, Java, php)
3. SQL database creation
4. Automatic text extraction (Phyton, Java, php)
5. (Human assisted) data correction / cleaning, imputation
6. Testing data quality (SPSS, STATA)
7. Data analysis and indicator creation (SPSS, STATA, R)

Blueprint for measuring institutionalised grand corruption in PP

1. Corruption Risk Index (CRI): generation and allocation of rents

2. Political Influence Indicator (PII): political influence on companies' market success

3. Political Control Indicator (PCI): direct political control of contractors

Corruption Risk Index (CRI)

- Probability of institutionalised grand corruption to occur

$$0 \leq CRI^t \leq 1$$

where 0=minimal corruption risk; 1=maximal observed corruption risk

- Composite indicator of 13 elementary risk (CI) indicators

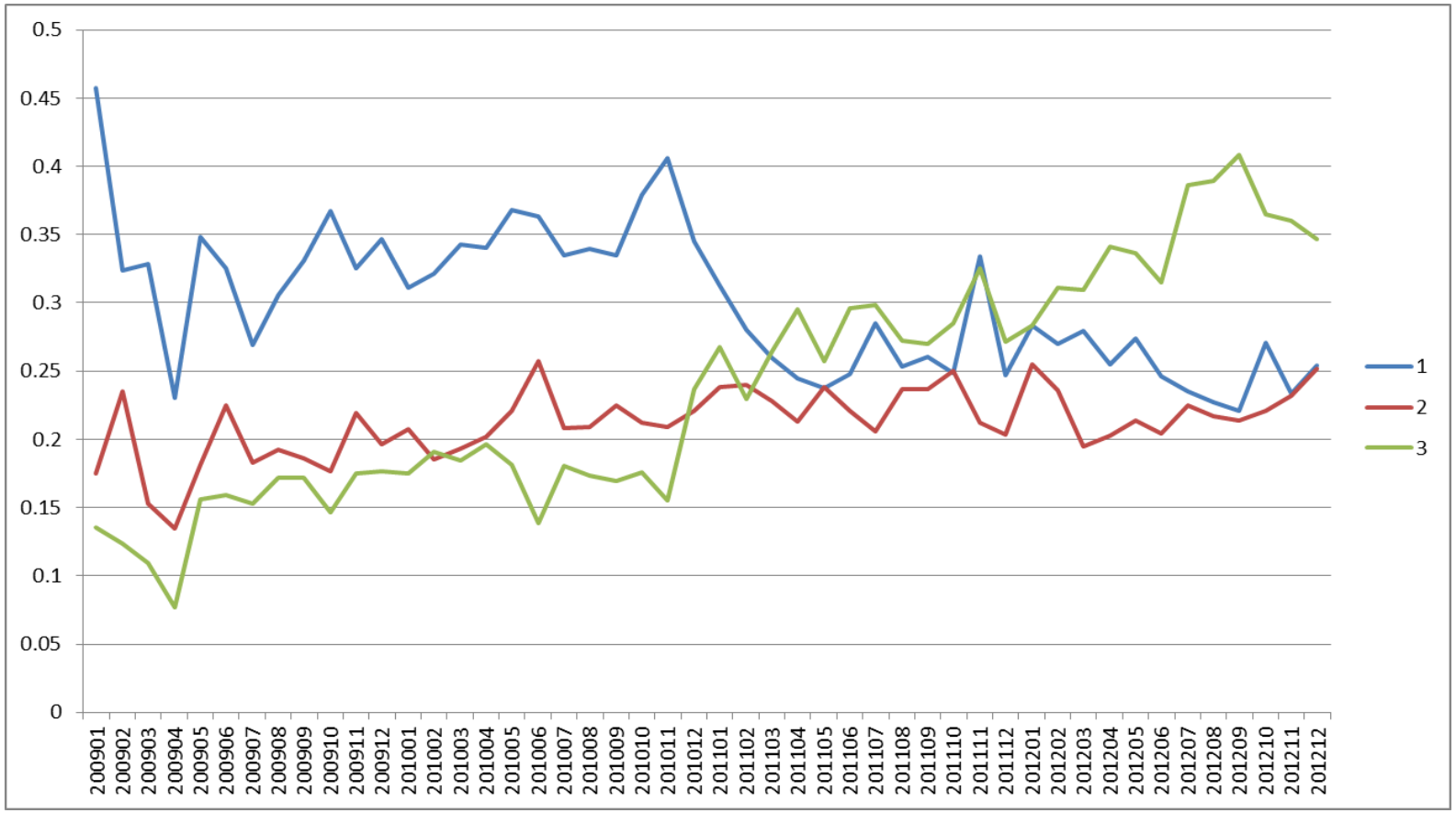
$$CRI^t = \Sigma_j w_j * CI_j{}^t$$

CRI construction

1. Wide set of potential components: 30 CIs

Examples of elementary indicators

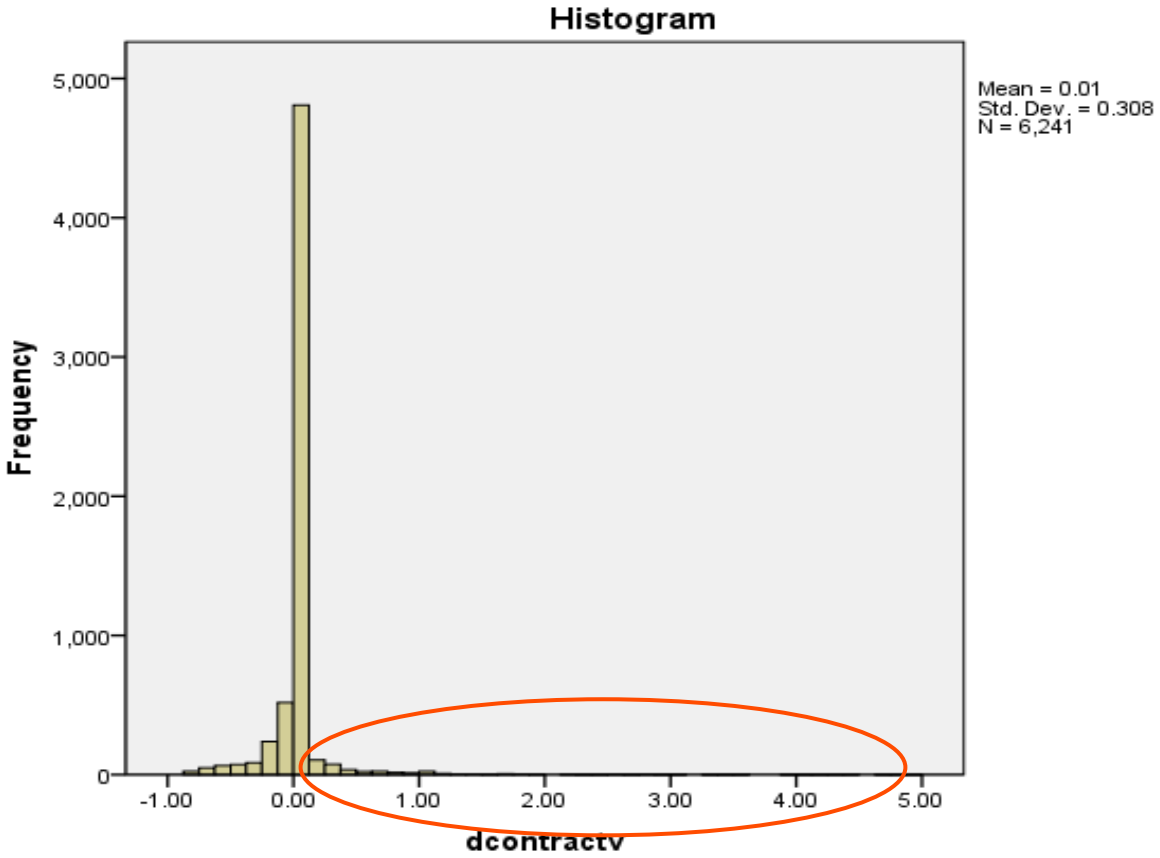
1. Number of submitted bids in Hungary (2009-2012)



CORRUPTION RESEARCH CENTER BUDAPEST

Examples of elementary indicators

2. Contract value increase during delivery in Hungary (2009-2012)



CORRUPTION RESEARCH CENTER BUDAPEST

CRI construction

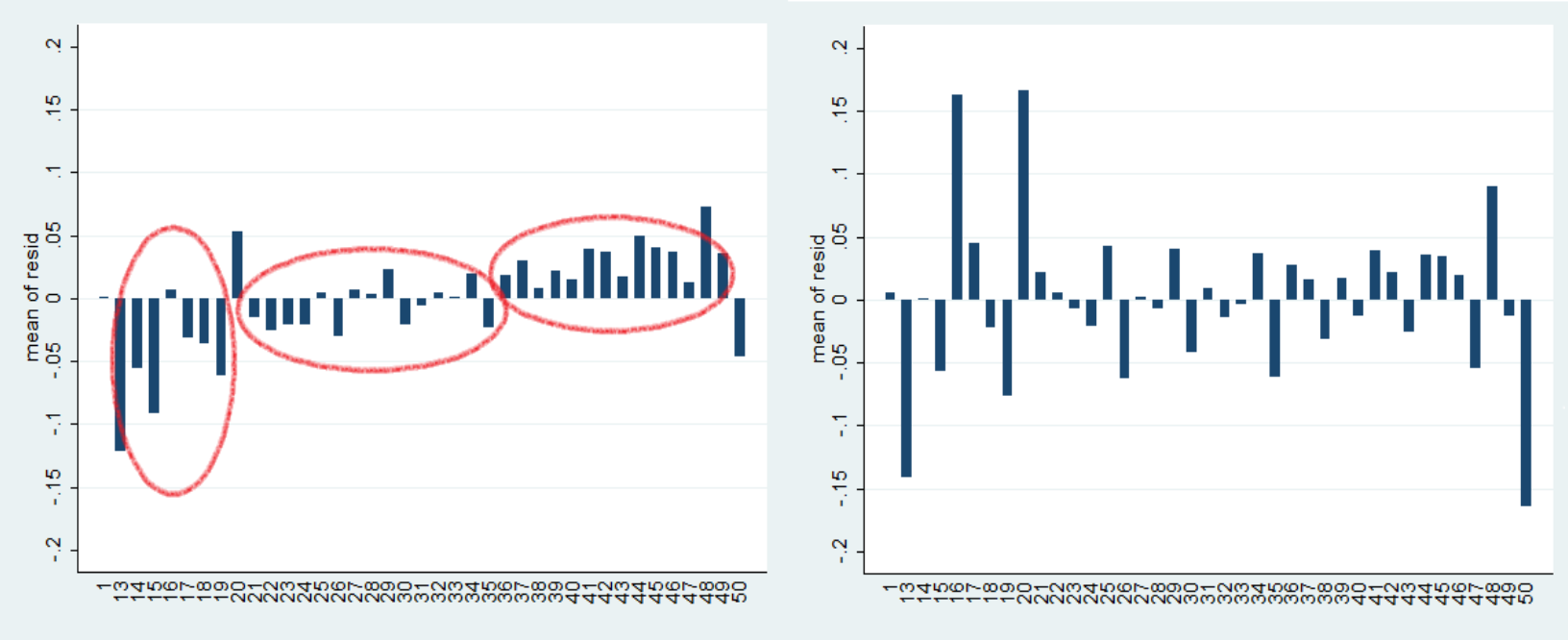
1. Wide set of potential components: 30 CIs
2. Narrowing down the list to the relevant components: 13 CIs
   - Set of regressions on single bidder and winner contract share (follow from definition!)

Regression setup

- Outcome variables
  - Single bidder (binary logistic regression)
  - Winner contract share (OLS)
- Explanatory variables:
  - Elementary corruption indicators
- Control variables:
  - Contract size
  - Type of market
  - Year
  - Authority type, xector, and status
  - Number of unique winners on the market

CRI-red flag identification

- Regressions define thresholds in continuous variables
- Example: relative price of tender documentation



CRI construction

1. Wide set of potential components: 30 CIs
2. Narrowing down the list to the relevant components: 13 CIs
   - Set of regressions on single bidder and winner contract share (follow from definition!)
3. CRI calculation: determining weights
   - Stronger predictor→higher weight
   - Norming to 0-1 band

Components of CRI

1. *Single bidder*
2. *Winner's contract share*
3. *Call for tender not published in official journal*
4. *Procedure type*
5. *Length of eligibility criteria*
6. *Lenght of submission period*
7. *Relative price of tender documentation*
8. *Call for tenders modification*
9. *Weight of non-price evaluation criteria*
10. *Annulled procedure re-launched subsequently*
11. *Length of decision period*
12. *Contract modification*
13. *Contract value/duration increase*

CRI composition in detail

Categorical variables using **thresholds**

Weights reflecting our limited understanding of the **exact** process

variable	component weight
single received/valid bid	0.096
no call for tenders published in official journal	0.096
procedure type	
ref. cat.=open procedure	0.000
1=invitation procedure	0.048
2=negotiation procedure	0.072
3=other procedures	0.096
4=missing/erroneous procedure type	0.024
relative length of eligibility criteria	
ref.cat.=length<-2922.125	0.000
1= -2922.125<length<=520.7038	0.024
2= 520.7038<length<=2639.729	0.048
3= 2639.729<length	0.072
4= missing length	0.096
short submission period	
ref.cat.=normal submission period	0.000
1=accelerated submission period	0.048
2=exceptional submission period	0.072
3=except. submission per. abusing weekend	0.096
4=missing submission period	0.024
relative price of tender documentation	
ref.cat.= relative price=0	0.000
1= 0<relative price<=0.0004014	0.000
2= 0.0004014<relative price<=0.0009966	0.096
3= 0.0009966<relative price<=0.0021097	0.064
4= 0.0021097<relative price	0.032
5=missing relative price	0.000
call for tenders modification(only before 01/05/2010)	0.096
weight of non-price evaluation criteria	
ref.cat.= only price	0.000
2= 0<non-price criteria weight<=0.4	0.000
3= 0.4<non-price criteria weight<=0.556	0.048
4= 0.556<non-price criteria weight<1	0.096
5=only non-price criteria	0.000
procedure annulled and re-launched subsequently	0.096
length of decision period	
ref.cat.= 44<decision period<=182	0.000
1= decision period<=32	0.064
2= 32<decision period<=44	0.032
4= 182<decision period	0.096
5= missing decision period	0.000
contract modified during delivery	0.096
contract extension(length/value)	
ref.cat.= c.length diff.<=0 AND c.value diff.<=0.001	0.000
2= 0<c. length d.<=0.162 OR 0.001<c.value d.<=0.24	0.096
3= 0.162<c. length diff. OR 0.24<c.value diff.	0.000
4= missing (with contr. completion ann.)	0.048
5= missing (NO contr. completion ann.)	0.000
winner's market share	0.096

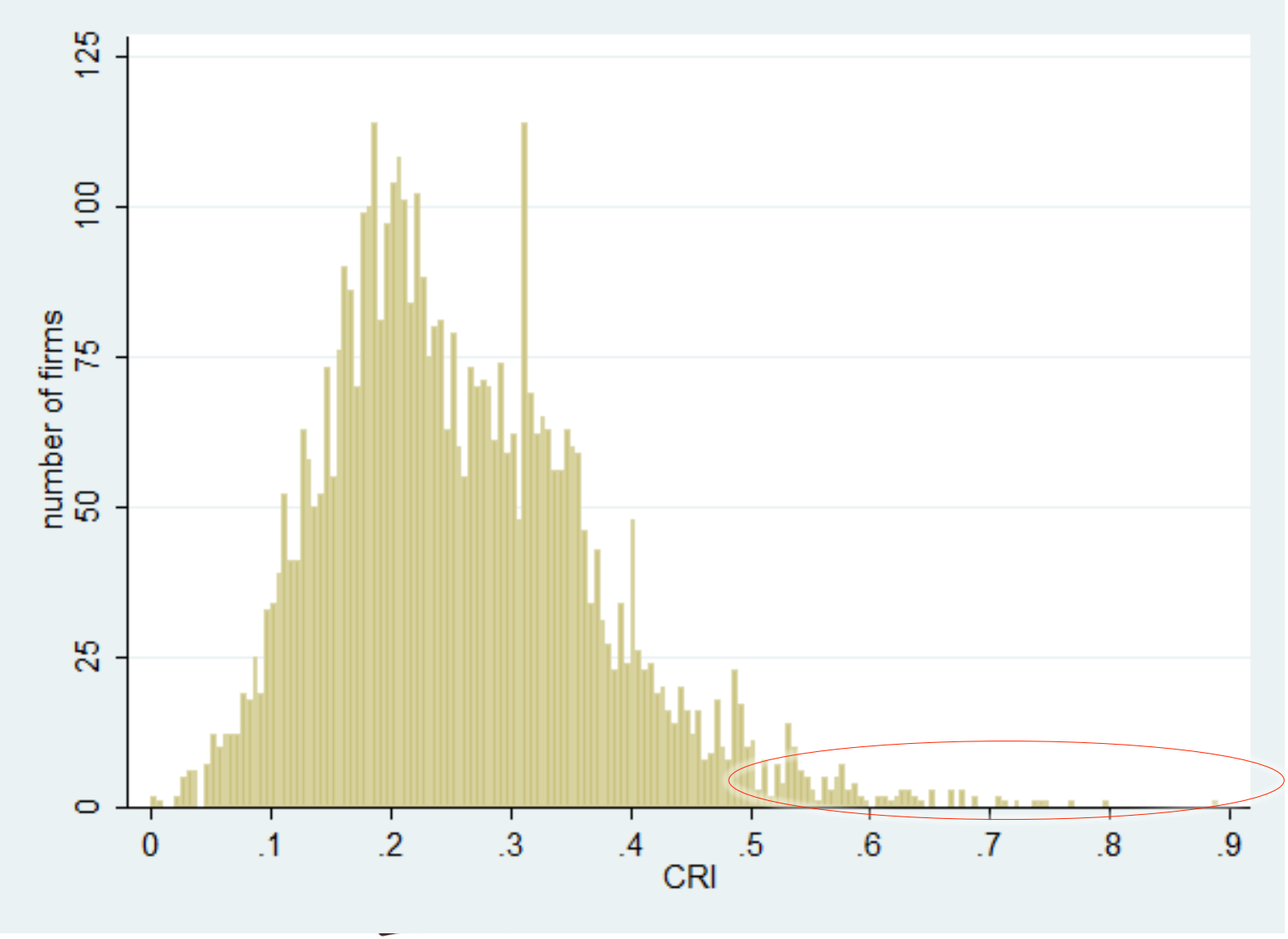
What kind of CRI distributions arise?

average CRI

Per winning bidder

2009-2012

Hungary



Political Influence Indicator (PII)

- Whether a company's market success depends on the political group in power

$$PII_i = \begin{cases} 1, \text{ if company } i \text{ is dependent on gov't} \\ 0, \text{ if company } i \text{ is NOT dependent on gov't} \end{cases}$$

PII construction

1. Baseline regressions
   - Explaining contract volume: BEFORE-AFTER gov't change

Basic regression setup

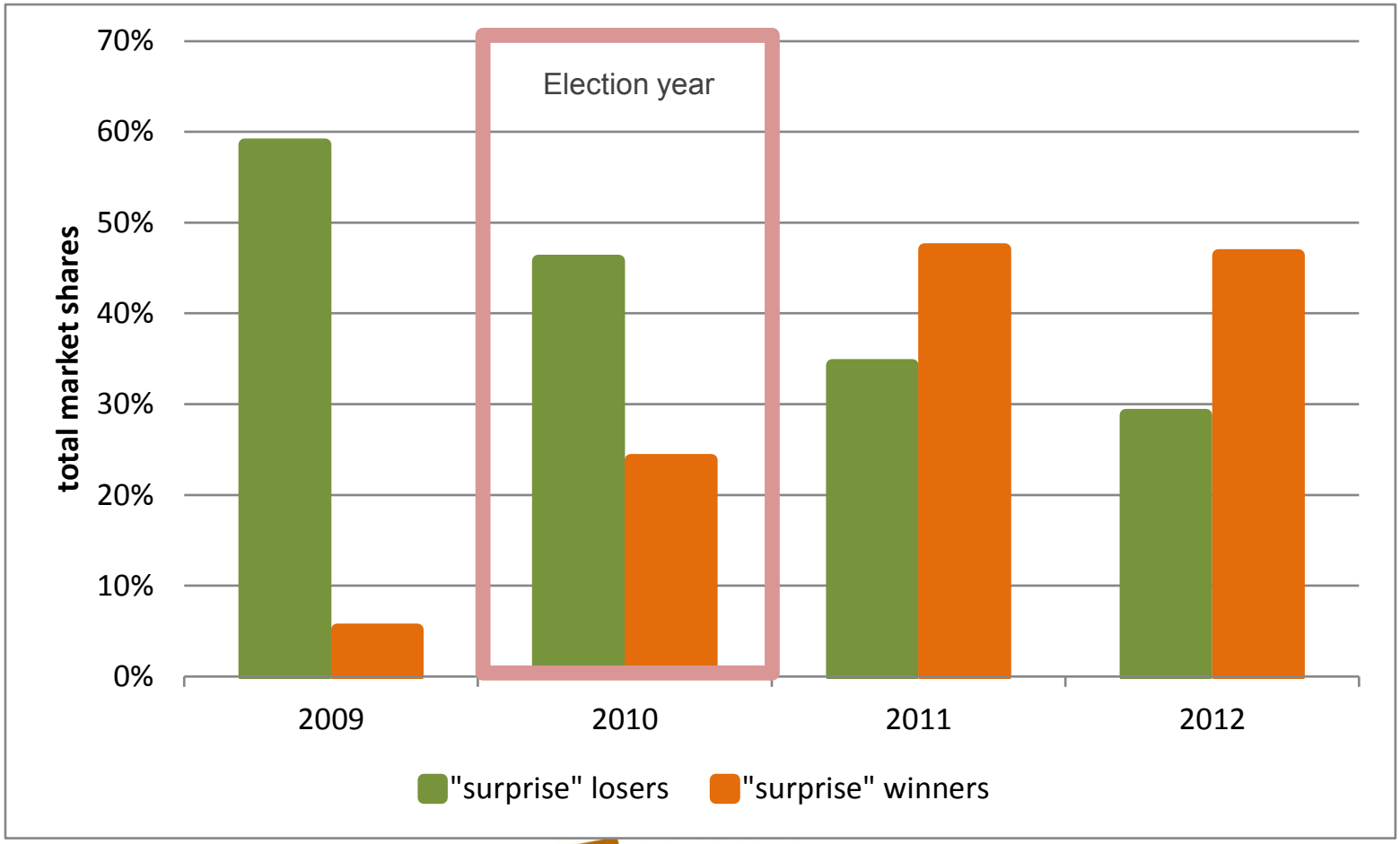
- Multilevel Modelling (main market as level)
- Outcome variable
  - logarithm of the difference of total contract value won in 2009 and 2011
- Company-level control variables:
  - location: county of company headquarters,
  - log employment (2009),
  - log turnover (2009),
  - log capital expenditure (2009), and
  - profit margin (2009)
- Market-level control variable
  - Hirschman-Herfindahl Index (2009)
- Separate analysis of entrants (without 2009 values)

PII construction

1. Baseline regressions
   - Explaining contract volume: BEFORE-AFTER gov't change
2. Benchmark regressions
   - Same regressions as in 1), but for periods WITHOUT gov't change
3. Marking companies
   - Significant and substantial differences between 1) and 2)

How does this look in pratice?

Hungary, total public procurement market, HU, 2009-2012



Political Control Indicator (PCI)

- Whether a company has direct political connections

$$PCI_i = \begin{cases} 1, \text{ if company } i \text{ has pol. connections} \\ 0, \text{ if company } i \text{ does NOT have pol. conn.} \end{cases}$$

PCI construction

1. Collecting names
   - Winners: company registry
   - Political officeholders: electoral registry, company registry, treasury records
2. Matching names/individuals
   - Biographical data
   - Statistical matching: name frequency, geographical distance
3. Marking companies

Indicator validity 1.

- Our corruption indicators co-vary
- CRI + PCI, HU, 2009-2012

Group	N	Mean CRI	Std. Err.	Std. Dev.	95% Conf.Interval	
0=*no* political connection	2900	0.254	0.002	0.111	0.250	0.258
1=politically connected	1449	0.265	0.003	0.110	0.260	0.271
combined	4349	0.258	0.002	0.111	0.254	0.261
difference (CRI1-CRI0)		**-0.011*** **	0.004		-0.018	-0.004

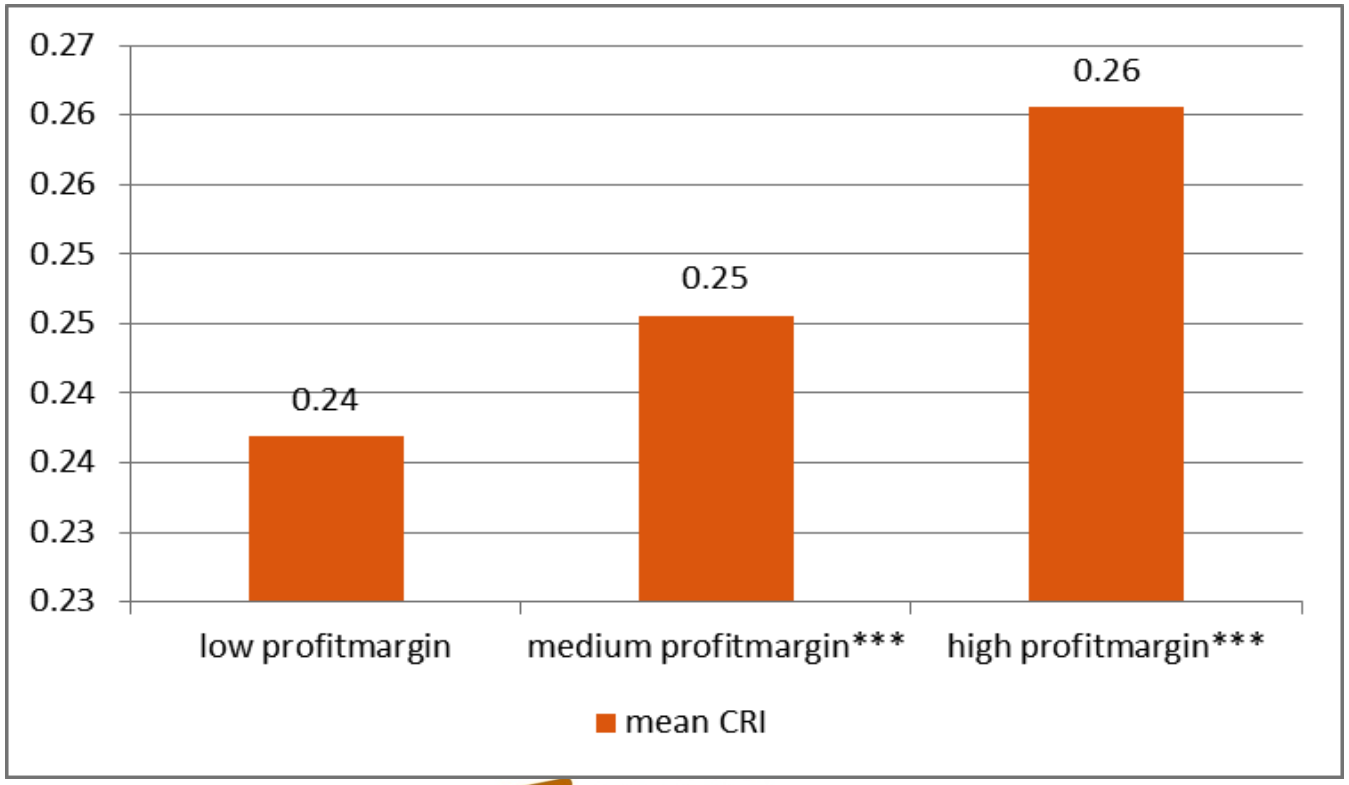
Indicator validity 1.

- Our corruption indicators co-vary
- CRI + PII, HU, 2009-2012

Group	N	Mean CRI	Std. Err.	Std. Dev.	95% Conf.Interval	
0=success *not* linked to government change	428	0.205	0.006	0.120	0.193	0.216
1=success linked to government change	2481	0.214	0.002	0.111	0.210	0.219
combined	2909	0.213	0.002	0.112	0.209	0.217
difference (CRI1-CRI0)		***0.010****	0.006		0.021	-0.002

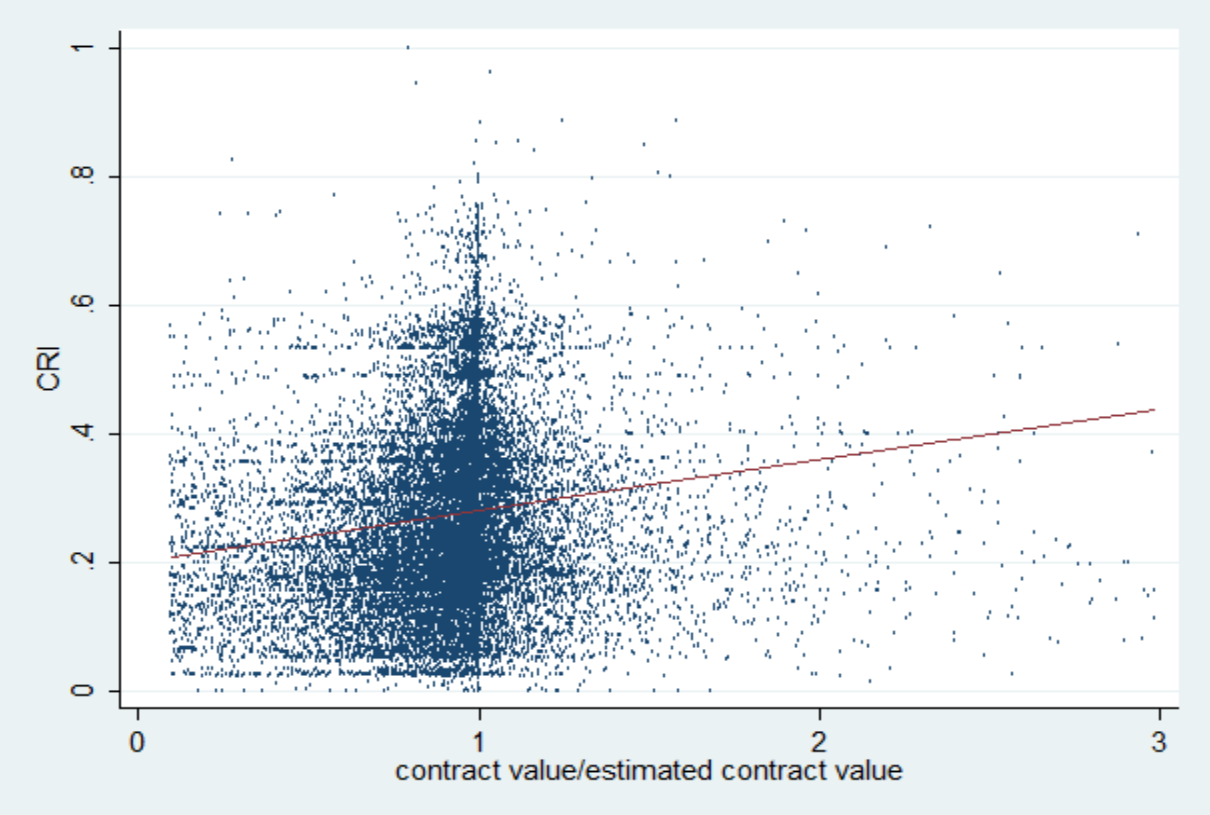
Indicator validity 2.

- Our indicators relate to external variables as expected: **rent extraction**
- Profitmargin + CRI in HU, 2009-2012



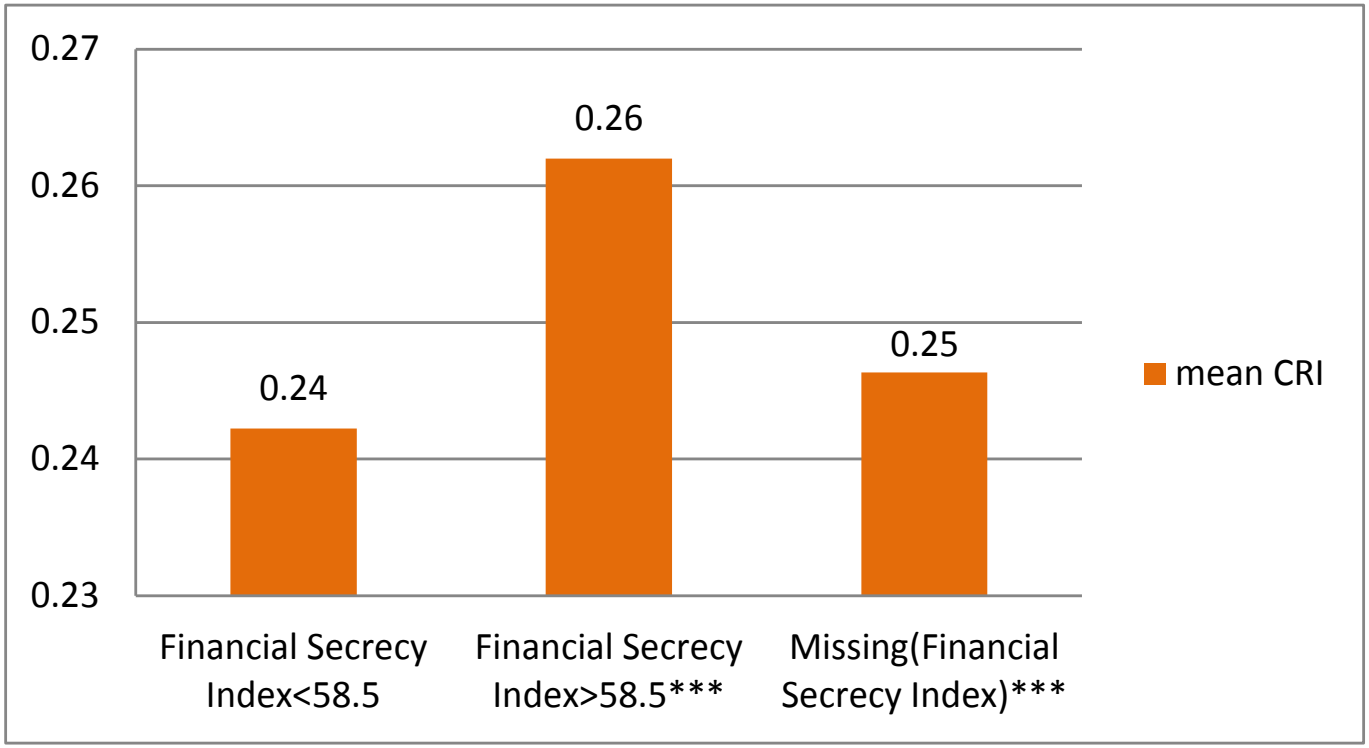
Indicator validity 2.

- Our indicators relate to external variables as expected: **rent extraction** from PP contracts
- Relative contract value + CRI in HU, 2009-2012



Indicator validity 2.

- Our indicators relate to external variables as expected: **money laundering, diversion of funds**
- Financial Secrecy Index + CRI in HU, 2009-2012



Limitations

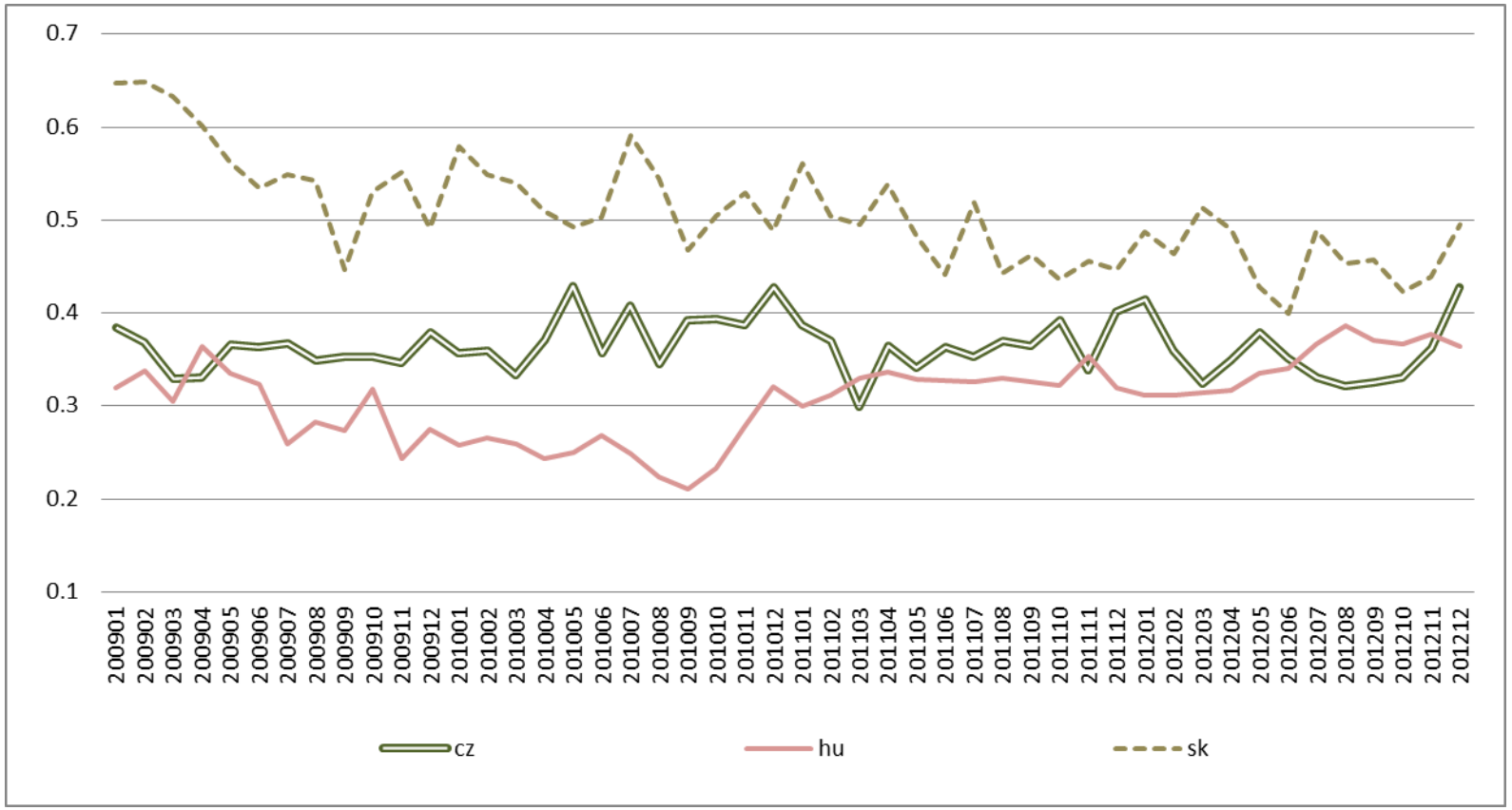
- You get what you measure: **no general indicator of corrupotion!**
- Reflexivity
- Two essential requirements
  - **Scope**: transparency is a preprequisite: minimum amount of coverage and detail is necessary
  - **Variance**: we need to compare corrupt to non-corrupt: some countries might not work
- Considering complex strategies for limiting competition: e.g. cartels

Applications – overview

1. **Evaluating countries:** against each other or the same country over time
2. **Evaluating large funding programmes:** e.g. EU structural funds in CEE
3. **Assessing the network structure of corruption:** e.g. identifying key points of policy intervention
4. **Evaluating regulatory or organisational reform:** e.g. loosening transparency regulations, integrity systems
5. **Risk-based audit of actors/transactions**

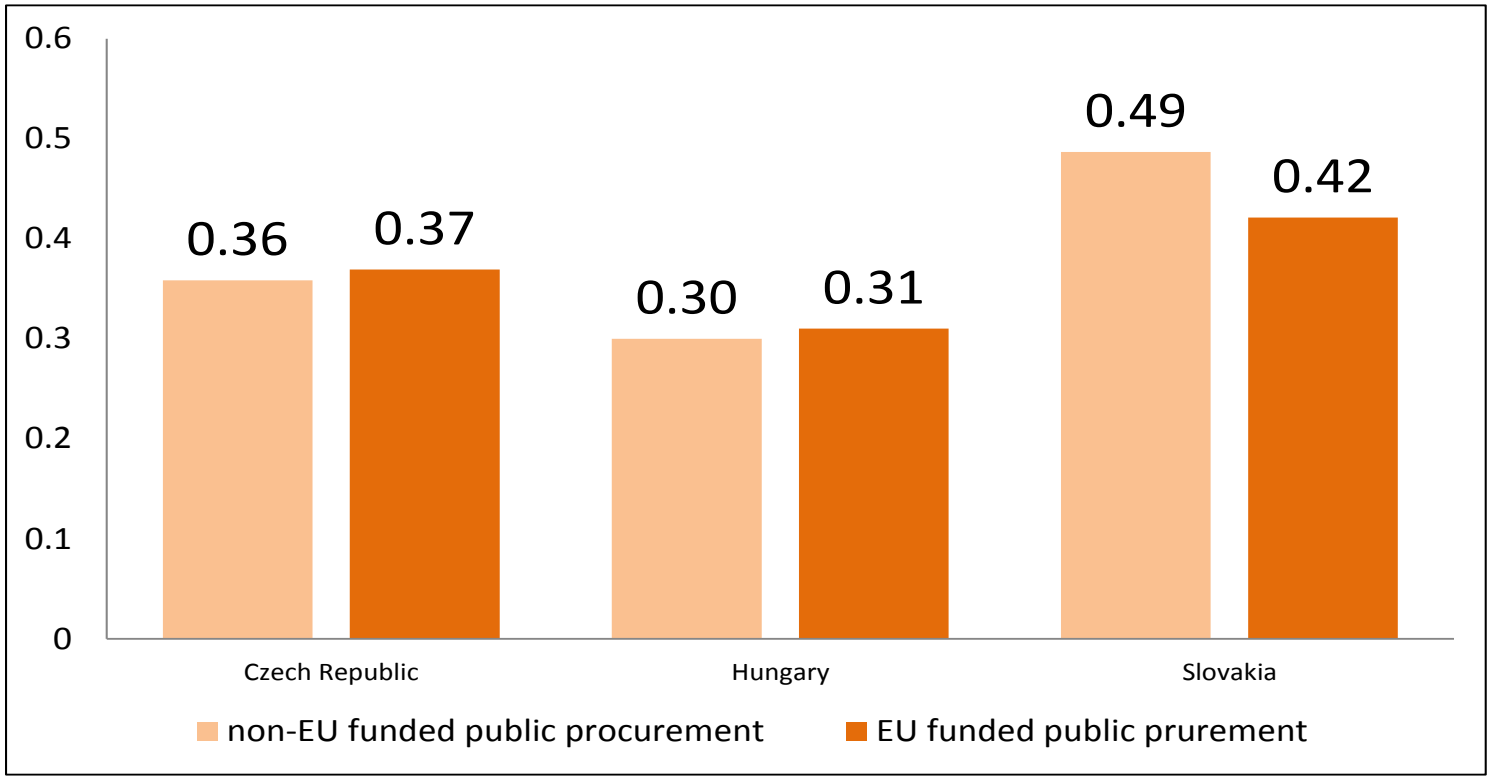
Applications 1. – tracking corruption over time and across countries

- Avg. CRI over time in CZ, HU, SK: 2009-2012



Applications 2. EU Funding in CEE

- EU structural funds' impact on corruption in CEE



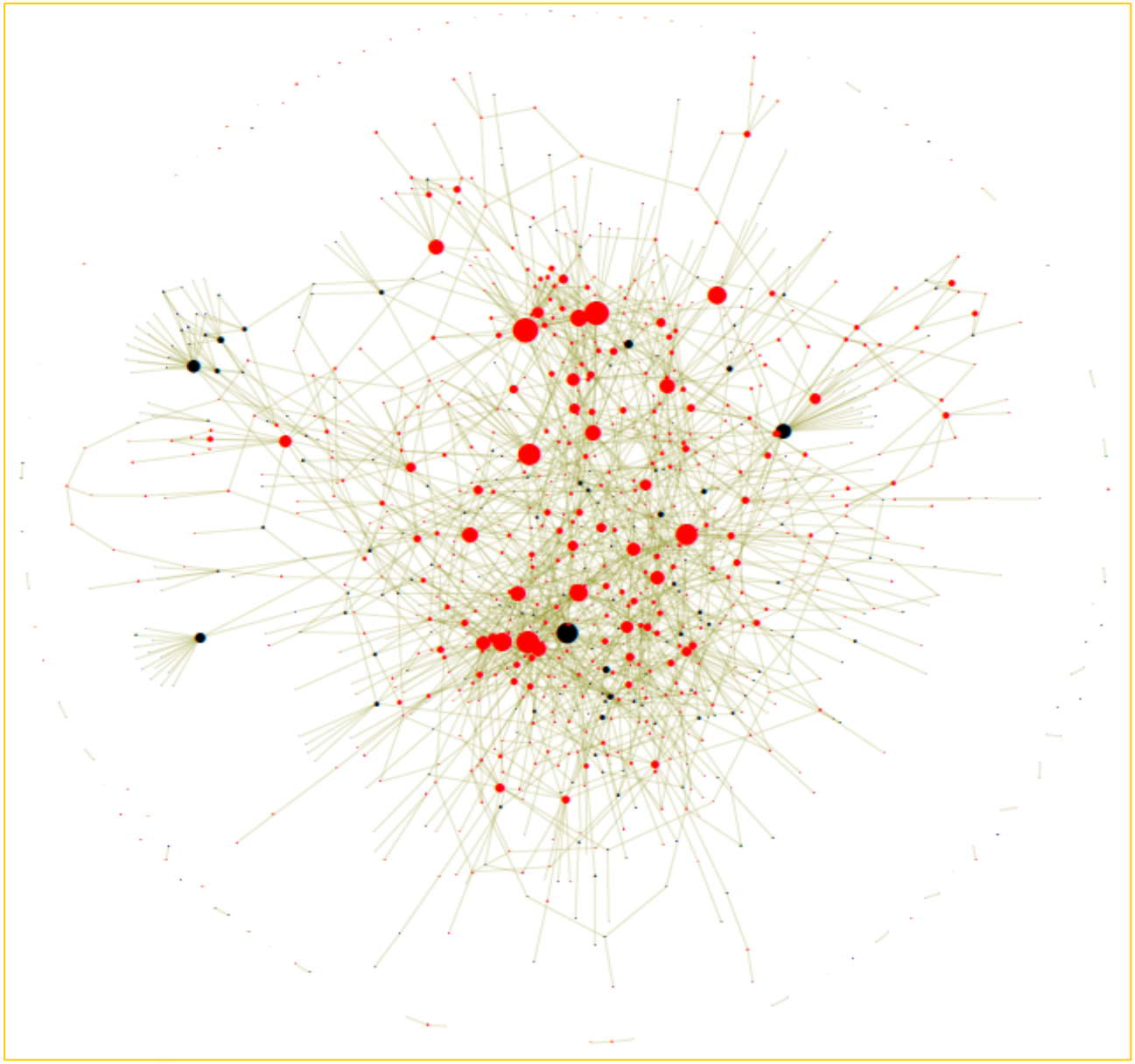
Applications 2. EU Funding in CEE

- Decomposing CRI differentials

variable/country	cz	sk	hu(comp)	hu(ext)
Winner contract share	++	++	++	++
Single bid	+	+	+	+
NO call for tenders published in o. journal	- -	-	-	-
Procedure type	- -	-/+	-	0
Length of submission period	- -	- -	- -	-/0
Length of decision period	-/+	-/+	-/0	-/0
Modification of call for tenders	+			0
Number of assessment criteria	-/0		-/+	
Weight of non-price evaluation criteria				++
Length of eligibility criteria				++
Relative price of documentation				-
Annulled procedure re-launched subsequently				-
Contract modification				++
Contract lengthening				- -

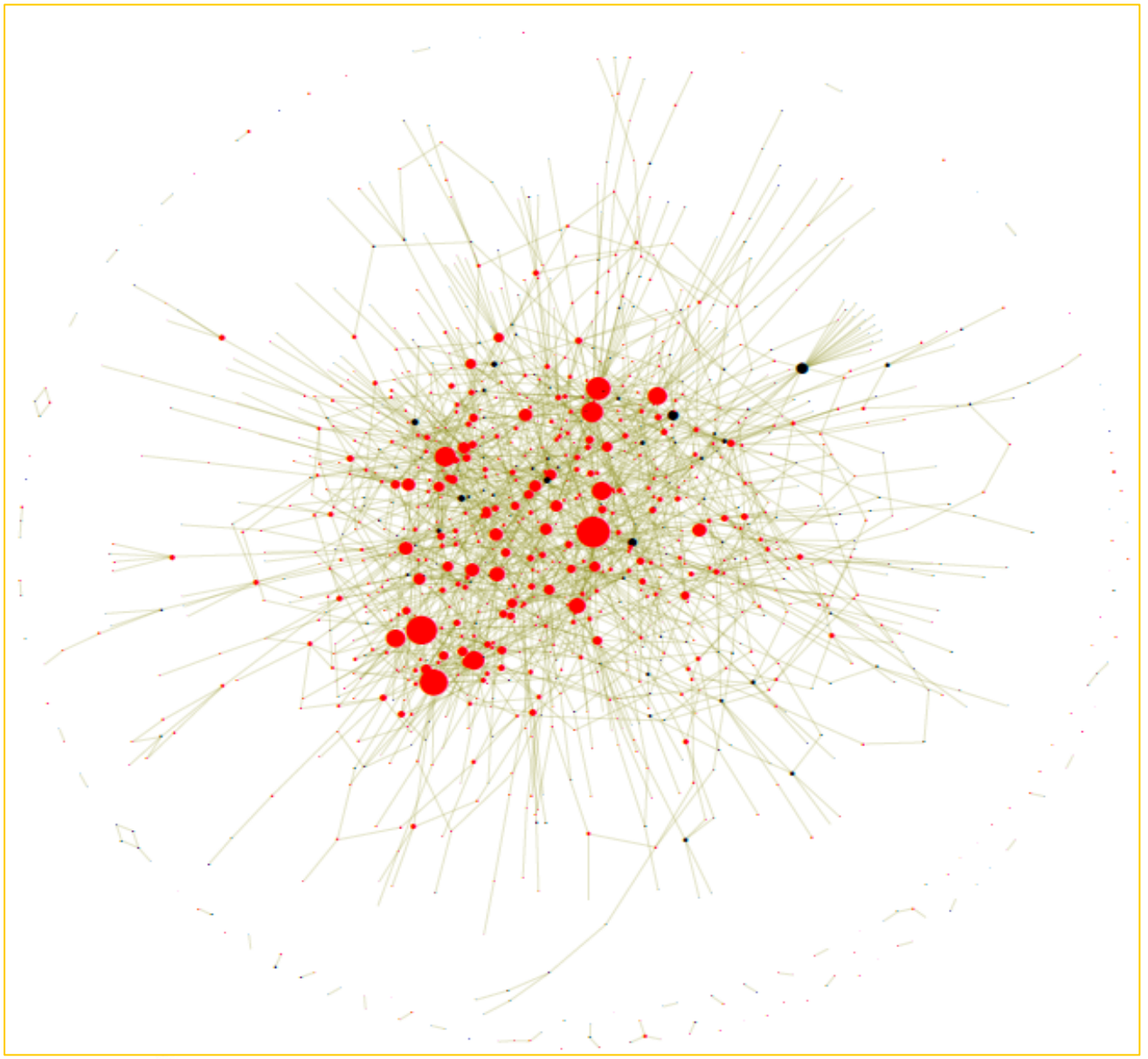
Applications 3. State capture

Captured org.s' network, HU, 2009-2010Q2



Applications 3.: State capture

Captured org.s' network, HU, 2011-201 2Q2



Applications 4.: civil service pay

Wages lower AND higher than average wages increase CRI

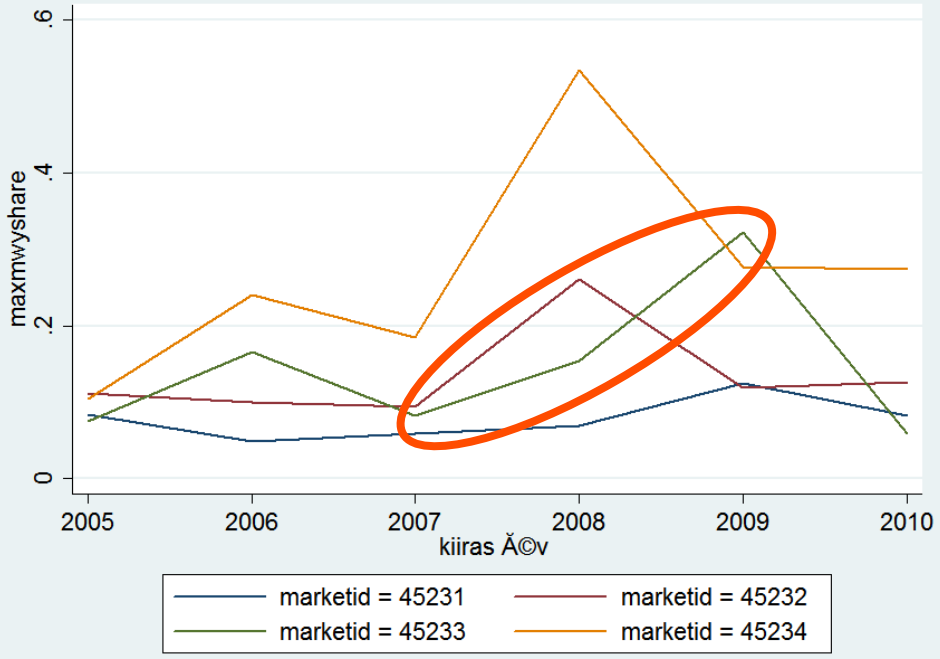
model #	1	2	3	4
dependent var.	CRI (org)			
independent var.	average real monthly wage (eur)			
linear	0.00002(0.22)	-0.00002(0.56)		
quadratic		0.00000(1.00)		
categorical				
ref.cat.:495<=w<576				
w<411			**0.0331(0.00)**	
411<=w<495			**0.0207(0.01)**	
576<=w<751			0.0123(0.17)	
751<=w			**0.0283(0.00)**	
lagged categorical				
ref.cat.:495<=w<576				
w<411				**0.0322(0.01)**
411<=w<495				**0.0298(0.01)**
576<=w<751				0.0136(0.29)
751<=w				0.0179(0.47)
control variables	transparency, size, type, sector, public procurement spending			
N	1679	1679	1679	925
R^2	0.11	0.11	0.12	0.12

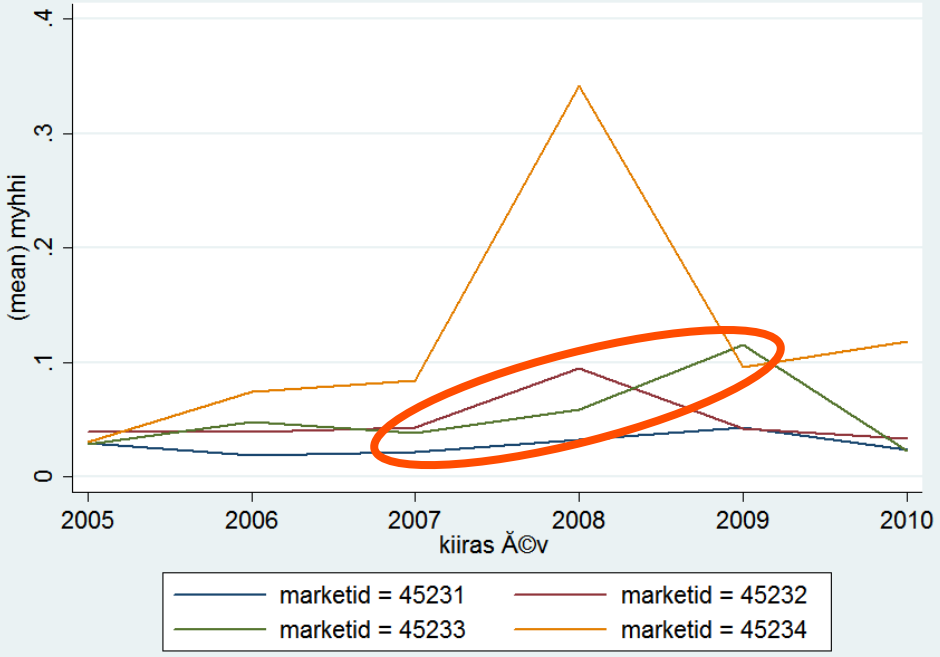
App. 5.:Initial red flags

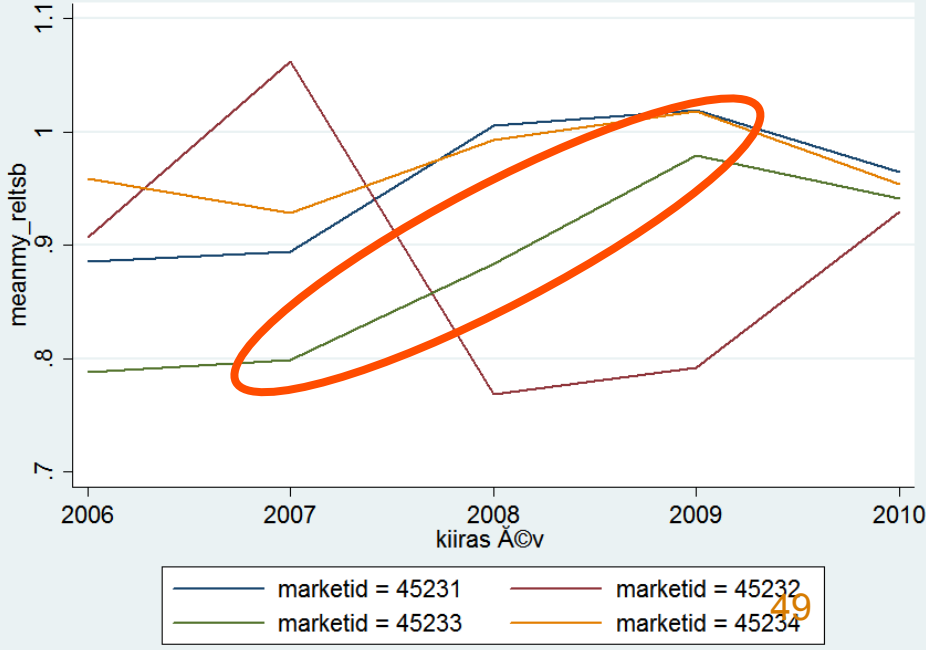
Market structure changing to monopolistic (leader market share and HHI)

Organised along geographical dimensions

Increased prices (relative contract value)





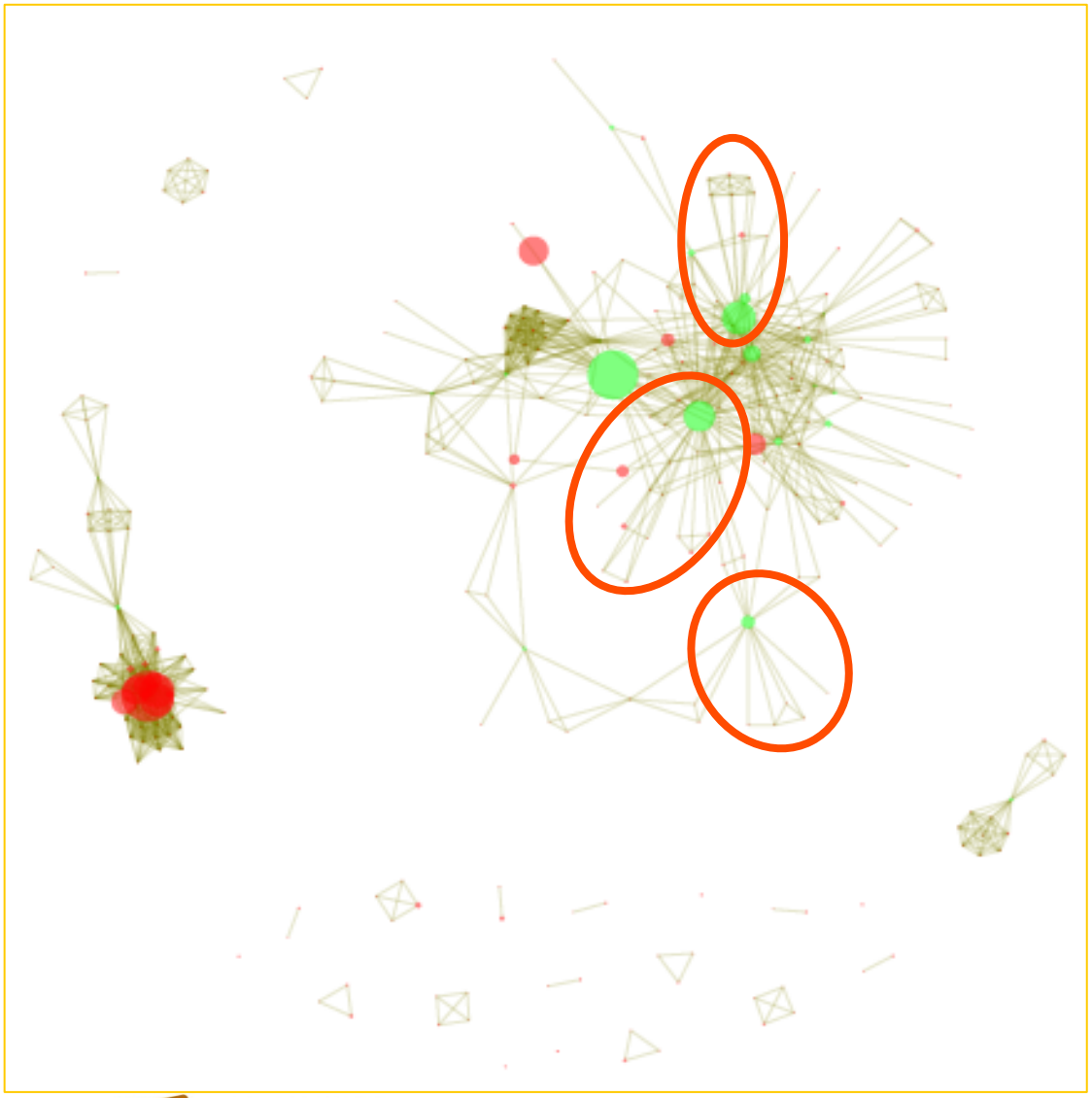


Appl. 5: co-bidding patterns benchmark

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

Applications 5.: cartel

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



Potential UK applications

- UK PP data:
  - contractsfinder, spendnetwork
  - Structured, rich data above EU threshold
  - Hard to get below EU threshold
  - Since 2008
- UK company and pol data:
  - Very good, easily accessible
- Issues to look at:
  - Revolving door
  - Local corruption
  - Bidding rings in large tenders

Looking forward to the discussion!

Further information about this approach

Corruption Research Center Budapest: www.crcb.eu

Published material:

Fazekas, M., Tóth, I. J. (2014), *In respectable society: on how elite configuration influences patterns of state capture in Hungary*. Conference paper, MPSA Annual Conference, Chicago, USA, 3 April 2014.

Fazekas, M., Tóth, I. J. (2014), *Three indicators of institutionalised grand corruption using administrative data*. Budapest: Corruption Research Centre.

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.

Fazekas, M., Chvalkovská, J., Skuhrovec, J., Tóth, I. J., & King, L. P. (2013). *Are EU funds a corruption risk? The impact of EU funds on grand corruption in Central and Eastern Europe*. CRC-WP/2013:03, Corruption Research Centre, Budapest.