CORRUPTION RESEARCH CENTER BUDAPEST

Mihály Fazekas[1], István János Tóth[2], Lawrence Peter King[3]

Anatomy of grand corruption:

A composite corruption risk index based on objective data

Working Paper series: CRC-WP/2013:02

Version 2.0

November 2013, Budapest, Hungary

The Corruption Research Center Budapest was created in November 2013 in response to the growing need for independent research on corruption and quality of government in Hungary. The central aim of the Center is to systematically explore the causes, characteristics, and consequences of low quality of government, corruption, and regulatory failure using an inter-disciplinary approach. In addition, the Center also aims to help citizens to hold governments accountable through the use of robust evidence. Our unique research approach combines qualitative and quantitative methods to analyse micro-level actor behaviour and generates novel hard data on the phenomena under scrutiny.

Corruption Research Center: http://www.crc.uni-corvinus.hu/

[1] University of Cambridge and Corruption Research Center, mf436@cam.ac.uk, www.mihalyfazekas.eu
[2] Hungarian Academy of Sciences and Corruption Research Center
[3] University of Cambridge

ABSTRACT

Anatomy of grand corruption: a composite corruption risk index based on objective data [4]

Although both the academic and policy communities have attached great importance to measuring corruption, most of the currently available measures are biased and too broad to test theory or guide policy. This article proposes a new composite indicator of grand corruption based on a wide range of elementary indicators. These indicators are derived from a rich qualitative evidence on public procurement corruption and a statistical analysis of a public procurement data in Hungary. The composite indicator is constructed by linking public procurement process 'red flags' to restrictions of market access. This method utilizes administrative data that is available in practically every developed country and avoids the pitfalls both of perception based indicators and previous 'objective' measures of corruption. It creates an estimation of institutionalised grand corruption that is consistent over time and across countries. The composite indicator is validated using company profitability and political connections data.

JEL classification: D72, D73, H57

Keywords: public procurement, grand corruption, corruption technique, composite corruption risk index

[4] The authors would like to express their gratitude for two EU funded projects at the Budapest Corvinus University (TAMOP 4.2.2.B and ANTICORRP (Grant agreement no: 290529)) even though they relied extensively on their voluntary contributions for realising this project. They would also like to express special thanks to colleagues at the Corruption Research Centre working on the Hungarian public procurement database (MakAB) for over three years, especially Kinga Csizmás, Ágnes Czibik, Zoltán Kelemen, and Tamás Uhrin. Furthermore, we would like to thank the colleagues at the University of Cambridge for their insightful comments on earlier drafts of this paper.

1. Introduction

Various corruption indices have received considerable academic, policy, and media attention, at least partially due to the central role the underlying phenomena play in the quality of democratic governance, the provision of public goods, economic growth, and equality. Understanding their importance, some international organisations regularly monitor corruption in their member countries (European Commission, 2011a) and even tie funding to performance on governance indicators including corruption (Andersson & Heywood, 2009; Radelet, 2002, 2003).

In the absence of robust objective measures, there are three major sources of corruption indicators to date: 1) surveys of corruption perceptions and attitudes (which are most widely used); 2) reviews of institutional and legal frameworks; and 3) detailed analyses and audits of individual cases. Unfortunately, each of these has serious deficiencies leaving us without any reasonably reliable and valid indicator of corruption allowing for comparing countries over time or exploring within country diversity.

In order to fill some of the gap between the demand for corruption indices and the dire state of the data currently available, the goal of this paper is *to develop a novel measure of institutionalised grand corruption* which:

1. solely derives from objective data describing behaviour,
2. is defined on the micro level such as individual transactions,
3. allows for consistent temporal comparisons within and across countries, and
4. rests on a thorough understanding of the corrupt rent extraction process.

In the context of public procurement, institutionalised grand corruption or legal corruption refer to the allocation and performance of public procurement contracts by bending prior explicit rules and principles of good public procurement in order to benefit a closed network while denying access to all others (for a related discussion see Kaufmann & Vincente, 2011; Mungiu-Pippidi, 2006; North, Wallis, & Weingast, 2009; Rothstein & Teorell, 2008).

The proposed indicator of institutionalised grand corruption fulfils all of the above criteria with potential for replication in most developed countries including every EU member state, Russia, and the US. Time series available in these countries range between 6-8 years. The approach makes use of micro-level data on individual public procurement procedures allowing for directly modelling corrupt actors' rent extraction activities. Institutionalised grand corruption in public procurement requires 1) the generation of corrupt rents and 2) the regular extraction of such rents. To achieve both of these, any corrupt group has to restrict competition prescribed by procurement laws to benefit a particular bidder multiple times. Hence, measuring the degree of competition restriction, recurrent contract awards to the same company, and the typical techniques used to achieve these goals allow for detecting institutionalised grand corruption consistently across countries, organisations and time.

The paper is structured as the follows: first, the literature on corruption measurement is reviewed; second, the proposed novel measurement approach is presented; third, Hungarian data and variables are summarized; fourth, the composite corruption risk index is constructed and some external validity measures offered; finally, conclusions and further research directions are provided.

2. Literature on measuring grand corruption

By now, an industry has emerged for measuring corruption. However, the available measurements are either fundamentally flawed or too narrow for testing theories of grand corruption and developing effective solutions to it.

In a broad sense, corruption indicators derive primarily from:

- Surveys of *attitudes, perceptions* and *experiences* of corruption among different stakeholders (e.g. general population, firms, experts);
- Reviews of *institutional features* controlling corruption in countries or individual organisations; and
- Audits and investigations of *individual cases* (see Kaufmann, Kraay, & Mastruzzi, 2006; Transparency International, 2012).

Among *perception and attitude surveys*, the two most widely used are the World Bank's Control of Corruption (Kaufmann, Mastruzzi, & Kraay, 2010) and Transparency International's Corruption Perceptions Index[5]. Both of these have received extensive criticism applicable to any similar survey (Andersson & Heywood, 2009; Kaufmann, Kraay, & Mastruzzi, 2007; Kurtz & Schrank, 2007a, 2007b; Lambsdorff, 2006). Without trying to be exhaustive, some of the key arguments include: perceptions may or may not be related to actual experience (Rose & Peiffer, 2012), they can be driven by general sentiment reflecting, for example economic growth (Kurtz & Schrank, 2007a) or media coverage of high profile corruption cases (Golden & Picci, 2005). Arguably, perceptions of grand corruption are even more unreliable than perceptions of everyday corruption since experts and citizens have almost no direct experience of this type of corruption. As both indicators and others of this type primarily derive from non-representative surveys, representativeness bias is likely to occur, in addition to reflexivity bias (i.e. respondents influenced by prior and future measurements) exaggerated by small sample sizes (Golden & Picci, 2005). These indicators vary surprisingly little over time given the large changes in underlying governance structures suggesting that they are too *in*sensitive to change (Arndt & Oman, 2006; Kurtz & Schrank, 2007a; Mungiu-Pippidi, 2011).

Surveys of experiences with corruption, that is low-level bribery, such as the Quality of Government Institute's regional survey (Charron, Dijkstra, & Lapuente, 2010) or surveys in Latin American countries (Seligson, 2002) while addressing some of the weaknesses of perception surveys fall short of a sufficient data source. A prime problem is non-response or false response to sensitive questions such as giving or receiving bribes. Most importantly, only a tiny fraction of the population has direct experience with grand corruption limiting the use of this method.

Reviews of institutions controlling corruption, while crucial in understanding the determinants of corruption, are, by design, not measuring corruption directly. In the absence of a precisely measured outcome variable, they have to rely on untested theories of which institutional features work.

[5] http://cpi.transparency.org/cpi2012/results/ (accessed: 16/1/2013)

Analyses of individual cases are highly reliable in establishing and explaining both petty and grand corruption, however, their narrow scope and lack of generalizability make them of only limited use for comparative purposes.

2.1 Objective measures of corruption

Some authors recognising the deficiencies of the above indicators have embarked on developing objective measures which rely on directly observable, hard indicators of behaviour that likely indicate corrupt behaviour (Table 1). These studies look into corruption in various contexts such as elections and high level politics or welfare services and redistributive politics. For example Olken (2007) uses independent engineers to review road projects and calculates the amount and value of missing inputs to determine corruption. More closely associated with our approach are those studies which focus on corruption in public procurement and bidding markets. For example, Golden & Picci (2005) propose a new measure of corruption based on the difference between the quantity of infrastructure and public spending on it. Other authors use some indicators also part of our composite indicator such as the use of exceptional procedure types (Auriol, Flochel, & Straub, 2011) or explicit scoring rules (Hyytinen, Lundberg, & Toivanen, 2008) or political connections of winning companies (Goldman, Rocholl, & So, 2013).

While these papers inspired our approach and point in the right direction, they cannot readily be scaled up to allow for temporal comparisons across countries and organisations. The reason is that they rely on a too narrow single indicator which may or may not be the primary vehicle for corrupt rent extraction depending on the regulatory framework in place (Olken & Pande, 2012). For example, corruption linked to exceptional procedure types may be easily removed by simply deleting the procedure from the procurement law, however it is unlikely that this alone would change the underlying corrupt phenomena much (Auriol et al., 2011). Instead, these and further elementary indicators have to be combined for meaningful temporal international comparisons.

Table 1. Summary of selected studies using objective indicators of corruption

paper	indicator used	Country	year	sector	potential for international comparison	part of CRI*
(Auriol et al., 2011)	Exceptional procedure type	Paraguay	2004-2007	general procurement	HIGH If procedure definitions can be aligned, international comparisons can be made widely	Yes
(Bandiera, Prat, & Valletti, 2009)	Price differentials for standard goods purchased locally or through a national procurement agency	Italy	2000-2005	various standardized goods (e.g. paper)	LOW Price data is not readily available in most countries, many countries don't have national procurement agencies, national procurement agencies are likely to be captured in many countries.	No
(Coviello & Gagliarducci, 2010)	Number of bidders Same firm awarded contracts recurrently Level of competition	Italy	2000-2005	general procurement	HIGH Number of bidders, recurrent contract award, and competitiveness of bids are available in many countries.	Yes
(Di Tella & Schargrodsky, 2003)	Difference in prices of standardized products such as ethyl alcohol	Brazil	1996-1997	health care	MEDIUM Detailed product-level price and quantity information is not readily available across many countries, but can be collected.	No
(Ferraz & Finan, 2008)	Corruption uncovered by federal audits of local government finances	Brazil	2003	federal-local transfers	LOW high quality audits, not influenced by powerful corrupt groups are unlikely to be available in many countries.	No
(Golden & Picci, 2005)	Ratio of physical stock of infrastructure to cumulative spending on infrastructure	Italy	1997	infrastructure	MEDIUM It is hard to compute comparable value of the stock of physical capital across countries different in the quality of infrastructure and geography.	No
(Goldman et al., 2013)	Political office holders' position on company boards	USA	1990-2004	general procurement	HIGH Company contract volumes can be estimated in many countries and publicly listed companies political connections can be traced relatively easily.	No**
(Hyytinen et al., 2008)	Number and type of invited firms Use of restricted procedure	Sweden	1990-1998	cleaning services	HIGH Both number of bidders and procedure types are readily available in many countries.	Yes
(Olken, 2006)	Difference between the quantity of in-kind benefits (rice) received according to official records and reported survey evidence	Indonesia	1998-1999	welfare spending	MEDIUM It is possible to design user surveys across a wide range of countries to track actual receipts, although it may be expensive.	No
(Olken, 2007)	Differences between the officially reported and independently audited prices and quantities of road construction projects	Indonesia	2003-2004	infrastructure (roads)	LOW Auditing large numbers of projects by independent engineers is costly and unlikely to allow for cross-country comparisons.	No
(Reinikka & Svensson, 2004)	Difference between block grants received by schools according to official records and user survey	Uganda	1991-1995	education	MEDIUM It is possible to design user surveys across a wide range of countries to track actual receipts, although it may be expensive.	No

*CRI=Corruption Risk Index, developed in this paper; **This approach is utilized in (Fazekas, Tóth, & King, 2013a).

3. The measurement approach

3.1 Corrupt rent extraction in public procurement

Institutionalised corruption's primary aim is earning corruption rents. Corruption rents in public procurement can be earned if and only if the winning contractor is a *pre-selected* company which earns *extra profit* due to higher than market price for the delivered quantity and/or quality.

The winning company has to be pre-selected in order to control rent extraction in an institutionalised manner. This rules out occasional corruption where the company is lured into corruption during the public procurement process. Extra profit has to be realised in order to create the pot of money from which rents can be paid.

In order to adequately measure extra profit; price, delivered quantity, and quality of deliveries has to be known with high precision. However, none of these three can adequately be measured. Price and quantity are publicly available, but they are comparable only for homogenous products such as electricity without laborious case-by-case analysis and even then it is difficult to arrive at accurate estimates. Quality cannot be reliably observed in official records without using expensive expert knowledge. Hence, *we can only measure the process of awarding contracts to pre-selected companies*.

Competition has to be eliminated or tilted in order to award the contract to the pre-selected company. Bypassing competition can be done in three primary forms, each corresponding to a phase of the public procurement process:

1. Limiting the set of bidders: submission phase;
2. Unfairly assessing bidders: assessment phase; and
3. Ex-post modifying conditions of performance[6]: delivery phase.

On the one hand, these three elementary corruption strategies can be combined in any way to reach the final desired outcome. For example, some bidders may be excluded with a tightly tailored eligibility criteria while the remaining unwanted bidders can simply be unfairly scored on subjective scoring items. On the other hand, once the desired outcome has been achieved at a given stage, there is no need for further corrupt actions which would increase the risk of detection with no additional benefit. For example, if the only company submitting a valid bid is the pre-selected company there is no need to modify contract content later to increase price.

3.2 Measurement model

[6] While modifying contract conditions does not belong to the set of company selection techniques, it can be part of an arsenal supporting the selection of the 'right' company. For example, the pre-selected company wins in a competitive process by promising low price and high quality knowing that later contract modifications will allow it to earn the agreed corruption rent.

Utilizing a public procurement database (for details see section 4), it is possible to measure a host of elementary indicators in relation to each of the above three stages of public procurement from which a composite indicator can be built (Fazekas, Tóth, et al., 2013a).

In order to most adequately model the company selection process, measurement is carried out on the level of individual contract award. Later, aggregation to organisation level per year can also be carried out to link procurement data to company profitability for example.

Likely *outcomes* of corrupt procurement procedures are defined for each of the above three main phases (see section 5.1). Indicators of likely corruption techniques to achieve these outcomes in each phase are also defined, which constitute the *inputs* for corrupt contract award and completion (see Fazekas, Tóth, et al., 2013b).

The *corrupt contract award process is modelled* using multiple regression linking likely corruption inputs (e.g. eligibility criteria tailored to one company) to likely corruption outcomes (e.g. only one company submitting a bid) in the presence of variables controlling for alternative explanations (e.g. number of competitors on the market). Our models linking corrupt inputs to outcomes in public procurement explain recurrent contract award to a pre-selected company with those corruption techniques which typically serve as means for corruptly eliminating competitors (Fazekas, Tóth, et al., 2013a).

The explanatory model linking corruption inputs to outcomes delivers a set of coefficients which represent the strength of association between each underlying likely corruption input and likely corruption outcome. Reliability of elementary corruption indicators is defined using their regression coefficients, as those corruption inputs which are more powerful in predicting probable corruption outcomes are more likely to signal corruption rather than noise. Falsely indicating corruption is minimised by dropping those indicators which didn’t prove to be powerful and significant predictors in the model and assigning lower component weights to those whose effect is only moderate.

In each country’s composite indicator, corruption outcomes, having no regression coefficients, receive *weight* of 1 reflecting their benchmark status in modelling the corruption process. Corruption outcomes measure most directly the underlying corrupt transactions hence their benchmark status. If overall model fit is adequate (i.e. passes standard tests of significance), the underlying model structure is verified supporting the conclusion that corruption outcome indicators are adequate themselves. Every powerful-enough corruption input receives a weight between 0 and 1, reflecting the size of its regression coefficient. This means that all weights are scaled compared to corruption outcomes.

For comparison across time and countries, both the list of components and component weights are kept constant unless there are differences in the institutional setup warranting any deviation. This is because some corruption inputs may be unused in some countries while widely used in others. Giving these different weights maximises the validity of the composite indicator while keeping measurement consistent across time and countries. As corruption techniques can substitute for each other, the different component weights reflect institutional features impacting on the form not the substance of institutionalised grand corruption (For details of comparative CRI see Fazekas, Chvalkovská, Skuhrovec, Tóth, & King, 2013).

Using the weights obtained from the measurement model, elementary indicators are simply summed to produce the corruption risk composite indicator of individual transactions. Summation reflects the view that any of the elementary corruption techniques is sufficient on its own to render a procedure corrupt; while multiple signs of corruption indicate higher corruption risks. Hence, we suggest the following formula for the composite indicator:

$$CRI^t = \Sigma_j\, w_j * CI_i^{\,t} \qquad (1)$$

$$\Sigma_j\, w_j = 1 \qquad (2)$$

$$0 \leq CRI^t \leq 1 \qquad (3)$$

$$0 \leq CI_j^t \leq 1 \qquad (4)$$

where CRI^t stands for the corruption risk index of transaction *t*, $CI_j^{\,t}$ represents the *j*th elementary corruption indicator observed in transaction *t*, and w_j represents the weight of elementary corruption indicator *j*. Elementary corruption indicators can be either corruption inputs or outputs.

Higher level units' such as organisations' CRI can be obtained by calculating the arithmetic average of their transactions' CRI in a given period (it is also possible to use contract values for weighting). The added value of aggregating CRI to a higher unit of observation such as an issuer of tenders is that it further increases our confidence in CRI. An organisation consistently displaying high CRI over time is likely to be actually a corrupt organisation rather than simply a victim of random fluctuations in the data.

4. Data

The database derives from Hungarian public procurement announcements of 2009-2012 (this database is referred to as PP henceforth). The data represent a complete database of all public procurement procedures conducted under Hungarian Public Procurement Law. PP contains variables appearing in 1) calls for tenders, 2) contract award notices, 3) contract modification notices, 4) contract completion announcements, and 5) administrative corrections notices. As not all of these kinds of announcements appear for each procedure, for example depending on procedure type, we only have the variables deriving from contract award notices consistently across every procedure. Comparable data sets exist or can be constructed from public records in all EU countries, the USA, and Russia for the last 6-8 years (Annex A with details).

The place of publication of these documents is the Public Procurement Bulletin which appears is accessible online[7]. As there is no readily available database, we used a crawler algorithm to capture the text of every announcement. Then, applying a complex automatic and manual text mining strategy, we created a structured database which contains variables with clear meaning and well-defined categories. As the original texts available online contain a range of errors, inconsistencies, and omissions, we applied several correction measures to arrive at a database of sufficient quality for scientific research. For a full description of database development, see Fazekas & Tóth (2012a) in Hungarian and in somewhat less detail Fazekas & Tóth (2012b) in English.

A potential limitation of our database is that it only contains information on public procurement procedures under the Hungarian Public Procurement Law as there is no central depository of other contracts. The law defines the minimum estimated contract value for its application depending on the type of announcing body and the kind of products or services to be procured (for example, from 1 January 2012, classical issuers have to follow the national regulations if they procure services for more than 8 million HUF or 27 thousand EUR). By implication, PP is a biased sample of total Hungarian public procurement of the period, containing only the larger and more heavily regulated cases. This bias makes PP well suited for studying more costly and more high stakes corruption where coverage is close to complete. Although, as removing contracts from the remit of the Public Procurement Law can in itself be part of corrupt strategies there remains some non-random bias in the data (for an estimation of this bias see (Fazekas, Tóth, et al., 2013b) and Figure 6 below).

As contract award notices represent the most important part of a procedure's life-cycle and they are published for each procedure under the Hungarian Public Procurement Law, their statistics are shown in Table 2 to give an overview of the database. It is noticeable that number and total value of contracts awarded has declined in the observation period. This is due to two parallel developments: 1) because of budget cuts since 2010, total public spending has declined; and 2) public procurement transparency has decreased since the new government entered office in 2010 (we will return to this point in section 6).

[7] See: http://www.kozbeszerzes.hu/nid/KE (in Hungarian)

Table 2. Main statistics of the analysed data – contracts

	2009	2010	2011	2012	**Total**
Total number of contracts awarded	10918	17914	14070	10342	53244
Total number of unique winners	3987	5617	5587	4923	13557
Total number of unique issuers	1718	2871	2808	2344	5519
Combined value of awarded contracts (million EUR) *	4604	3834	1856	1298	11592

Source: PP

*Notes: * = a 300 HUR/EUR uniform exchange rate was applied for exchanging HUF values.*

5. Building blocks: the corruption process' outcomes and inputs

5.1 Indicators of corruption outcomes

The key outcome of institutionalised corruption in public procurement, which we are measuring here, is contract performance by a pre-selected company. This corruption outcome can be secured at the procurement process'

1. Submission phase: only the pre-selected bidder submits a bid; or
2. Assessment phase: contract award to the pre-selected bidder;

As it is extremely rare that the company awarded a contract is changed during the delivery phase, the corruption outcome at the delivery phase[8] could be treated as fully determined by phases 1 and 2. Three outcome indicators are proposed to capture the full scale of institutionalised public procurement corruption where outcomes of any prior stage also serve as an inputs to later stages (Table 3). The corrupt outcome of the submission phase - only the pre-selected bidder submits a bid – is indicated by *whether a single bid was submitted to the tender*. In single submitted bid contracts, the issuer has an exceptionally large leeway to award the contract in a way which serves corrupt rent extraction. The corrupt outcome of the assessment phase - contract award to the pre-selected bidder – can only partially be captured by a quantitative indicator: *exclusion of all but one received bid*. Much of the award process such as scoring bidders is not extensively reported in public records hence the lack of further direct outcome indicators. In order to capture the final corruption outcome more appropriately, a further indicator is proposed which signals repeated contract award to the same company throughout multiple procedures: *winner's share of issuer's contracts* during the 12 month period before the contract award in question.

Table 3. Summary of outcome indicators

phase	indicator name	Definition
submission	single bidder	1=1 bid received, 0=more than 1 bid received
assessment	exclusion of bids	1=1 bid NOT excluded, 0=more than1 bid NOT excluded
overall	winner's share of issuer's contracts	12-month total contract value of winner / 12-month total awarded contract value (by issuer)

5.1.1 Single bidder

Issuers of tenders are free to choose the bidder of their preference; however, they are prescribed to maximise value for money, most importantly through soliciting competing bids. Corruption arises when competition is blocked in order to earn corruption rent. The most obvious signal that there was absolutely no competition for a public contract is when a tender received only 1 bid. Interview evidence from Hungary suggests that tenders with only 2-3 bids are also highly likely to be prone to corruption, as one public procurement adviser working in the industry for over a decade put it: "it is easy, just bring two friends with whom we can agree on the exact content of their bids". Focusing only on single bidder contracts is,

[8] If corruption is not institutionalised the delivery phase may well be the location of forming corrupt links. This, however, falls outside the remit of our measurement model.

therefore, a conservative approach in line with the goal of delivering a lower bound estimate of large-scale corruption.

There are two potential criticisms to this indicator: 1) The single bidder indicator also signals corruption in cases when there was truly only one bidder capable of performing the task, but no corruption took place. While this is a serious weakness of the indicator, it is considered to be only of marginal magnitude as the overwhelming majority of products procured by governments are ordinary and widely produced such as office stationery, cars, national roads, or IT support services (less than 5% of contracts were awarded on markets with 3 or fewer companies). In addition, robustness checks of our models, excluding markets with a small number of competitors, warrant that *this concern is of minor importance*. 2) Some authors contend that a single bidder has no incentive to give a bribe (Soreide, 2002). However, in an environment of systemic corruption, a single bidder tender is the ideal outcome created by colluding bidders and issuers, especially if the same single bidder wins contracts repeatedly (see section 5.1.3).

5.1.2 Exclusion of all but one bidder

It is possible that a corrupt issuer didn't manage to deter all but one bidder from submitting a bid, in which case it can still award the contract to the 'well-connected' bidder if it manages 1) to exclude the bids of all unwanted bidders on administrative or formal grounds (Heggstad & Froystad, 2011); or 2) to unfairly assess the bids to favour a particular bidder. As there is no direct evidence available in public records for the latter, the assessment phase's corruption outcome indicator captures only the former. Having a single valid bid tender can be heavily associated with corruption for, by and large, the same reasons as for single submitted bid (see section 5.1.1). Counter-arguments follow the same lines too. This similarity between the two measures, while conveying additional information, is also supported by regression results (Table 9).

5.1.3 Winner's share of issuer's contracts

While there is no separate indicator for the delivery phase, we develop a likely corruption outcome measure for the public procurement corruption process as a whole. The ultimate goal of large-scale institutionalised corruption is to repeatedly award contracts to the same company or companies controlled by the corrupt group (Heggstad & Froystad, 2011). By implication, winner's share of issuer's contracts indicates the likelihood of such corruption. As the primary location of collusion and capture is the individual public organisation disbursing public funds, this variable is defined as the ratio of contract value the winner won from a given issuer to the total value of contracts awarded by the given issuer throughout a 12-month period.

Using winner's share within issuer's contracts (or winner's contract share as we will call it to remain succinct) as corruption indicator is likely to suffer from disturbances in periods when a new dominant group takes control of public organisations with its new clientele, for example when a new government comes into office. Changes of dominant, captor groups are expected to be rare events, hence, this downward bias may only be moderate (and controlling for year of contract award in the below regressions captures much of this potential bias). Moreover, this indicator also underestimates corruption when the corrupt network uses multiple companies for extracting rents. Interviews indicate that combining company ownership groups' contract volumes accounts for most of this bias.[9]

Table 4. Descriptive statistics for the three outcome variables, 2009-2012, markets with at least 3 competitors

	mean	min	max	st. deviation	N
single received bid	0.30	0.00	1.00	0.46	51012
single valid bid	0.37	0.00	1.00	0.48	41277
winner's share of issuer's contracts	0.31	0.00	1.00	0.40	37399

Source: PP

5.2 Indicators of corruption inputs

According to our measurement model, the above outlined likely outcomes of the corruption process at least partially result from corruption techniques such as tailoring eligibility criteria to one company. These corruption techniques are interpreted as corruption inputs to the corruption process in public procurement which aims at purporting institutionalised grand corruption. A much wider set of corruption techniques in public procurement and their expected effects are extensively discussed in Fazekas et al. (2013)[10]. This section only provides a brief summary of 1) those factors which turned out to be powerful predictors in the below regressions in line with our prior expectations; and 2) of the theoretical expectations linking each input to each outcome.

14 input factors[11] are considered when building the models accounting for outcomes of the corruption process (variable definitions in Table 5, descriptive statistics in Table 6and Table 7). These capture key characteristics of the public procurement process from the beginning of the submission phase until the end of delivery.

[9] A further potential bias comes from collusion between bidding firms which tends to be based on product market rather than public organisation, hence it is deemed a relatively minor problem. An ongoing research project of the authors aims at separating corruption from cartel which is expected to deliver high quality evidence on this potential bias.
[10] Fazekas et al. (2013) discusses these indicators already applied to a group of contracts such as contracts awarded by an issuer over a period of time, while here they are interpreted on contract-level. This is only a formal difference without changing the logic of analysis.
[11] Note that single bidder contract is both an outcome of the submission phase as well as an input to the corruption process at later procurement stages.

Table 5. Summary of corruption inputs (higher score indicates greater likelihood of corruption)

phase	indicator name	indicator definition
submission	Single bidder contract	0=more than one bid received 1=ONE bid received
	Call for tender not published in official journal	0=call for tender published in official journal 1=NO call for tenders published in official journal
	Procedure type	0 =open procedure 1=invitation procedure 2=negotiation procedure 3=other procedures (e.g. competitive dialogue) 4=missing/erroneous procedure type
	Length of eligibility criteria	number of characters of the eligibility criteria MINUS average number of characters of the given market's eligibility criteria
	Length of submission period	number of days between publication of call for tenders and submission deadline
	Relative price of tender documentation	price of tender documentation DIVIDED BY contract value
	Call for tenders modification	0=call for tenders NOT modified 1=call for tenders modified
assessment	Exclusion of all but one bid	0=at least two bids NOT excluded 1=all but one bid excluded
	Weight of non-price evaluation criteria	proportion of NON-price related evaluation criteria within all criteria
	Annulled procedure re-launched subsequently*	0=contract awarded in a NON-annulled procedure 1=contract awarded in procedure annulled, but re-launched
	Length of decision period	number of working days between submission deadline and announcing contract award
delivery	Contract modification	0=contract NOT modified during delivery 1=contract modified during delivery
	Contract lengthening	relative contract extension (days of extension/days of contract length)
	Contract value increase	relative contract price increase (change in contract value/original, contracted contract value)

* Combining annulations by the issuer and the courts

Following from the discussion in (Fazekas, Tóth, et al., 2013b) specific expectations are formulated linking each input to each output (Table 8). Single received bid and single valid bid outcomes are discussed jointly because the theoretical considerations are very similar and the regressions unravel largely the same findings.

The expectations are formulated in a general linear form, for example, the shorter the submission period is the more likely that only one bid was received. However, many of the continuous variables are indeed not a continuous measure of corruption risks, rather there are critical thresholds beyond which corruption risks greatly increase. For example, a submission period of 5 days compared to 15 days is likely to convey higher corruption risks while a submission period of 35 days compared to 45 days may carry little to no information regarding corruption. By implication, behind any of our linear hypotheses lies the expectation of finding the thresholds which best capture spikes in the probability of a corruption outcome hence corruption risks.

In every case, the input variables are defined in a way that their higher values are expected to signal higher corruption risks. However, some of the corruption inputs are typically used as 'corrective action' later on in the procurement process to fix the failed attempts at bending competition earlier. These factors are expected to have negative association with corruption outcomes of earlier stages. For example, if only the well-connected company submitted a bid there is no need for subsequently modifying the contract as the corrupt bidder could set the price and quality allowing for corrupt rent extraction. However, if there was real competition at the submission phase the well-connected bidder is likely to be forced to submit a competitive bid with little scope for earning extra profit; hence the need for subsequent contract modification.

Table 6. Descriptive statistics of corruption inputs, 2009-2012, markets with at least 3 unique winners

	mean	min	max	sd	N
Single bidder contract	0.301	0.00	1.00	0.46	51012
Exclusion of all but one bid	0.367	0.00	1.00	0.48	41277
Call for tender not published in official journal	0.388	0.00	1.00	0.49	51823
Length of submission period	10.842	-7594.84	21594.88	3266.15	29215
Relative price of tender documentation	0.003	0.00	0.20	0.01	16743
Call for tenders modification	0.109	0.00	1.00	0.31	31726
Annulled procedure re-launched subsequently	0.061	0.00	1.00	0.24	55217
Weight of non-price evaluation criteria	0.216	0.00	1.00	0.33	51823
Length of decision period	90.871	0.00	1004.00	120.24	28605
Contract modification	0.189	0.00	1.00	0.39	51823
Contract lengthening	0.014	-0.97	30.29	0.26	16238
Contract value increase	0.079	-0.80	5.00	0.53	6547

Source: PP

Table 7. Distribution of procedure type, 2009-2012, markets with at least 3 unique winners

	N	%
open	31,007	59.83
invitation	906	1.75
negotiation	9,510	18.35
other	5,760	11.11
missing/error	4,640	8.95
Total	51,823	100

Source: PP

Table 8. Summary of the expected direction of and grounds for the relationships between corruption inputs and outputs

Phase	INPUT/OUTPUT	single received / valid bid		winner's share within issuer's contracts	
		direction	reason	direction	reason
Submission	Single bidder contract	**not relevant**	not relevant	+	Single received bid contracts make it easier for issuers to repeatedly award contracts to the same well-connected company.
	Call for tender not published in official journal	+	Not publishing the call for tenders in the official journal makes it less likely that eligible bidders notice the bidding opportunity and bid.	+	Not publishing the call for tenders in the official journal weakens competition allowing the issuer to more easily award contracts repeatedly to a well-connected company.
	Procedure type	+	Non-open procedures, which are less transparent and require less open competition, create more opportunities to limit the range of bids received and to exclude bids.	+	Non-open procedures, which are less transparent and require less open competition, create more opportunities for issuers to repeatedly award contracts to the same well-connected company.
	Length of eligibility criteria	+	Lengthy, hence complex, eligibility criteria allows issuers to tailor the tender to a single company and to exclude unwanted bids.	+	Lengthy, hence complex, eligibility criteria allows issuers to benefit a well-connected company, for example by keeping less competitive bidders in competition.
	Exceptionally short submission period	+	A short submission period leaves less time hence make it harder for non-connected companies to bid and to submit a bid.	+	A short submission period leaves less time hence make it harder for non-connected companies to bid successfully whereas a well-connected firm can use its inside knowledge to win repeatedly.
	Relative price of documentation	+	Relatively expensive tender documentation makes bidding more expensive and hence deters bidders from bidding except for the well-connected company which is close to certain of its success.	+	Relatively pricey tender documentation weakens competition allowing the issuer to more easily award contracts repeatedly to a well-connected company.
	Call for tenders modification	+	Modifying call for tenders allows for excluding unwanted bidders by changing eligibility criteria once the interested bidders are known.	+	Strategic modification of the call for tenders favours the well-connected company further increasing its market share.
Assessment	Exclusion of all but one bid	**not relevant**	not relevant	+	Single valid bid contracts make it easier for issuers to repeatedly award contracts to the same well-connected company.
	Weight of non-price evaluation criteria	+	Non-price related evaluation criteria tend to be more subjective, allowing issuers to favour the well-connected company. Apparently unfair assessment criteria deters bidders.	+	Non-price related evaluation criteria tend to be more subjective, allowing issuers to favour the well-connected company, hence repeatedly awarding contracts to the same company.
	Annulled procedure re-launched subsequently*	-	If unwanted bidders couldn't be deterred from bidding and their bids couldn't be excluded, annulling and subsequently re-launching the tender allows issuer to correct its failed attempt to eliminate competition.	+	If unwanted bidders couldn't be deterred from bidding and their bids couldn't be excluded, annulling and subsequently re-launching the tender allows issuer to more successfully award the contract to a well-connected company.
	Length of decision period	+	Overly lengthy decision period signals extensive legal challenges to the tender, suggesting that the issuer attempted to limit competition.	+	Lengthy decision periods signal extensive legal challenge to the tender, suggesting that the issuer wants to award the contract to a well-connected company.
Delivery	Contract modification	-	If competition couldn't be eliminated, the well-connected firm can still win with a competitive offer, but subsequent contract modification(s) still allow it to collect extra profit.	+	Contract modification(s) suggests that the issuer corruptly favour a well-connected company, potentially repeatedly.
	Contract lengthening	-	If competition couldn't be eliminated, the well-connected firm can still win with a competitive offer, but subsequent contract lengthening still allows it to collect extra profit.	+	A contract lengthening suggests that the issuer corruptly favour a well-connected company, potentially repeatedly.
	Contract value increase	-	If competition couldn't be eliminated, the well-connected firm can still win with a competitive offer, but subsequent contract value increase still allows it to collect extra profit.	+	A contract value increase suggests that the issuer corruptly favour a well-connected company, potentially repeatedly.

6. Composite corruption risk index

This section discusses 1) the regressions modelling institutionalised grand corruption in public procurement, 2) derives component weights for composite indicator building, and 3) provides validity tests for the resulting composite indicator.

The regressions' primary purpose is to validate whether corruption inputs could contribute to outputs in line with our theoretical expectations reflecting institutionalised grand corruption on the procurement market. They provide the primary source of internal validity of the composite indicator. As different phases of the procurement process are intertwined with each other, most appropriate analytical technique would be Structural Equation Modelling (Hoyle, 2012). However, this technique cannot easily handle large numbers of binary variables among dependent and independent variables and many non-linear relationships, hence, we opted for modelling each stage separately, but using partially overlapping variable sets. For outcomes single received bid and single valid bid, we used binary logistic regression; while for the winner's contract share outcome, we used linear regression.

In any regression, a significant and large coefficient is interpreted as indicating that the given corruption input is typically used for reaching the corruption output even after taking into account alternative explanations, such as contract size or length, and all other corruption inputs. This still means that it can be used for other, non-corrupt purposes in atypical cases; conversely, all the non-significant and weak explanatory factors may still be used for corrupt purposes, albeit only exceptionally.

Component weights of the composite indicator are derived from regression coefficients; whereby, the larger coefficient means higher component weight. This reflects the view that the more often a corruption input is used in combination with corruption outcomes the more confident we can be that institutionalised grand corruption lies behind its use.

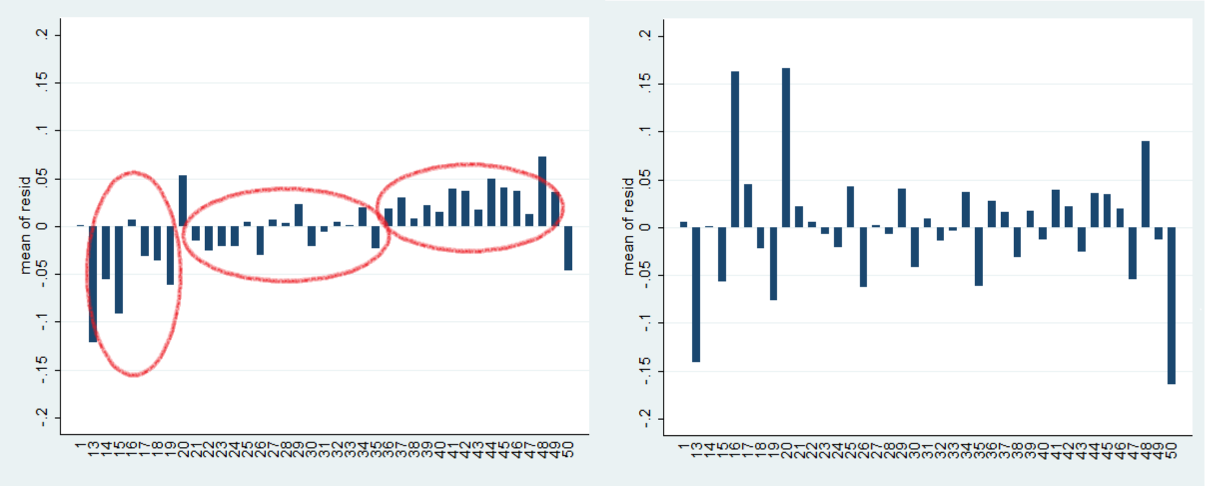
6.1 Modelling corrupt rent extraction: component weights

Regression models were built based on the above theoretical expectations by entering each corruption input and controls step-by-step. Here, only final regression results are reported for the sake of brevity. The regressions are fitted only one markets with at least 3 different winners in 2009-2012, that is where there is surely enough adequate competitors present. As the validity of all three outcome variables crucially hinges on the availability of suitable competitors, robustness checks are presented in Annex B excluding markets with less than 38 and 110 different winners throughout 2009-2012. The conclusions are substantially the same on the restricted samples too.

Thresholds in continuous variables were identified in an iterative process: first, a model was fitted using the linear continuous predictor; second, jumps in residual values were identified using residual distribution graphs. For example, average residual values of the regression using all the control variables plus the linear continuous measure of the relative price of documentation for predicting single received bid are depicted in Figure 1, left panel. It clearly indicates that there are three distinctive groups of relative document prices. For the lowest region, ranging between approximately the 24^{th} and 40^{th} percentiles, the model overestimates the probability of a single received bid, while it is the opposite case for the region between the 70th and 100th percentiles. These suggest at least three distinct categories. The right panel of Figure 1 shows the same residual distribution after the categorical measure of relative document price replaced its continuous version in the model with categories following the cut-points identified earlier. No clear pattern remains in the residual distribution, suggesting most non-linearity has been accounted for by the categorical measure of relative document price. A similar procedure was followed in the case of every continuous variable; if necessary completing multiple iterations of searching for thresholds.

In order to preserve the full population of observations, we always included a missing category in every categorical variable. In addition, this also helped measuring corruption inputs as concealing relevant tender information from bidders or the wider public often serves as a corruption technique.

Figure 1. Mean regression residuals by two-percentiles of relative price of documentation, left panel: linear prediction; right panel: prediction after taking into account non-linearity



Source: PP

When deciding on whether a variable is significant in the model, we used significance values from *Monte Carlo random permutation simulations* (Good, 2006), even though standard Fisher significance tests would have led to the same conclusions in most cases. This is because standard Fisherian significance tests are appropriate for statistical inference from a random sample to a population. However, our public procurement database contains the full population of interest, that is there is no sample. While some observations have been removed purposefully from the public domain hence from the database (a corruption risk on its own which is certainly far from random) this cannot be reflected by Fisher significance tests. Permutation tests are widely used in the natural as well as the social sciences, for example in social network analysis where data typically relates to full populations and observations are not independent of each other (Borgatti, Everett, & Johnson, 2013). The Monte Carlo random permutation simulation randomly reassigns the outcome variable to observations multiple times and calculates the regression coefficients each time. By doing so, it obtains a distribution of each regression coefficient when the outcome is truly random. The probability of the actual test statistic falling outside this random distribution, therefore, represents the probability of observing the relationship when the effect is truly random. A low significance level indicates that it is highly unlikely that the observed regression coefficient could be the result of a random process – a very intuitive interpretation.

Five different regressions are reported in Table 9, two binary logistic regressions on single received bid and two binary logistic regressions on single valid bid, following the same structure:

$$Pr(\text{single bidder}_i = 1) = \frac{1}{1+e^{-Z_i}} \quad (5)$$

$$Z_i = \beta_0 + \beta_{1j}S_{ij} + \beta_{2k}A_{ik} + \beta_{3l}D_{il} + \beta_{4m}C_{im} + \varepsilon_i \quad (6)$$

where single bidder$_i$ equals 1 if the *i*th contract awarded had only one bidder and 0 if it has more; Z_i represents the logit of a contract being a single bidder contract; β_0 is the constant of the regression; S_{ij} is the matrix of *j* corruption inputs of the submission phase for the *i*th contract such as length of submission period; A_{ik} stands for the matrix of *k* corruption inputs of the assessment phase for the *i*th contract such weight of non-price evaluation criteria; D_{il} stands for the matrix of *l* corruption inputs of the delivery phase for the *i*th contract such contract lengthening; C_{im} stands for the matrix of *m* control variables for the *i*th contract such as the number of competitors on the market; ε_i is the error term; and β_{1j}, β_{2k} , β_{3l}, and β_{4m} represent the vectors of coefficients for explanatory and control variables.

In addition to the four logistic regression models in Table 9, a linear regression on winner's share within issuer's contracts is reported following the structure:

$$Y_i = \beta_0 + \beta_{1j}S_{ij} + \beta_{2k}A_{ik} + \beta_{3l}D_{il} + \beta_{4m}C_{im} + \varepsilon_i \quad (7)$$

where Y_i represents winner's share within issuer's contracts; β_0 is the constant of the regression; S_{ij} is the matrix of *j* corruption inputs of the submission phase for the *i*th contract such as length of submission period; A_{ik} stands for the matrix of *k* corruption inputs of the assessment phase for the *i*th contract such weight of non-price evaluation criteria; D_{il} stands for the matrix of *l* corruption inputs of the delivery phase for the *i*th contract such contract lengthening; C_{im} stands for the matrix of *m* control variables for the *i*th contract such as the number of competitors on the market; ε_i is the error term; and β_{1j}, β_{2k} , β_{3l}, and β_{4m} represent the vectors of coefficients for explanatory and control variables.

The main differences among regressions are the outcome variables and whether the sample also includes withdrawn contracts (models 2 and 4). Withdrawn contracts couldn't be included in regressions on winner's share within issuer's contracts as they would have inflated contract values of 12 month periods. Each regression includes the full list of controls and predictors having non-missing values in the given sample. Control variables account for the most obvious alternative explanations to our corrupt outcomes:

- type of product procured using 40 different CPV[12] divisions which control for differences in technology and market standards;
- number of winners throughout 2009-2012 on the product market using a matrix of 820 CPV categories at level 3 and 4 geographical regions using NUTS[13] definitions

[12] CPV=Common Procurement Vocabulary. For more info see: http://simap.europa.eu/codes-and-nomenclatures/codes-cpv/codes-cpv_en.htm

[13] NUTS=Nomenclature of territorial units for statistics. For more info see: http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction

which makes sure that our findings on single bidders and winner's share within issuer's contracts are not driven by the low number of competitors available on the market.

- year of contracting which by and large proxies the changes in the legal framework and government in power;
- log real contract value (2009 constant prices) and contract length, both controlling for the differences emanating from contract size and complexity;
- whether the contract is a framework contract which have specific regulations and procedural rules[14]; and
- issuer type, sector, and status controlling for the regulatory as well as the institutional specificities of different issuers.

The regressions are performed on a restricted sample in order for the regressions to adequately fit a corrupt rent extraction logic as opposed to market specificities or inexperience with public procurement:

- markets with at least 3 unique winners throughout 2009-2012 for markets defined by a matrix of 820 CPV categories at level 3 and 4 geographical regions using NUTS definitions; and
- issuers awarding at least 3 contracts in the 12 months period prior to the contract award in question.

By and large, our hypotheses are supported by regressions, warranting the construction of a composite indicator reflecting systematically corrupt public procurement (Table 9).[15] First, the *single received or valid bid* is a powerful predictor of winner's share within issuer's contracts. Those contracts with a single bid tend to be awarded to winners with 1.8% higher share within issuer's contracts on average compared to contracts with more than one bids. This significant effect confirms that restricting the number of bids to one can support corrupt rent extraction on a recurrent basis. The magnitude of the impact is modest which is not surprising as restricting competition at the submission phase is only one of many ways to bent competition in public procurement.

Second, *not publishing the call for tenders in the official journal* increases the probability of single received and valid bids and the winner's contract share in every regression in line with expectations. For example, in model 1 and 3, it increases the average probability of a single received bid contract award by 14.8%-16.9% which is one of the strongest impact across models.

Third, every *non-open procedure type* carries a higher corruption risk than open procedures in terms of single received and valid bids and winner's contract share, supporting and further refining our theoretical expectations. Other, exceptional procedures carry the highest corruption risks adding 2.9% to winner's share within issuer's contracts compared to open procedures. Invitation and negotiation procedures are powerful and significant predictors in the regressions explaining single bidder contracts, but they have weak or counterintuitive impacts in the winner's contract share regressions which suggests that their main effect is

[14] For details see:? http://ec.europa.eu/internal_market/publicprocurement/docs/explan-notes/classic-dir-framework_en.pdf

[15] Of course, a number of further corruption inputs identified in Fazekas, Tóth, et al. (2013) are not presented here as they turned out to be either insignificant or too small.

likely to come through number of bidders. Invitation procedures appear to have about twice as strong effect on the probability of a single bidder contract award (7.1%-7.8%) as negotiation procedures (2.7%-5.9%).

Fourth, *relative length of eligibility criteria* behaves as expected with more lengthy, thus complex, criteria associated with higher probability of a single bidder contract and higher winner contract share. The effect of criteria length around the market average length seems weak, but positive indicating that there may be markets where complex criteria is frequently used to deter bidders. Criteria lengths considerably higher than market average are especially strongly associated with higher probability of single bidder contracts and higher winner contract share. For example, criteria length above market average by 520-2639 characters[16] increases probability of a single received bid by 10.4%-11.9% and the winner's share within issuer's contracts by 1.3% compared to the shortest criteria-length group. Interestingly, the call for tenders which are published, but don't contain eligibility criteria at the section where it is prescribed by law, are associated with especially high corruption risks: 9%-16% higher probability of single received bid contract compared to the shortest character length group. This signals that making eligibility criteria less visible deters bidders.

Fifth, the shorter the *submission period* the higher the probability of single received and valid bids and winner contract share in line with expectations. This relationship appears in distinct jumps around legally prescribed thresholds and the abuse of weekends. The exceptionally short submission period abusing weekends is one of the most powerful predictors in all of the models. It increases the winner's share within issuer's contracts by 7.6% and the probability of single valid bid by 17.2%-19.8%. Similar to criteria length, not displaying visibly and clearly the submission deadline is associated with very high corruption risks, for example 16%-24% higher probability of single received bid. As the effect is negligible on winner contract share, this corruption technique's impact arises primarily in the submission phase.

Sixth, more *expensive tender documents* increase both the probability of single bidder contracts and winner contract share in line with expectations. Compared to free documentation, document prices between 0.04%-0.1% of the contract value increase the probability of single received bid by 2.9%-3.4% and increase winner's share within issuer's contracts by 3.5%. Even more expensive tender documents have a stronger impact in the single bidder regressions, but insignificant and small effect in the winner contract share regression. This indicates that their main effect is exercised in the submission phase. The effect of the cheapest tender documentation is ambiguous across regressions. Missing tender documentation price is insignificant in most regressions. Therefore, these categories receive a zero weight in the composite indicator.

Seventh, call for tenders modifications behave according to expectations only for the period of the previous government (before 01/05/2010)[17], that is it increases the probability of single bidder contracts and the winner's market share. While it takes on a considerable significant negative coefficient under the current government' period. These differences signal the

[16] Standard deviation of character lengths from the population mean is 3435 for the whole 2009-2012 period. So, eligibility criteria 2639 characters above its market average is about three quarters standard deviation difference.

[17] Restricted sample results are not reported here. Regression outputs can be obtained from the authors.

changing role call for tenders modifications may play in corrupt rent extraction in response to changing regulatory (e.g. new Public Procurement Law entering into force soon after the new government entered into force) and political climate such judicial review of modifications (interviews indicate that the regulations and practice of judicial review of procurement tenders changed considerably after the new government entered office). Call for tenders modifications receive a positive weight in the composite indicator only for the pre-May 2012 period reflecting a conservative approach.

Table 9. Regression results on contract level, 2009-2012, average marginal effects reported for models 1-4 and unstandardized coefficients for model 5, nr. of winners >=3

models	1	2	3	4	5
Independent vars / dependent vars	single received bid	single received bid	single valid bid	single valid bid	winner's 12 month market share
single received/valid bid					0.018***
P(Fisher)					0.000
P(permute)					0.000
no call for tenders published in official journal	0.169***	0.14***	0.148***	0.121***	0.039***
P(Fisher)	0.000	0.000	0.000	0.001	0.040
P(permute)	0.000	0.000	0.000	0.000	0.000
procedure type					
ref. cat.=open procedure					
1=invitation procedure	0.078***	0.071***	0.069***	0.06***	-0.032*
P(Fisher)	0.126	0.122	0.301	0.308	0.259
P(permute)	0.000	0.000	0.000	0.000	0.015
2=negotiation procedure	0.027***	0.03***	0.059***	0.058***	0.009*
P(Fisher)	0.064	0.036	0.002	0.001	0.379
P(permute)	0.000	0.000	0.000	0.000	0.030
3=other procedures	0.275***	0.274***	0.257***	0.258***	0.029***
P(Fisher)	0.000	0.000	0.000	0.000	0.001
P(permute)	0.000	0.000	0.000	0.000	0.000
4=missing/erroneous procedure type	0.021**	0.028***	0.011	0.017	-0.008
P(Fisher)	0.134	0.049	0.484	0.270	0.256
P(permute)	0.000	0.000	0.140	0.055	0.155
length of eligibility criteria					
ref.cat.=length<-2922.125					
1= -2922.125<length<=520.7038	0.062***	0.046***	0.028*	0.019	0.001
P(Fisher)	0.009	0.044	0.328	0.505	0.942
P(permute)	0.000	0.000	0.015	0.065	0.895
2= 520.7038<length<=2639.729	0.119***	0.104***	0.07***	0.063***	0.013
P(Fisher)	0.000	0.000	0.026	0.041	0.427
P(permute)	0.000	0.000	0.000	0.000	0.110
3= 2639.729<length	0.138***	0.124***	0.077***	0.071***	0.014
P(Fisher)	0.000	0.000	0.021	0.035	0.418
P(permute)	0.000	0.000	0.000	0.000	0.105
4= missing length	0.16***	0.09***	0.05***	0.018***	0.048***
P(Fisher)	0.000	0.007	0.247	0.621	0.045
P(permute)	0.000	0.000	0.000	0.000	0.000
short submission period					
ref.cat.=normal submission period					
1=accelerated submission period	0.02***	0.022***	0.005	0.007	0.014***
P(Fisher)	0.067	0.051	0.715	0.581	0.028
P(permute)	0.000	0.000	0.525	0.335	0.000
2=exceptional submission period	0.086***	0.09***	0.076***	0.084***	0.047***
P(Fisher)	0.005	0.002	0.025	0.009	0.163
P(permute)	0.000	0.000	0.000	0.000	0.000
3=except. submission per. abusing weekend	0.189***	0.216***	0.172***	0.198***	0.076***
P(Fisher)	0.000	0.000	0.001	0.000	0.087
P(permute)	0.000	0.000	0.000	0.000	0.000
4=missing submission period	0.24***	0.16***	0.082***	0.028	-0.009
P(Fisher)	0.000	0.000	0.088	0.490	0.743
P(permute)	0.000	0.000	0.000	0.055	0.545
relative price of tender documentation					
ref.cat.= relative price=0					
1= 0<relative price<=0.0004014	-0.003	-0.01	-0.02	-0.042***	0.062***
P(Fisher)	0.902	0.598	0.371	0.060	0.001
P(permute)	0.860	0.360	0.130	0.000	0.000
2= 0.0004014<relative price<=0.0009966	0.034***	0.029**	0.016	-0.005	0.035***
P(Fisher)	0.095	0.128	0.419	0.796	0.013
P(permute)	0.000	0.005	0.225	0.715	0.000
3= 0.0009966<relative price<=0.0021097	0.035***	0.031***	0.027*	0.008	0.009
P(Fisher)	0.079	0.097	0.155	0.677	0.412
P(permute)	0.000	0.000	0.025	0.495	0.230
4= 0.0021097<relative price	0.058***	0.049***	0.03**	0.012	0.000
P(Fisher)	0.005	0.012	0.092	0.487	0.989
P(permute)	0.000	0.000	0.005	0.235	0.985

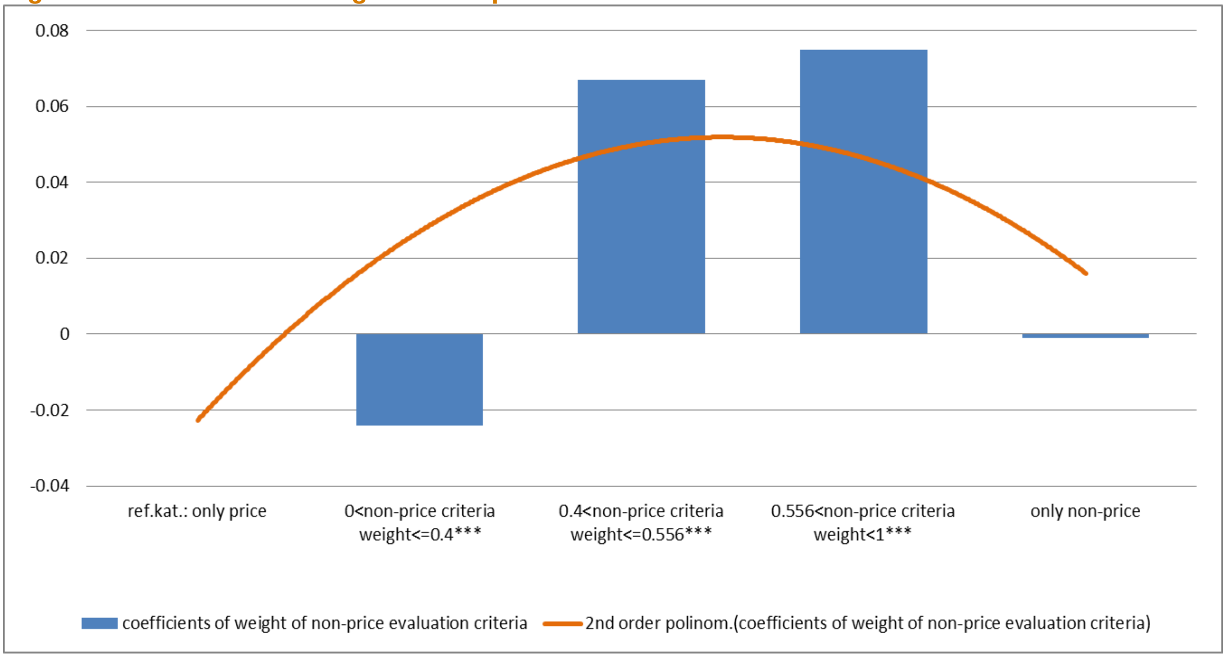
models	1	2	3	4	5
5=missing relative price	-0.011	-0.001	-0.004	-0.017	-0.008*
P(Fisher)	0.651	0.971	0.834	0.389	0.451
P(permute)	0.195	0.940	0.605	0.065	0.190
call for tenders modified	-0.032***	-0.036***	-0.043***	-0.043***	0.017***
P(Fisher)	0.059	0.029	0.039	0.033	0.032
P(permute)	0.000	0.000	0.000	0.000	0.000
weight of non-price evaluation criteria					
ref.cat.= only price					
2= 0<non-price criteria weight<=0.4	-0.024***	-0.019***	-0.043***	-0.034***	-0.002
P(Fisher)	0.053	0.121	0.004	0.019	0.782
P(permute)	0.000	0.000	0.000	0.000	0.705
3= 0.4<non-price criteria weight<=0.556	0.067***	0.069***	0.05***	0.05***	0.028***
P(Fisher)	0.000	0.000	0.004	0.002	0.006
P(permute)	0.000	0.000	0.000	0.000	0.000
4= 0.556<non-price criteria weight<1	0.075***	0.076***	0.078***	0.075***	0.038***
P(Fisher)	0.000	0.000	0.000	0.000	0.000
P(permute)	0.000	0.000	0.000	0.000	0.000
5=only non-price criteria	-0.001	0.001	-0.012	-0.012	0.007***
P(Fisher)	0.947	0.938	0.464	0.465	0.265
P(permute)	0.925	0.885	0.175	0.190	0.220
procedure annulled and re-launched		-0.112***		-0.031*	
P(Fisher)		0.000		0.357	
P(permute)		0.000		0.010	
length of decision period					
ref.cat.= 44<decision period<=182					
1= decision period<=32	0.085***	0.078***	0.121***	0.117***	0.013**
P(Fisher)	0.000	0.000	0.000	0.000	0.059
P(permute)	0.000	0.000	0.000	0.000	0.005
2= 32<decision period<=44	0.037***	0.032***	0.046***	0.047***	0.016***
P(Fisher)	0.002	0.004	0.001	0.000	0.028
P(permute)	0.000	0.000	0.000	0.000	0.000
4= 182<decision period	0.142***	0.147***	0.155***	0.161***	0.046***
P(Fisher)	0.000	0.000	0.001	0.001	0.002
P(permute)	0.000	0.000	0.000	0.000	0.000
5= missing decision period	-0.043***	-0.02	-0.036***	-0.016	0.022*
P(Fisher)	0.076	0.324	0.251	0.549	0.120
P(permute)	0.000	0.090	0.000	0.095	0.025
contract modified during delivery	-0.004	-0.004	-0.026***	-0.024***	0.015***
P(Fisher)	0.718	0.726	0.028	0.032	0.016
P(permute)	0.465	0.430	0.000	0.000	0.000
contract extension(length/value)					
ref.cat.=c.length diff.<=0 AND c.value diff.<=0.001					
2=0<c.length d.<=0.16 OR 0.001<c.value d.<=0.24	-0.064***	-0.061***	-0.02	-0.026	-0.01
P(Fisher)	0.000	0.001	0.359	0.204	0.405
P(permute)	0.000	0.000	0.175	0.060	0.355
3= 0.16<c. length diff. OR 0.24<c.value diff.	-0.008	-0.017	0.007	0.000	-0.006
P(Fisher)	0.701	0.373	0.753	0.986	0.550
P(permute)	0.580	0.125	0.675	0.985	0.450
4= missing (with contr. completion ann.)	-0.023**	-0.022**	-0.017*	-0.017*	-0.002
P(Fisher)	0.176	0.176	0.315	0.289	0.782
P(permute)	0.005	0.005	0.045	0.015	0.715
5= missing (NO contr. completion ann.)	-0.01	-0.011*	0.003	0.005	0.003
P(Fisher)	0.394	0.296	0.773	0.623	0.709
P(permute)	0.120	0.050	0.610	0.340	0.565
constant included in each regression; control variables: product market (cpv divisions); number of winners on the market (market defined by cpv level 4 & nuts 1) year of contract award; log real contract value; contract length; framework contract; issuer type, sector, and status (public or private)					
N	48853	52390	39309	42607	20653
R2/pseudo-R2	0.1038	0.0998	0.1022	0.0986	0.2433

*Source: PP; Note: * p<0.05; ** p<0.01; *** p<0.001; clustered standard errors clustered by issuer for P(Fisher), Monte Carlo random permutation simulations for P(permute) (200 permutations) using stata 12.0*

Eight, the effect of the *weight of non-price evaluation criteria* turned out to be somewhat different from expectations. Instead of a clearly positive relationship, we found an inverted U-shape relationship (Figure 2). This can be interpreted using our interview evidence:

stipulating only or predominantly price-related evaluation criteria warrants fair competition, hence, it is associated with lower corruption risks. While majority subjective criteria suggests rigged competition deterring bidders and increasing winner contract share. Only non-price evaluation criteria combined with fixed price is most likely complying with certain industry standards such as IT procurement without signalling heightened corruption risks (Fazekas, Tóth, et al., 2013b). Hence, only the two categories with positive coefficient receive non-zero weight in the composite indicator.

Figure 2. Effect sizes of weight of non-price evaluation criteria from model 1

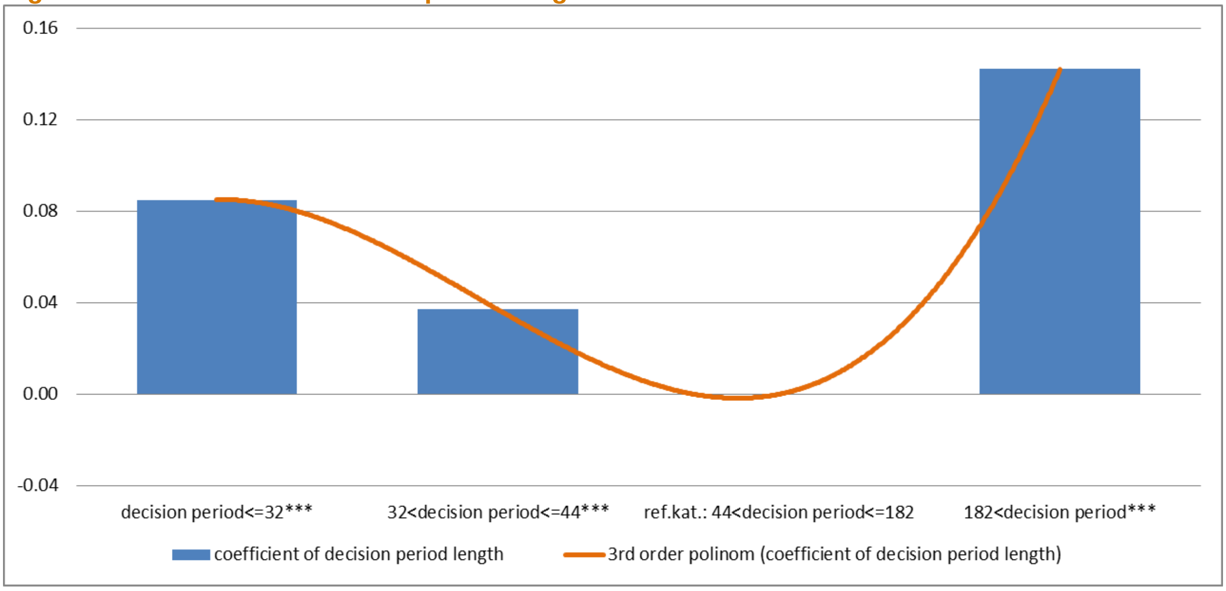


Source: PP
Note: p<0.05; ** p<0.01; *** p<0.001*

Ninth, *annulling and re-launching procedures* has the expected sign for both single received and single valid bid outcomes, but its effect cannot be determined on winner contract share due to technical complexities. Annulling a contract award is associated with 3.1%-11.2% *lower* probability of single bidder contract award, that is contract awards are annulled and re-launched more often when there were multiple bidders. This is completely contradictory to the prescriptions of the EU Public Procurement Directive or the Hungarian Public Procurement Law, but in line with a corrupt rent extraction logic.

Tenth, the effects of *decision period length* on probability of single bid and winner contract share are both somewhat different from our expectations. It seems that the relationship follows a U-shaped pattern with average decision period lengths (between 40th and 90th percentile) having the lowest corruption risk (Figure 3). Compared to this reference category, exceptionally long decision periods and exceptionally short decision periods are both associated with high corruption risks. Decision periods longer than 182 working days result in 14.2%-16.1% higher probability of single bid contract and 4.6% higher winner's share within issuer's contracts. Decision periods shorter than 32 working days are associated with 7.8%-12.1% higher probability of single bid contract and 1.3% higher winner contract share. Decision periods between 32 and 44 working days have a somewhat weaker effect than exceptionally short decision periods. These results suggest that there are two mechanisms at play. First, exceptionally short decision periods may indicate rushed through decisions and the corresponding high corruption risks. Second, exceptionally long decision periods may signal multiple legal challenges and troubled decision making hence high corruption risks. While the missing category is significant in some models, its effect is far from clear, thus, it cannot be included in the composite indicator.

Figure 3. Effect sizes of decision period length from model 1



Source: PP
Note: p<0.05; ** p<0.01; *** p<0.001*

Eleventh, *contract modification* has the expected relationships with probability of single bid and winner contract share albeit effect sizes are small in general and insignificant for model 1-2. Modifying contract at least once after contract award is associated with 2.4%-2.6% *lower* probability of single valid bid and 1.5% higher winner's share within issuer's contracts. This indicates that a competitive contract award procedure may necessitate contract modification to assure rent extraction.

Twelfth, *increasing contract length* and increasing the contract value after contract award had to be considered together due to low number of relevant observations. These two techniques can be combined in as much as they represent two parallel methods for increasing the profitability of a contract, that is making delivery cheaper by extending the completion deadline or making price higher by increasing contract value. Contract extension (length/value) display the expected relationships, but effects are insignificant for the winner contract share regression.

Compared to contracts which were performed within the timeframe of delivery and original contract price (less than 0.1% value increase), contracts with 0%-16.2% longer delivery period *or* 0.1%-24% higher contract value were associated with 6.1%-6.4% *lower* probability of single received bid. For contracts which were extended even more the effects are insignificant which may signal that excessive project overruns are more often due to non-corrupt reasons such as low state capacity. For contracts whose contract completion announcement didn't contain the prescribed final contract length or final contract value information the probability of single bid was 1.7%-2.3% lower which is a moderately strong impact. This suggests that competitive tendering makes it more necessary to hide the final total performance potentially not according to original contractual terms. Hence, contract extensions of moderate magnitude and missing information are included in the composite indicator.

Based on these regression results the variables and their categories could be selected which will make up the composite corruption risk index (CRI). First, all three corruption outcomes could be part of CRI because the regressions accounting for them are of adequate quality (i.e. formal tests of model appropriateness are affirmative). Second, as mentioned earlier, outcome variables get the weight of 1 reflecting their benchmark status. Qualitative evidence clearly underlines that any of the corruption inputs (i.e. corruption techniques) is sufficient on its own to render a procurement procedure corrupt. Therefore, each significant corruption input receives the weight of 1. In order to reflect coefficient sizes of categories in each corruption input, we ranked categories of each variable with the most impactful category receiving weight 1 and the others proportionately lower weights. For example, if there are four significant categories of a variable, then they would get weights 1, 0.75, 0.5, and 0.25. Finally, we normed each component weight so that the resulting composite indicator falls between 0 and 1 (Table 10). This was achieved in two steps: component weights were divided by the total number of components (N=13), then the resulting score was divided by its observed maximum (CRI[raw]=0.805). This rescaling assures that the minimum (maximum) of the score corresponds to the lowest (highest) corruption risks observed. The upper end of the scale may be too conservative as the combined presence of 3-4 corruption inputs and/or outputs (CRI=0.27-0.36) is already almost certainly very corrupt according to our interviewees[18].

[18] Calculating CRI for court decisions which established corruption in public procurement could serve as a more robust upper bound for the CRI scale. Further work is in progress.

Table 10. Component weights of CRI reflecting variable and category impact on corruption outcomes, normed to have an overall sum of 1

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
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	0.000
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)	
weight of non-price evaluation criteria	0.000
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

Source: PP

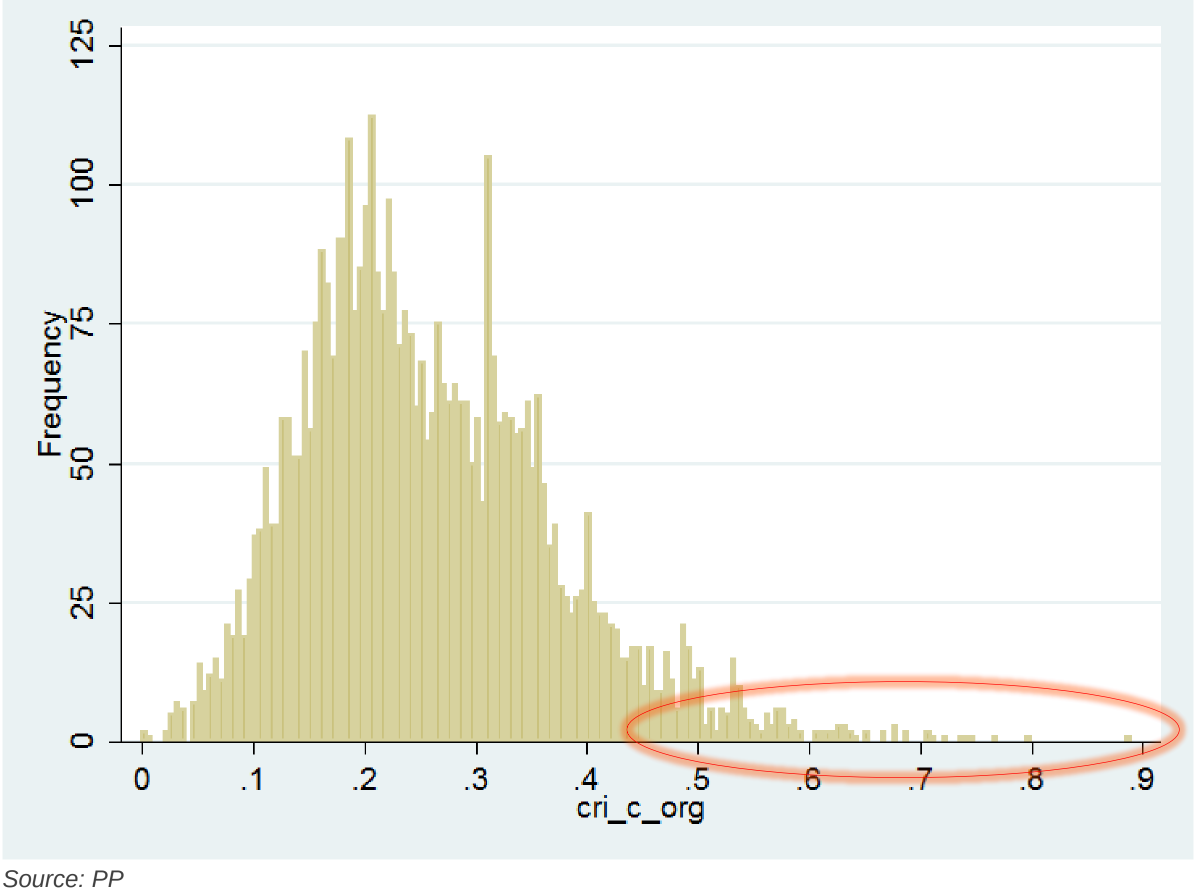
Note: If the call for tenders or contract fulfilment announcements are missing, the index is reweighted to only reflect the available variables (i.e. proportionately increasing the weight of observed variables).

6.2 Validating the corruption risk index

Validating CRI will take several years of work, here only elementary validating procedures are done. First, we look at the cross-sectional and time-series distribution of CRI to see if it behaves in any apparently unusual way. Second, the relationship between the amount of spending not reported in the PP database and CRI on the organisational level is explored to gauge the possible extent of distortion due to missing observations. Third, profitability and turnover growth of winning firms with different CRI are analysed. Fourth, political control of winning companies is collated with their CRI. Fifth, average CRI of companies whose market success seems to be strongly determined by the government in power is compared with those whose success is largely unaffected by government change (Fazekas, Tóth, et al., 2013a).

First, applying the weights specified in Table 10, each contract receives a corruption risk index (CRI) falling into a 0–1 band. Calculating the average CRI of each winning firm results in a CRI distribution which doesn't deviate extensively from a normal distribution, albeit it has a long tail to the right (Figure 4). These companies with CRI higher than approximately 0.4-0.5 represent particularly high corruption risks and hence deserve attention in later research.

Figure 4. Frequency distribution of winners according to CRI, 2009-2012[19], N=4430



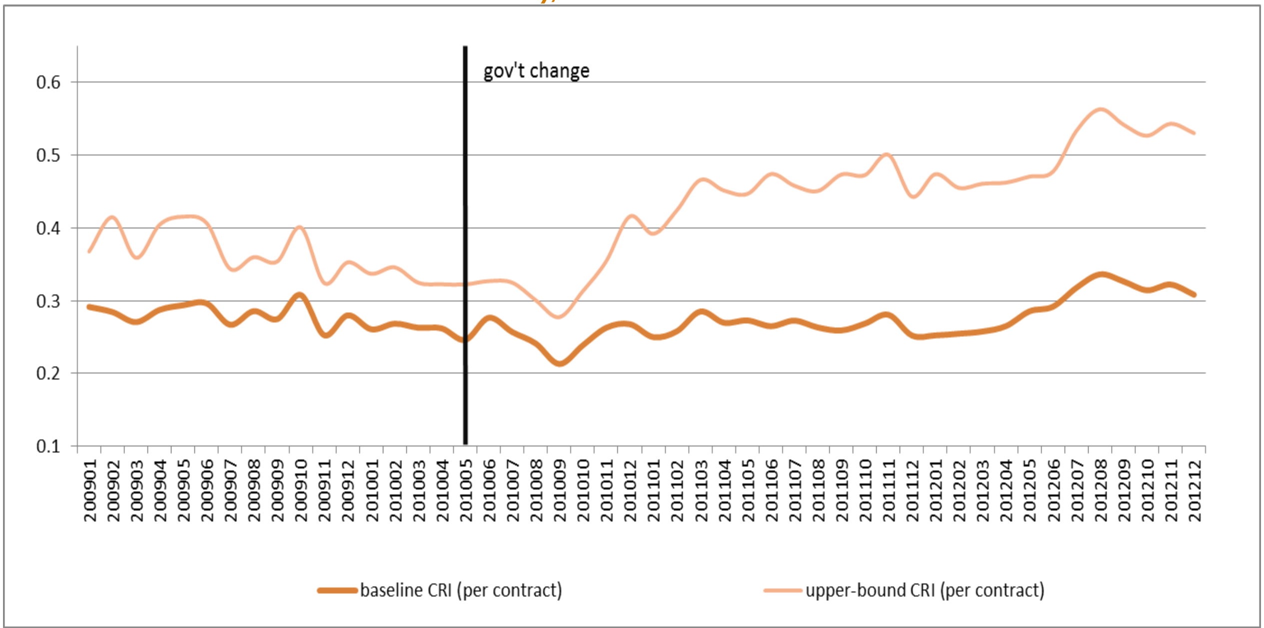
Source: PP

A simple test of indicator reliability is whether it displays any unexpected jumps at particular points in time or whether it reflects drastic changes known to impact on corruption. As CRI is

[19] In order to calculate CRI for 2009 where the 12-month values of winner's share within issuer's contracts is not available we had to input this variable using model 5 in Table 9.

defined for individual contract awards, monthly time series can be developed by calculating the CRI of the average contract. Such aggregation leads to a CRI time-series which is stable over time while showing some interesting variation from month to month (Figure 5). For example, it displays a spike just after the new government came into power which is primarily driven by contract modifications and longer decision periods. These are expected when dominant corrupt networks succeed each other and the newcomer tries to gain control of as many active sources of rent extraction as possible.

Figure 5. Monthly average CRI, 1/1/2009 – 31/12/2012 (averaging using the number and value of contracts awarded in each month), N=43642



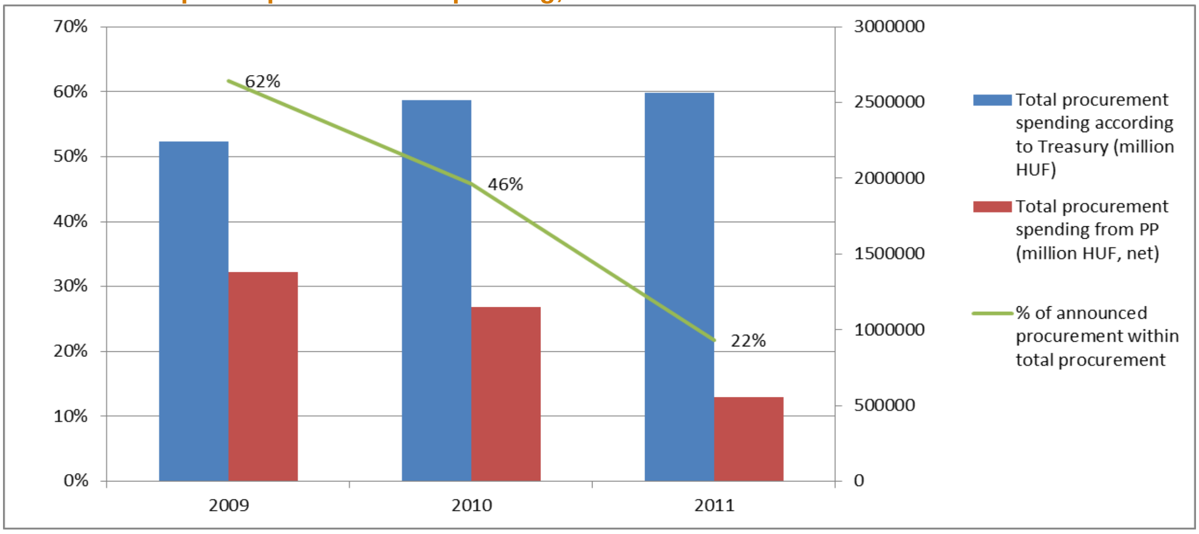
Source: PP

CRI declined between January 2009 and September 2010, but has increased since then which may provide hints at the performance of the new Fidesz government (Figure 5); although public procurement follows distinct cycles around elections hence comparisons are more appropriate at the same points in each cycle. Most interestingly, the Fidesz government has introduced a range of changes to the public procurement law which decreased transparency in at least three ways: 1) introducing less stringent requirements to publish call for tenders; 2) removing the requirement to publish contract fulfilment announcements; and 3) making it easier to move contracts outside the public procurement law for example by invoking national security concerns. Each of these can be tracked with our data creating an alternative estimate for CRI.

The baseline CRI is simply reweighted if call for tenders or contract fulfilment announcements are not available by relying on the available variables more extensively. However, as limiting transparency is a corruption technique confirmed by qualitative as well as quantitative evidence, it is reasonable to assume that the non-observed announcements are as risky as the highest corruption risk announcements observed. Under such a scenario, the starkly increasing corruption risks become visible after the Fidesz government takes power (Figure 5).

It is also possible to track the ratio of public procurement spending announced in the Public Procurement Bulletin to total public procurement spending (Figure 6). Since, the Fidesz government took power in 2010, this ratio has been cut by a half to reach only 22%. Once again, knowing that contracts awarded outside the remit of the Public Procurement Law represent higher corruption risks (for a detailed discussion see Fazekas, Tóth, et al., 2013b), it seems that corruption risks have increased between May 2010 and December 2012.

Figure 6. Public procurement spending announced in the Public Procurement Bulletin and total public procurement spending, 2009-2011

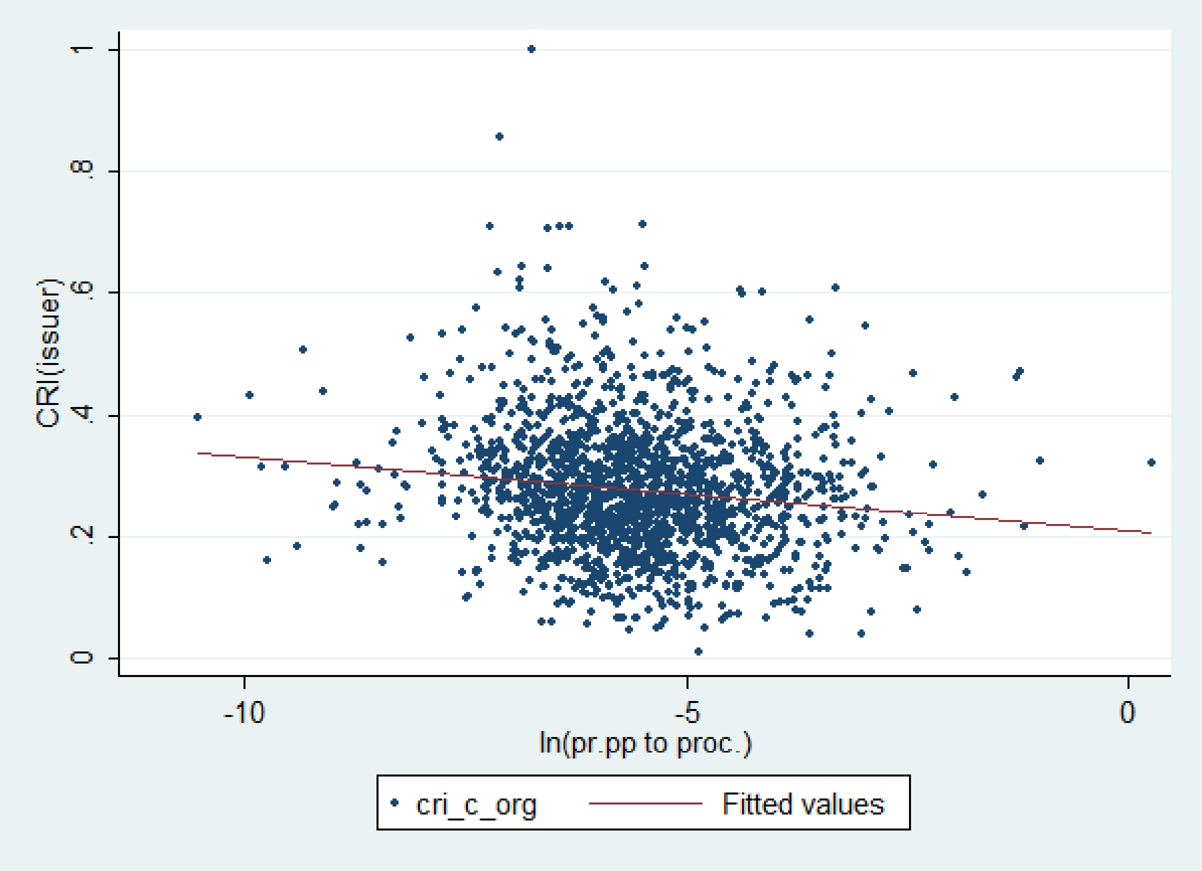


Source: PP

Notes: for details of calculating total procurement spending from Treasury annual budget accounts see: (Audet, 2002; European Commission, 2011b). The ratio reported is only an estimation as spending as announced in PP refers to the total lifetime of the contract while Treasury accounts contain only the spending accrued in a given year. Further reason for imprecision of the ratio is that the set of institutions submitting accounts to the Treasury and those subject to the Public Procurement Law are somewhat different.

Second, as qualitative evidence points out that removing contracts and procedures from the remit of the Public Procurement Law and hence the public domain is a corruption technique on its own, it is possible that the PP database is a biased sample of all the contracts and procedures relevant for analysing institutionalised grand corruption. It is possible to calculate the total estimated public procurement spending for each public organisation using Treasury data on individual organisations' annual budget breakdowns. By exploring the relationship between the amount of missing spending and average CRI per organisation, we get an insight into the potential bias due to missing data. The natural logarithm of the ratio of total procurement spending (Treasury records) to reported public procurement spending (Public Procurement Bulletin) is weakly negatively correlated with average organisational CRI (r^2=-0.12) (Figure 6Figure 7). This implies that the missing data bias is in line with our overall conservative approach of developing a lower bound estimate of institutionalised grand corruption, at least on the level of organisations. In addition, the overall weak relationship indicates that this bias is mostly due to random factors rather than systematic avoidance of transparency.

Figure 7. Issuer annual mean CRI and log total procurement to procurement reported in the Public Procurement Bulletin, 2009-2012, N=1717

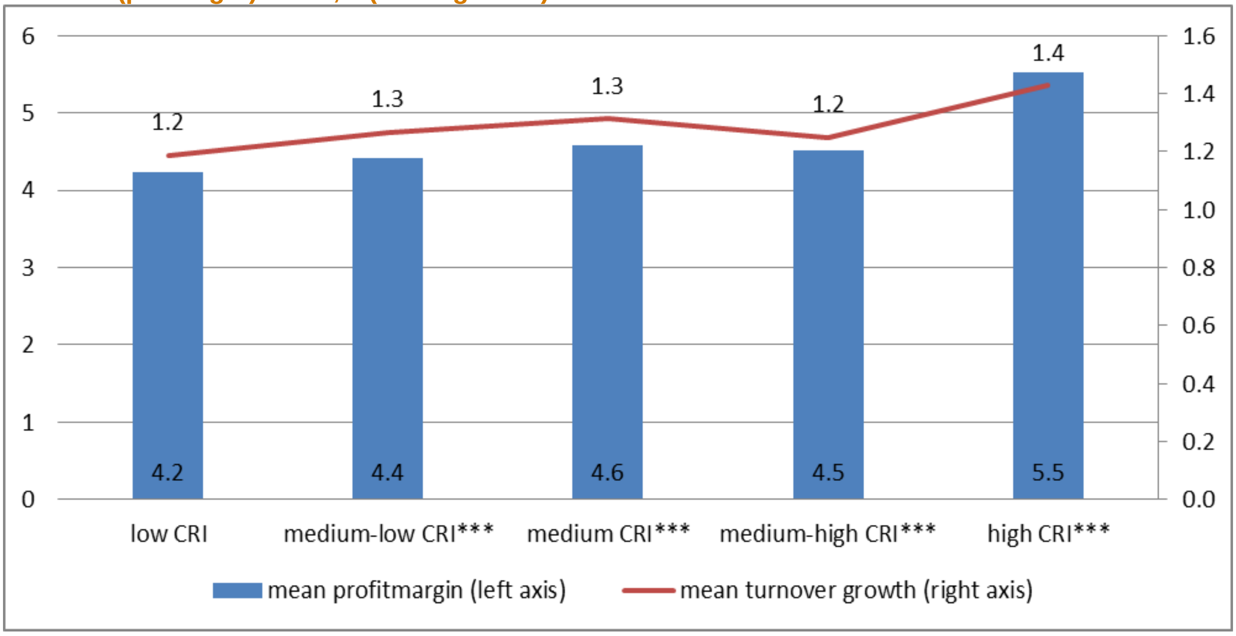


Source: PP

Third, we expect high CRI companies to earn higher profit and increase their turnover quicker than their low CRI peers because the primary aim of institutionalised grand corruption, which we are measuring with CRI, is to generate extra profit considerably above market average. However, we believe this relationship is likely to be only of moderate magnitude and probabilistic as high corruption companies are often hiding their profits and turnover through offshore companies, chains of subcontractors, and tax fraud. These have been confirmed by interviews in Hungary.

Simple comparisons of companies falling in the quintiles of CRI reveal a relationship in line with expectations (Figure 8). Percentile comparisons are preferable to simple correlations as corruption may have a non-linear effect on profitability and turnover growth (linear correlation coefficients are 0.04 and 0.02). Companies of highest CRI (0.35<CRI<0.87) are more profitable than any other company group, but the difference is especially large when compared to the lowest CRI companies (0<CRI<0.16): 1.3% points higher profit margin or 30% more profitable (1.3/4.4). Turnover growth, that is turnover in t_1 divided by turnover in t_0, is characterised by the same relationship with CRI. The highest CRI group has a 24% higher growth rate than the lowest CRI group. To some up, public procurement suppliers designated as high corruption risk companies by our corruption risk index are both more profitable and increase their turnover quicker than companies of the lowest corruption risk group. The fact that the relationship is particularly pronounced when comparing the two ends of the CRI distribution suggests that extremities of the CRI distribution may be the most precise in signalling institutionalised grand corruption.

Figure 8. Mean profit margin and mean turnover growth by CRI quintiles, 2009-2012, N (pr.margin)=3097; N(turno.growth)=2894



Source: PP

Note: $p<0.05$; ** $p<0.01$; *** $p<0.001$ designate the significance of the difference from the "low CRI" group. Significance levels computed using Monte-Carlo random permutations (300 repetitions) with stata*

Fourth, we expect that companies with political connections to display higher corruption risks as the primary vehicle for maintaining institutionalised grand corruption is to have strong ties between powerful political and business actors. We mapped the owners and manager of each company winning in 2009-2012 (15% of companies were either unidentifiable or we lacked the relevant data) and matched them with key political officeholders of public organisations existing in the period (for full list of institutions and offices see Annex C). The matching was done between more than 35000 owners/managers of winning firms and more than 10000 political officeholders based on full name[20]. Matching solely on name is obviously prone to random error which is nevertheless set aside for the present analysis by assuming that name frequency is not correlated with CRI. Those companies which have or had at least one owner or manager holding a political office at any point in time were designated as politically connected firms.

In line with our expectations, politically connected firms are of higher CRI (Table 11), they have a higher CRI by 0.01 on average than companies without political connections. While this difference is relatively small, increasing the precision of identifying political connections will shed more light at the validity of CRI. The magnitude of group differences may also signal that political connections serve as a means to corruption only in some cases while in others the politicians just picking profitable companies winning procurement contracts.

[20] Matching based on publicly available biographical data will be available in a later version of this paper.

Table 11. Comparisons of mean CRI of politically connected and not connected firms, 2009-2012

Group	N	Mean CRI	Std. Err.	Std. Dev.	95% Conf.Interval	
0=no political connection	2687	0.254	0.002	0.113	0.250	0.258
1=politically connected	1318	0.264	0.003	0.112	0.258	0.270
combined	4005	0.257	0.002	0.113	0.254	0.261
difference (CRI1-CRI0)		***0.010****	0.004		0.017	0.003

Source: PP

Note: p<0.05; ** p<0.01; *** p<0.001; Significance levels computed using Monte-Carlo random permutations (300 repetitions) with stata*

Fifth, it is possible to predict the total contract volumes of companies winning public procurement contracts between 2009-2012, and hence to identify those companies which win considerably more or less when the government changed in 2010 controlling for company characteristics such as prior investment and main market (Fazekas, Tóth, et al., 2013a). While more work is needed to reliably carry out this analysis, we expect that those companies whose market success highly depends on who is in power, i.e. latent political connections, display higher CRI. This is because institutionalised grand corruption is likely to be strongest where political connections are present. A simple comparison of the two groups' CRIs reveal a relationship in line with our expectations (Table 12). Companies with government dependent contract volume have 0.01 or 5% higher CRI than those whose contract volume is unaffected by which government is in power. While this difference is relatively small, it supports the claim that latent political connections translate into institutionalised grand corruption as measured by CRI.

Table 12. Comparisons of mean CRI[21] of companies whose market success does or does not depend on the which government is in power, 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

Source: PP

Note: p<0.05; ** p<0.005; *** p<0.001, Significance levels computed using Monte-Carlo random permutations (300 repetitions) with stata*

[21] Unlike in other validation tests, this test makes use of CRI aggregated by contract value rather than number of contracts. Hence, its meaning is closer to 'corruption risk index of the average HUF won' rather than average corruption risk index of the average contract won'. The reason for using contract value-based aggregation is that identification of companies as government-dependent is done using their contract volumes hence contract value aggregated CRI is more consistent with the company identification strategy. Findings are qualitatively the same with the alternative aggregation method.

7. Conclusions and the uses of the indicators

The analysis demonstrated that it is feasible and fruitful to construct a corruption risk index (CRI) at the micro-level based on objective behavioural data only. Initial evidence confirms the validity of CRI. The great advantage of our approach is that a large amount of data is available for research across every developed country for the last 6-8 years, opening up a new horizon for comparative corruption research. In addition, such comparative research will be able to use a conceptually much clearer concept whose measurement avoids the pitfalls of subjective indicators as well as prior objective indicators.

Almost every corruption input displayed a relationship with corruption outcomes in line with prior expectations (Table 13). Robust models linking corruption inputs to outputs allowed for deriving component weights for CRI composed of 14 variable groups neither of which dominates the resulting index (linear correlation coefficients between corruption inputs and CRI range between 0.01 and 0.57). The strength of this approach is that any change of regulation impacting on the relative costs of a corruption technique compared to other techniques leaves our CRI robust, as the increasing use of measured substitutive corruption techniques are adequately reflected. This characteristic of our CRI is particularly useful when comparing different countries of diverse regulatory environments and power constellations between elite groups. Further comparative work will use the same set of variables and regression setup in order to identify country- and period-specific parameters, as for example character-length of eligibility criteria tailored to a single company is likely to vary across countries and time with different regulatory institutions while the underlying institutionalised corruption may remain the same (Fazekas, Chvalkovská, et al., 2013).

Table 13. Summary of regression results

Phase	INPUT/OUTPUT	single received/ valid bid	winner market share
		empirical direction of relationship	
submission	Single bidder contract	not relevant	+
	Call for tenders not published in official journal	+	+
	Procedure type	+	+
	Length of eligibility criteria	+	+
	Exceptionally short submission period	+	+
	Relative price of documentation	+	+
	Call for tenders modification(only before 01/05/2010)	+	+
assessment	Exclusion of all but one bid	not relevant	+
	Weight of non-price evaluation criteria	∩	∩
	Annulled procedure re-launched subsequently*	-	not tested
	Length of decision period	U	U
delivery	Contract modification	-	+
	Contract extension (length/value)	-	0

Source: PP

We expect subsequent research to further validate CRI collating it to additional measures of grand corruption in more detail in Hungary and replicate measurement and analysis in other countries (work is ongoing for Czech Republic, Slovakia, Romania, and Russia).

Bibliography

Andersson, S., & Heywood, P. M. (2009). The Politics of Perception: Use and Abuse of Transparency International's Approach to Measuring Corruption. *Political Studies*, *57*, 746–767.

Arndt, C., & Oman, C. (2006). *Uses and Abuses of Governance Indicators*. Paris: OECD.

Audet, D. (2002). Government Procurement: A Synthesis Report. *OECD Journal on Budgeting*, *2*(3), 149–194.

Auriol, E., Flochel, T., & Straub, S. (2011). Public Procurement and Rent-Seeking: The Case of Paraguay. TSE Working Papers: 11-224, Toulouse: Toulouse School of Economics (TSE).

Bandiera, O., Prat, A., & Valletti, T. (2009). Active and Passive Waste in Government Spending: Evidence from a Policy Experiment. *American Economic Review*, *99*(4), 1278–1308.

Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). *Analyzing Social Networks*. London: SAGE Publications.

Charron, N., Dijkstra, L., & Lapuente, V. (2010). Mapping Quality of Government in the European Union: A Study of National and Sub-National Variation. Gothenburg.

Coviello, D., & Gagliarducci, S. (2010). Building Political Collusion: Evidence from Procurement Auctions. IZA DP No. 4939, Bonn: Institute for the Study of Labor (IZA).

Di Tella, R., & Schargrodsky, E. (2003). The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires. *Journal of Law and Economics*, *46*(1), 269–292.

European Commission. (2011a). *Establishing an EU Anti-corruption reporting mechanism for periodic assessment ("EU Anti-corruption Report")*. Brussels.

European Commission. (2011b). *Public Procurement Indicators 2010*. Brussels.

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

Fazekas, M., & Tóth, I. J. (2012a). *Hibák, javítások és előzetes eredmények - magyarországi közbeszerzések 2010-2011*. Corruption Research Centre, Budapest.

Fazekas, M., & Tóth, I. J. (2012b). *Public Procurement, Corruption and State Capacity in Hungary – objective measures and new insights*. Corruption Research Centre, Budapest.

Fazekas, M., Tóth, I. J., & King, L. P. (2013a). 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., Tóth, I. J., & King, L. P. (2013b). 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.

Ferraz, C., & Finan, F. (2008). Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. *The Quarterly Journal of Economics*, *123*(2), 703–745. doi:10.1162/qjec.2008.123.2.703

Golden, M. A., & Picci, L. (2005). Proposal for a New Measure of Corruption, illustrated with Italian data. *Economics & Politics*, *17*(1), 37–75. doi:10.1111/j.1468-0343.2005.00146.x

Goldman, E., Rocholl, J., & So, J. (2013). Politically Connected Boards of Directors and The Allocation of Procurement Contracts. *Review of Finance*, *17*(5), 1617–1648. doi:10.1093/rof/rfs039

Good, P. I. (2006). *Resampling Methods A Practical Guide to Data Analysis* (3rd editio.). Berlin: Birkhauser.

Heggstad, K. K., & Froystad, M. (2011). *The basics of integrity in procurement*. Bergen, Norway: U4 Anti-Corruption Resource Centre Chr. Michelsen Institute (CMI).

Hoyle, R. H. (Ed.). (2012). *Handbook of Structural Equation modeling*. New York: Guildford Press.

Hyytinen, A., Lundberg, S., & Toivanen, O. (2008). Politics and Procurement: Evidence from Cleaning Contracts. HECER Discussion paper No. 233.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2006). Measuring Corruption: Myths and realities. *Development*.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2007). Growth and Governance: A Reply. *The Journal of Politics*, *69*(2), 555–562.

Kaufmann, D., Mastruzzi, M., & Kraay, A. (2010). The Worldwide Governance Indicators. Methodology and Analytical Issues. Washington, DC.

Kaufmann, D., & Vincente, P. D. (2011). Legal Corruption. *Economics & Politics*, *23*(2), 195–219. doi:10.1111/j.1468-0343.2010.00377.x

Kurtz, M. J., & Schrank, A. (2007a). Growth and Governance: Models, Measures, and Mechanisms. *The Journal of Politics*, *69*(2), 538–554.

Kurtz, M. J., & Schrank, A. (2007b). Growth and Governance: A Defense. *Journal of Politics*, *69*(2), 563–569.

Lambsdorff, J. G. (2006). Measuring Corruption – The Validity and Precision of Subjective Indicators (CPI). In C. Sampford, A. Shacklock, C. Connors, & F. Galtun (Eds.), *Measuring Corruption* (pp. 81–100). Aldershot, UK: Ashgate.

Mungiu-Pippidi, A. (2006). Corruption: Diagnosis and Treatment. *Journal of Democracy*, *17*(3), 86–99.

Mungiu-Pippidi, A. (Ed.). (2011). *Contextual Choices in Fighting Corruption: Lessons Learned*. Oslo: Norwegian Agency for Development Cooperation.

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.

Olken, B. A. (2006). Corruption and the costs of redistribution: Micro evidence from Indonesia. *Journal of Public Economics*, *90*(4–5), 853–870.

Olken, B. A. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, *115*(2), 200–249.

Olken, B. A., & Pande, R. (2012). *Corruption in Developing Countries*. Boston, Massachusetts.

Radelet, S. (2002). *Qualifying for the Millennium Challenge Account*.

Radelet, S. (2003). Bush and Foreign Aid. *Foreign Affairs*, *82*(5), 104–117.

Reinikka, R., & Svensson, J. (2004). Local Capture: Evidence From a Central Government Transfer Program in Uganda. *The Quarterly Journal of Economics*, *119*(2), 678–704.

Rose, R., & Peiffer, C. (2012). Paying Bribes to Get Public Services: A Global Guide to Concepts and Survey Measures. Glasgow, UK: Centre for the Study of Public Policy.

Rothstein, B., & Teorell, J. (2008). What Is Quality of Government? A Theory of Impartial Government Institutions. *Governance*, *21*(2), 165–190.

Seligson, M. A. (2002). The impact of corruption on regime legitimacy: a comparative study of four Latin American countries. *Journal of Politics*, *64*, 408–433.

Soreide, T. (2002). *Corruption in public procurement Causes, consequences and cures*. Bergen, Norway.

Transparency International. (2012). *Gateway: Mapping The Corruption Assessment Landscape*. Berlin: Transparency International.

Annex A - Availability of public procurement data

Table 14. Overview of contract-level public procurement data availability in selected countries and regions, 2000-2012

Country	Data-source	Key online source	Minimum threshold (2012, classical issuer, services, EUR)[22]	Period	Availability
Czech Republic	Ministerstvo pro místní rozvoj ČR	http://www.isvzus.cz/usisvz/	39,000	2006-2012	structured data readily available and partially cleaned
EU	Tenders Electronic Daily	http://ted.europa.eu/	130,000	2005-2012	structured data partially available and cleaned
Germany	Bund.de- Verwaltung Online	http://www.bund.de/DE/Ausschreibungen/ausschreibungen_node.html	130,000[23]	2010-2012[24]	raw data available, not cleaned
Hungary	Közbeszerzési Értesítő	http://www.kozbeszerzes.hu/	27,300	2005-2012	structured data available and partially cleaned
Romania	eLicitatie	http://www.e-licitatie.ro/	30,000	2007-2012	raw data available, not cleaned
Russia	Goszakupki	www.zakupki.gov.ru	2,500	2006-2012[25]	structured data partially available and cleaned
Slovakia	Úrad pre verejné obstarávanie	http://tender.sme.sk/en/	30,000	2005-2012	structured data readily available and partially cleaned
UK	UK Contracts Finder	http://www.contractsfinder.businesslink.gov.uk/	11,600	2000-2012	raw data available, not cleaned
US	Federal Procurement Data System - Next Generation	https://www.fpds.gov/fpdsng_cms/	1,850	2004-2012	structured data readily available and partially cleaned

[22] National currencies are converted into EUR using official exchange rates of 5/2/2013 of the European Central Bank.
[23] It was increased from 30,000 EUR during the economic crisis.
[24] Earlier data have to be requested from the relevant bodies.
[25] 2006-2010 only for some regions.

Annex B - Robustness checks

The most convincing alternative explanation to this paper's interpretation of regressions as models of corrupt contract award states that products and services bought by public agencies are highly specific. Therefore, both single bidder and high share of the winner within the issuer's contracts are driven by the lack of adequate suppliers rather than corruption. In order to control for this important confounding factor each regression contains the number of winners on the market throughout 2009-2012 as an explanatory factor. In addition, this annex reports regressions on restricted samples which include contracts for products and services procured on markets with more than 2, 9, and 37 winners in 2009-2012. The cut-points 2 and 37 were defined using the same technique of identifying thresholds in continuous variables as spelled out in section 6.1. The cut-point of 9 was added arbitrarily in order to display an intermediary value.

To define the number of adequate competitors on a market, an appropriate definition of market has to be found. We defined markets along two dimensions: 1) the nature of product or service procured, and 2) the geographical location of contract performance. CPV codes differentiate over 3000 products and services as detailed as eggs (03142500-3) or potatoes (03212100-1). While we aim at being conservative in market definition, such level of detail is surely excessive. Exploiting the hierarchical nature of CPV classification, level-4 categories were selected as suitable for market definition, because the distribution of winners throughout 2009-2012 suggested that there are a large number of markets with a fairly small winners. Contracts were awarded in 820 level-4 CPV categories such as crops, products of market gardening and horticulture (0311) or construction materials (4411). Even though Hungary is a relatively small country interviewees suggested that there may be geographical frontiers of markets. Hence, we used 3 NUTS-1 regions plus the whole country to define markets along a geographical dimension (national reach typically requires an extensive set of local offices warranting an effective market barrier). Taken together, these resulted in 820*4=3280 distinct markets.

To define how many suitable competitors a market has, we simply calculated the winners of each market throughout 2009-2012. This is a conservative estimate as bidders who never won, for example because they were too expensive, but submitted valid bids were not taken into account. As some companies may have gone bankrupt or been bought by others, this estimation strategy may also be somewhat upward biased; therefore in some regressions we excluded markets with very many competitors.

The below tables demonstrate the robustness of our models to excluding markets with specific products and services (**Hiba! A hivatkozási forrás nem található.**, Table 15, and Table 16). Each of the findings in these alternative specifications remain unchanged compared to the main regressions, while indicators of goodness of fit improve somewhat.

Table 15. Regression results on contract level, 2009-2012, average marginal effects reported for models 1-4 and unstandardized coefficients for model 5, nr. of winners >=38

models	1	2	3	4	5
Independent vars / dependent vars	single received bid	single received bid	single valid bid	single valid bid	winner's 12 month market share
single received/valid bid					0.027***
P(Fisher)					0.000
P(permute)					0.000
no call for tenders published in official journal	0.173***	0.131***	0.167***	0.128***	0.057***
P(Fisher)	0.000	0.000	0.000	0.002	0.000
P(permute)	0.000	0.000	0.000	0.000	0.000
procedure type					
ref. cat.=open procedure					
1=invitation procedure	0.065***	0.06***	0.067***	0.058***	-0.021
P(Fisher)	0.224	0.206	0.332	0.339	0.471
P(permute)	0.000	0.000	0.000	0.000	0.195
2=negotiation procedure	0.025***	0.03***	0.066***	0.063***	0.013
P(Fisher)	0.14	0.074	0.002	0.002	0.235
P(permute)	0.000	0.000	0.000	0.000	0.055
3=other procedures	0.305***	0.3***	0.282***	0.281***	0.031***
P(Fisher)	0.000	0.000	0.000	0.000	0.003
P(permute)	0.000	0.000	0.000	0.000	0.000
4=missing/erroneous procedure type	0.03**	0.039***	0.019	0.026***	-0.008
P(Fisher)	0.062	0.017	0.315	0.165	0.379
P(permute)	0.000	0.000	0.060	0.000	0.275
length of eligibility criteria					
ref.cat.=length<-2922.125					
1= -2922.125<length<=520.7038	0.054***	0.033***	0.02	0.009	0.014
P(Fisher)	0.067	0.227	0.556	0.784	0.233
P(permute)	0.000	0.000	0.105	0.420	0.175
2= 520.7038<length<=2639.729	0.125***	0.106***	0.079***	0.07***	0.022
P(Fisher)	0.000	0.001	0.031	0.052	0.114
P(permute)	0.000	0.000	0.000	0.000	0.070
3= 2639.729<length	0.135***	0.116***	0.079***	0.068***	0.025
P(Fisher)	0.000	0.001	0.049	0.087	0.106
P(permute)	0.000	0.000	0.000	0.000	0.025
4= missing length	0.151***	0.057***	0.03	-0.008***	0.041*
P(Fisher)	0.001	0.132	0.540	0.841	0.052
P(permute)	0.000	0.000	0.060	0.000	0.015
short submission period					
ref.cat.=normal submission period					
1=accelerated submission period	0.023***	0.025***	0.005	0.009	0.015***
P(Fisher)	0.048	0.028	0.719	0.530	0.045
P(permute)	0.000	0.000	0.515	0.260	0.010
2=exceptional submission period	0.08***	0.089***	0.047***	0.065***	0.012
P(Fisher)	0.028	0.006	0.265	0.090	0.514
P(permute)	0.000	0.000	0.020	0.000	0.500
3=except. submission per. abusing weekend	0.136***	0.193***	0.088*	0.153***	0.039
P(Fisher)	0.019	0.004	0.131	0.013	0.423
P(permute)	0.000	0.000	0.045	0.000	0.520
4=missing submission period	0.28***	0.163***	0.123***	0.047*	-0.014
P(Fisher)	0.000	0.000	0.030	0.308	0.641
P(permute)	0.000	0.000	0.000	0.015	0.495
relative price of tender documentation					
ref.cat.= relative price=0					
1= 0<relative price<=0.0004014	-0.003	-0.013	-0.019	-0.047***	0.056***
P(Fisher)	0.901	0.531	0.463	0.053	0.010
P(permute)	0.855	0.295	0.165	0.000	0.000
2= 0.0004014<relative price<=0.0009966	0.022	0.016	0.011	-0.019	0.038***
P(Fisher)	0.361	0.455	0.673	0.418	0.015
P(permute)	0.070	0.195	0.440	0.175	0.000
3= 0.0009966<relative price<=0.0021097	0.038***	0.031***	0.022	-0.005	0.012
P(Fisher)	0.121	0.135	0.346	0.839	0.388
P(permute)	0.000	0.005	0.120	0.720	0.245
4= 0.0021097<relative price	0.07***	0.055***	0.044***	0.015	0.003
P(Fisher)	0.005	0.009	0.055	0.482	0.803
P(permute)	0.000	0.000	0.000	0.160	0.765

models	1	2	3	4	5
5=missing relative price	-0.005	0.005	0.001	-0.02	-0.012*
P(Fisher)	0.856	0.828	0.983	0.416	0.304
P(permute)	0.565	0.620	0.970	0.065	0.180
call for tenders modified	-0.015	-0.02*	-0.013	-0.016	0.005
P(Fisher)	0.441	0.288	0.617	0.538	0.610
P(permute)	0.090	0.030	0.185	0.105	0.515
weight of non-price evaluation criteria					
ref.cat.= only price					
2= 0<non-price criteria weight<=0.4	0.002	0.005	-0.024***	-0.017***	-0.003
P(Fisher)	0.882	0.718	0.176	0.316	0.722
P(permute)	0.675	0.405	0.000	0.000	0.585
3= 0.4<non-price criteria weight<=0.556	0.091***	0.091***	0.071***	0.069***	0.047***
P(Fisher)	0.000	0.000	0.001	0.000	0.000
P(permute)	0.000	0.000	0.000	0.000	0.000
4= 0.556<non-price criteria weight<1	0.102***	0.102***	0.095***	0.086***	0.045***
P(Fisher)	0.000	0.000	0.000	0.000	0.000
P(permute)	0.000	0.000	0.000	0.000	0.000
5=only non-price criteria	-0.005	-0.002	-0.008	-0.009	0.001
P(Fisher)	0.711	0.900	0.672	0.615	0.893
P(permute)	0.530	0.840	0.520	0.360	0.865
procedure annulled and re-launched		-0.098***		-0.027*	
P(Fisher)		0.001		0.422	
P(permute)		0.000		0.035	
length of decision period					
ref.cat.= 44<decision period<=182					
1= decision period<=32	0.075***	0.067***	0.123***	0.119***	0.014*
P(Fisher)	0.000	0.000	0.000	0.000	0.110
P(permute)	0.000	0.000	0.000	0.000	0.020
2= 32<decision period<=44	0.03***	0.023***	0.04***	0.042***	0.021***
P(Fisher)	0.030	0.067	0.012	0.003	0.019
P(permute)	0.000	0.000	0.000	0.000	0.000
4= 182<decision period	0.133***	0.147***	0.179***	0.187***	0.05***
P(Fisher)	0.000	0.000	0.000	0.000	0.005
P(permute)	0.000	0.000	0.000	0.000	0.000
5= missing decision period	-0.057***	-0.024*	-0.053***	-0.022	0.032**
P(Fisher)	0.027	0.249	0.114	0.418	0.112
P(permute)	0.000	0.010	0.000	0.060	0.005
contract modified during delivery	-0.005	-0.003	-0.034***	-0.029***	0.023***
P(Fisher)	0.678	0.765	0.013	0.028	0.001
P(permute)	0.400	0.545	0.000	0.000	0.000
contract extension(length/value)					
ref.cat.=c.length diff.<=0 AND c.value diff.<=0.001					
2=0<c. length d.<=0.162 OR 0.001<c.value d.<=0.24	-0.069**	-0.063***	-0.017	-0.026	-0.011
P(Fisher)	0.000	0.000	0.524	0.269	0.445
P(permute)	0.005	0.000	0.400	0.110	0.475
3= 0.162<c. length diff. OR 0.24<c.value diff.	-0.005	-0.015	0.022	0.011	-0.008
P(Fisher)	0.842	0.468	0.367	0.605	0.523
P(permute)	0.735	0.335	0.220	0.520	0.575
4= missing (with contr. completion ann.)	-0.01	-0.008	-0.009	-0.007*	-0.001
P(Fisher)	0.549	0.634	0.655	0.707	0.883
P(permute)	0.190	0.340	0.260	0.395	0.825
5= missing (NO contr. completion ann.)	-0.01	-0.013*	0.005	0.007	0.005
P(Fisher)	0.412	0.252	0.712	0.594	0.582
P(permute)	0.100	0.030	0.480	0.255	0.380
constant included in each regression; control variables: product market (cpv divisions); number of winners on the market (market defined by cpv level 4 & nuts 1) year of contract award; log real contract value; contract length; framework contract; issuer type, sector, and status (public or private)					
N	33440	36977	27067	30365	13019
R2/pseudo-R2	0.1183	0.1101	0.1074	0.1024	0.2558

Source: PP

Note: p<0.05; ** p<0.01; *** p<0.001; clustered standard errors clustered by issuer for P(Fisher), Monte Carlo random permutation simulations for P(permute) (200 permutations) using stata*

Table 16. Regression results on contract level, 2009-2012, average marginal effects reported for models 1-4 and unstandardized coefficients for model 5, nr. of winners >=110

models	1	2	3	4	5
Independent vars / dependent vars	single received bid	single received bid	single valid bid	single valid bid	winner's 12 month market share
single received/valid bid					0.034***
P(Fisher)					0.000
P(permute)					0.000
no call for tenders published in official journal	0.201***	0.136***	0.18***	0.114***	0.032
P(Fisher)	0.000	0.001	0.000	0.010	0.150
P(permute)	0.000	0.000	0.000	0.000	0.085
procedure type					
ref. cat.=open procedure					
1=invitation procedure	0.066*	0.054***	0.071**	0.05**	-0.054*
P(Fisher)	0.276	0.304	0.350	0.451	0.196
P(permute)	0.010	0.000	0.005	0.010	0.030
2=negotiation procedure	0.019*	0.023**	0.06***	0.056***	0.032***
P(Fisher)	0.328	0.208	0.009	0.009	0.051
P(permute)	0.010	0.005	0.000	0.000	0.000
3=other procedures	0.314***	0.309***	0.29***	0.287***	0.037***
P(Fisher)	0.000	0.000	0.000	0.000	0.007
P(permute)	0.000	0.000	0.000	0.000	0.000
4=missing/erroneous procedure type	0.023**	0.037***	0.009	0.02	-0.004
P(Fisher)	0.235	0.062	0.685	0.376	0.741
P(permute)	0.010	0.000	0.410	0.080	0.660
length of eligibility criteria					
ref.cat.=length<-2922.125					
1= -2922.125<length<=520.7038	0.057***	0.029*	0.016	-0.004	0.008
P(Fisher)	0.081	0.345	0.620	0.896	0.565
P(permute)	0.000	0.015	0.215	0.785	0.605
2= 520.7038<length<=2639.729	0.122***	0.093***	0.075***	0.056***	0.02
P(Fisher)	0.001	0.006	0.038	0.121	0.247
P(permute)	0.000	0.000	0.000	0.000	0.195
3= 2639.729<length	0.136***	0.107***	0.078***	0.052**	0.027*
P(Fisher)	0.000	0.003	0.047	0.178	0.140
P(permute)	0.000	0.000	0.000	0.005	0.035
4= missing length	0.18***	0.039***	0.059**	-0.009***	0.018
P(Fisher)	0.001	0.325	0.276	0.829	0.527
P(permute)	0.000	0.000	0.005	0.000	0.380
short submission period					
ref.cat.=normal submission period					
1=accelerated submission period	0.021**	0.025***	0.001	0.006	0.014
P(Fisher)	0.116	0.062	0.966	0.715	0.177
P(permute)	0.010	0.000	0.955	0.605	0.060
2=exceptional submission period	0.064***	0.086***	0.025	0.062**	0.015
P(Fisher)	0.063	0.006	0.550	0.120	0.660
P(permute)	0.000	0.000	0.310	0.005	0.585
3=except. submission per. abusing weekend	0.122*	0.204***	0.073	0.169**	-0.027
P(Fisher)	0.067	0.008	0.255	0.016	0.501
P(permute)	0.010	0.000	0.150	0.005	0.765
4=missing submission period	0.316***	0.165***	0.157***	0.053*	0.004
P(Fisher)	0.000	0.001	0.007	0.273	0.907
P(permute)	0.000	0.000	0.000	0.010	0.885
relative price of tender documentation					
ref.cat.= relative price=0					
1= 0<relative price<=0.0004014	0.012	-0.007	-0.022	-0.063***	0.036
P(Fisher)	0.720	0.765	0.502	0.029	0.168
P(permute)	0.410	0.615	0.240	0.000	0.070
2= 0.0004014<relative price<=0.0009966	0.03*	0.014	0.003	-0.04*	0.022
P(Fisher)	0.349	0.555	0.934	0.146	0.269
P(permute)	0.025	0.255	0.895	0.015	0.140
3= 0.0009966<relative price<=0.0021097	0.048***	0.032*	0.01	-0.029	-0.004
P(Fisher)	0.123	0.193	0.717	0.258	0.834
P(permute)	0.000	0.020	0.580	0.070	0.735
4= 0.0021097<relative price	0.102***	0.069***	0.057***	0.009	-0.005
P(Fisher)	0.001	0.005	0.032	0.707	0.768
P(permute)	0.000	0.000	0.000	0.540	0.700
5=missing relative price	0.002	0.01	-0.011	-0.039***	-0.038***
P(Fisher)	0.965	0.687	0.717	0.146	0.033
P(permute)	0.850	0.305	0.405	0.000	0.000
call for tenders modified	-0.023*	-0.028***	-0.019	-0.02	0
P(Fisher)	0.211	0.118	0.489	0.456	0.989
P(permute)	0.025	0.000	0.125	0.095	0.990

models	1	2	3	4	5
weight of non-price evaluation criteria					
ref.cat.= only price					
2= 0<non-price criteria weight<=0.4	-0.013	-0.005	-0.047***	-0.031***	-0.008
P(Fisher)	0.433	0.729	0.017	0.087	0.456
P(permute)	0.085	0.425	0.000	0.000	0.270
3= 0.4<non-price criteria weight<=0.556	0.074***	0.077***	0.044***	0.048***	0.049***
P(Fisher)	0.001	0.000	0.043	0.017	0.007
P(permute)	0.000	0.000	0.000	0.000	0.000
4= 0.556<non-price criteria weight<1	0.124***	0.124***	0.112***	0.102***	0.077***
P(Fisher)	0.000	0.000	0.000	0.000	0.000
P(permute)	0.000	0.000	0.000	0.000	0.000
5=only non-price criteria	0.011	0.014	0.01	0.005	-0.004
P(Fisher)	0.486	0.355	0.631	0.795	0.751
P(permute)	0.310	0.115	0.525	0.675	0.720
procedure annulled and re-launched		-0.076***		-0.025	
P(Fisher)		0.007		0.445	
P(permute)		0.000		0.100	
length of decision period					
ref.cat.= 44<decision period<=182					
1= decision period<=32	0.03***	0.033***	0.084***	0.089***	0.005**
P(Fisher)	0.044	0.015	0.000	0.000	0.688
P(permute)	0.000	0.000	0.000	0.000	0.610
2= 32<decision period<=44	0.023*	0.019*	0.024*	0.03**	0.01
P(Fisher)	0.167	0.212	0.173	0.051	0.441
P(permute)	0.025	0.035	0.015	0.005	0.305
4= 182<decision period	0.116***	0.143***	0.138***	0.159***	0.055***
P(Fisher)	0.002	0.000	0.007	0.001	0.013
P(permute)	0.000	0.000	0.000	0.000	0.000
5= missing decision period	-0.082***	-0.035***	-0.084***	-0.038***	0.016
P(Fisher)	0.000	0.088	0.020	0.177	0.461
P(permute)	0.000	0.000	0.000	0.000	0.440
contract modified during delivery	0	0.001	-0.027***	-0.023**	0.022***
P(Fisher)	0.973	0.922	0.065	0.102	0.015
P(permute)	0.945	0.835	0.000	0.005	0.000
contract extension(length/value)					
ref.cat.=c.length diff.<=0 AND c.value diff.<=0.001					
2=0<c. length d.<=0.162 OR 0.001<c.value d.<=0.24	-0.052**	-0.048**	0.006	-0.01	-0.022
P(Fisher)	0.012	0.012	0.856	0.719	0.252
P(permute)	0.005	0.005	0.775	0.580	0.225
3= 0.162<c. length diff. OR 0.24<c.value diff.	-0.028	-0.035*	0.007	-0.005	-0.023
P(Fisher)	0.311	0.119	0.813	0.858	0.192
P(permute)	0.130	0.025	0.715	0.790	0.185
4= missing (with contr. completion ann.)	0.001	0.002	0.015	0.015	0
P(Fisher)	0.961	0.900	0.495	0.457	0.995
P(permute)	0.945	0.830	0.240	0.195	0.985
5= missing (NO contr. completion ann.)	-0.004	-0.009	0.011	0.011	-0.01
P(Fisher)	0.767	0.454	0.490	0.416	0.372
P(permute)	0.655	0.195	0.240	0.190	0.220
constant included in each regression; control variables: product market (cpv divisions); number of winners on the market (market defined by cpv level 4 & nuts 1) year of contract award; log real contract value; contract length; framework contract; issuer type, sector, and status (public or private)					
N	22276	25813	18273	21584	7806
R2/pseudo-R2	0.1442	0.1272	0.1274	0.1148	0.2448

Source: PP

Note: p<0.05; ** p<0.01; *** p<0.001; clustered standard errors clustered by issuer for P(Fisher), Monte Carlo random permutation simulations for P(permute) (200 permutations) using stata*

Annex C – List of political offices considered for political connection measurement

The full list of institutions and positions can be obtained from the data provider, the government owned MTI Hungarian News Agency, which maintains a database of the most significant political office holders of the country for more than 20 years.

For more information see: http://mkk.mti.hu/

Table 17. List of institutions and positions of the political office holder database, 2010-2011

Institution	Position
Ministries	minister, secretary of state, vice-secretary of state, ministerial councillor,
Constitutional court	members and leaders
County courts	president, vice- president
Supreme court	President, vice-president, spokesperson
Prosecutors' Office	Chief prosecutor, vice-chief prosecutor, spokesperson
Municipalities	Major, vice-major, notary
County governments (new “kormányhivatal” too)	president, vice-president, notary
Regional police	Chief
National police headquarters	Chief, vice-chief, spokesperson
Minority governments	president, vice-president, head of office head of secretary
National medical service	Chief doctor, chief pharmacist
National Healthcare Fund	Director, vice-director
Army headquarters	Marshal, Vice-marshal
Treasury	President, vice-president, head of finances
Tax Administration	President, vice-president, spokesperson
Office of the president	President of the state, heads of every bureau of the office
State Audit Office	President, vice-president, chief director, director of finances
Regional Development Councils	presidents, member of governing committee
Office of the parliament	Head of office, heads of offices
Ombudsmen offices	Ombudsmen, heads of offices
National headquarters of Prisons	National chief, national vice-chief,
Competition Authority	President, vice-president, head of secretary
Central statistical office	president, vice-president
Other regulatory agencies and background institutes	top-management (2-3 positions)