# The extent and cost of corruption in transport infrastructure. New evidence from Europe

Mihály Fazekas[1], Bence Tóth[2]

Accepted for publication in Transportation Research Part A: Policy and Practice. Vol. 113, July 2018, pp. 35–54

ABSTRACT

*Transport infrastructure provision from roads to waterways involves large amounts of public funds in very complex projects. It is hardly a surprise that all across Europe, but especially in high corruption risk countries, it is a primary target of corrupt elites. This article provides a state-of-the-art review of the literature on the cost of corruption and estimates the level of corruption risks and associated costs in European infrastructure development and maintenance in 2009-2014 using novel data on over 40,000 government contracts. Two forms of corruption costs are investigated in the empirical section: 1) distorting spending structure and project design, and 2) inflating prices. Findings indicate that corruption steers infrastructure spending towards high value as opposed to small value investment projects. It also inflates prices by 30-35% on average with largest excesses in high corruption risk regions. Contrary to perceptions, corruption risks in infrastructure are decoupled to a considerable extent from the national corruption environment. Source data and risk scores are made downloadable at digiwhist.eu/resources/data.*

**Keywords**: corruption, transport infrastructure, public procurement, Europe, spending structure, price effect

[1] Corresponding author, University of Cambridge, mf436@cam.ac.uk, Address: Flat 3, Aumbrey Building, 95 Eastway, London, E9 5HE, UK

[2] University College London, bence.toth.16@ucl.ac.uk

# 1. Introduction

Improving the coverage and quality of transport infrastructure is key for economic development as it is a core public good supporting economic transactions, hence growth in general. As building such infrastructure involves large amounts of public funds in highly complex projects which are only comprehensible to a few experts, it is a primary target of corrupt elites. This is amply evidenced by scandals and trials, perception surveys and increasingly objective proxies of corruption. For example, a series of corruption scandals have rocked Spanish politics since 2014 when the so-called Gurtel case or the Operación Punica led to the trial and imprisonment of over 90 politicians and businessmen on charges of mishandling government contracts, many of which linked to transportation infrastructure development (Charron, Dahlström, Fazekas, & Lapuente, 2017). Or take the numerous scandals in Italy involving government contracting for large infrastructure such as the infamous, never built Messina Straits Bridge[3]. Corruption in transport infrastructure delivery is hugely important as it not only can lead to large amounts wasted, but also compromise their intended beneficial effects. Transportation infrastructure investments such as roads, railways, airports or maritime infrastructure represent a large part of public spending. Based on OECD statistics[4], the investment and maintenance spending on transport infrastructure amounted to roughly 1% of the GDP in years between 1995 and 2013 in OECD countries. Corruption can also distort spending structure and project design leading to low quality provision and unnecessary investments (Kenny, 2006, 2009b; Mauro, 1997).

The challenge of corruption in transport infrastructure building and maintenance is typically investigated in the context of developing countries where the issues are perceived to be more wide-spread (Collier, Kirchberger, & Söderbrom, 2015). However, given the qualitative and quantitative evidence of corruption in infrastructure across the EU (Becker, Egger, & Ehrlich, 2012) and the region's increased emphasis on renewing its transport infrastructure (e.g. "Junker-plan"), it is timely to review what we know and provide new insights. In particular, this article aims (i) to give a comprehensive overview of the existing literature on transport infrastructure corruption across the globe, and (ii) to assess corruption risks and associated costs in transport infrastructure development across Europe using novel objective corruption proxies.

The contribution of this article is fourfold. First, it carefully lays out the three main theoretical impact mechanisms through which corruption imposes costs on infrastructure provision: spending structure, prices, and delivery time and quality. To our surprise, there is a considerable theoretical and empirical unity in the literature regarding the nature and amount of costs corruption poses to transport infrastructure development across the globe. Second, it uses a novel dataset comprising tens of thousands of transport infrastructure tenders published by national governments and the European Commission, recording every single government contract regulated by the EU Public Procurement Directives (DG GROWTH, 2015). This data is ideal for analysing corruption risks in transport infrastructure investment, as every large project has to go through a formal tendering process, leaving us with a close to complete database of large-scale infrastructure delivery all across Europe. Such an unprecedented detail of infrastructure delivery reveals considerable variation in terms of corruption risks both across and within EU countries. Third, corruption also impacts on spending structure with increased regional corruption risks being associated with increased spending on large projects. As larger projects allow for concentrating corrupt rents in the hands of a few, it is possible that the distorted spending structure serves corrupt goals rather than genuine investment needs. Fourth, increased corruption risks substantially inflate prices in Europe. New road construction projects' unit costs are increased in high corruption risk tenders by up to 30-35% compared to low corruption risk ones.

The article is organised as follows: first, we synthetize the existing literature on transport infrastructure-related corruption. Second, the data used is discussed with particular focus on pros and cons of using 'Big Data'. Third, the empirical analysis proceeds by highlighting simple comparisons across sub-sectors of transport infrastructure and European regions then systematically assessing the impact corruption has on prices and spending structure. Finally, conclusions are drawn.

# 2. Conceptual frame

## *2.1 What do we know currently about transport infrastructure and corruption?*

Unfortunately, the literature on transport infrastructure corruption is scattered.[5] While theoretical works deliver clear expectations, most research in this area had to rely on country-level perception-based indices which are particularly inadequate in this context as the general population or experts have very little direct experience

[3] http://www.independent.co.uk/news/world/europe/italys-85bn-bridge-to-nowhere-8317312.html
[4] http://stats.oecd.org/Index.aspx?DataSetCode=ITF_INV-MTN_DATA#
[5] Although, Le, Shan, Chan, & Hu (2014) provides an overview on recent corruption research dealing with the construction industry, the literature focusing especially on transport infrastructure is only partly discussed.

with transport infrastructure development and can observe corruption in it to a limited degree (Kenny, 2006; Olken, 2009). Nevertheless, based on the available theoretical and empirical evidence we can draw up the key expectations about the costs of corruption in transport infrastructure which we explore on EU data and a small number of benchmark estimations from diverse countries can also be enumerated.

Investment into transport infrastructure is thought to be of particularly high corruption risk based on perception data, high profile scandals, and theoretical considerations (Golden & Picci, 2005; Kenny, 2007). Infrastructure development tends to imply large, long-term and complex projects each of which are conducive to corruption. In large projects, even a small fraction of the investment value amounts to large corruption rents making them particularly attractive (Rose-Ackerman, 1999). In case of long term investments such as transport infrastructure, the situation is further complicated as the gains of corruption – e.g. through building in less/low quality material – are realized early on, while costs arise only later. Complex projects are characterised by high degrees of information asymmetry which makes it harder to detect misconduct in terms of inflated prices, inferior quality, or sluggish delivery (Golden & Picci, 2005; Kenny, 2007). In addition, complex projects can require highly specialised skills and capacities which give rise to monopoly power and pricing making the detection and punishment of misconduct even more difficult. However, the results of transport infrastructure projects in terms of bridges and roads are highly visible to voters and donor agencies and infrastructure failure has the potential to damage many lives leading to investigations of construction works delivered. These together should in principle curb corruption in transport infrastructure development at least to some degree. As some of these characteristics such as complexity are also present in other types of government contracting such as IT development or legal services, *it has to be investigated to what degree transport infrastructure delivery is prone to corruption risks compared to other sectors in Europe*. This is a fundamental question as there is no study comparing transport investment with other investment types in terms of corruption prevalence using objective data, and relying on perceptions and high-profile scandals may be misleading (Kenny, 2006; Olken, 2009).

Corruption can occur at any phase of the investment cycle inflicting different costs on societies and implying different mitigation strategies (Benitez, Estache, & Soreide, 2010; Kenny, 2006, 2009a). Strategic planning for new projects, the tendering process, or the contract implementation phase each is prone to corruption. Nevertheless, corruption in transport infrastructure provision can compromise public goals in at least three direct ways: 1) distorting spending structure and project design; 2) inflating prices for a given quality; and 3) contributing to delayed and low quality provision, in extreme cases non-completion. Each of these are reviewed briefly in order to focus to the subsequent empirical analysis.

First, corruption in transport infrastructure provision is likely to **distort spending structure**, in particular to bias public investment towards high value, high complexity investments into new infrastructure as opposed to spending on maintenance and operations. Spending distortions can have an alternative explanation as discussed by another strand of political economy literature focusing on temporal commitment problems (see e.g. Dixit, 1998). Voters may force governments to commit to inefficiently high value projects (i.e. new transport infrastructure instead of maintenance) so that the choice of future governments is restricted, even though it would not have been preferred ex post (Glazer, 1989). Interestingly, such long-term commitment devices may also serve clientelism in as much as large projects cement long term rents enjoyed by a favoured community even if political leadership changes.

This expected distortion is demonstrated by Tanzi & Davoodi (1997) who show that higher level of corruption in a country (measured by perception indices) is associated with increased public investment, but with lower expenditures on operations and maintenance. Similarly, Mauro (1998) shows, that country-level corruption is negatively associated with the share of education-related government expenditure in GDP, and this relationship is robust to a number of alternative explanations such as prior level of development. In an approach more closely associated to this article's empirical analysis, Lukács & Fazekas (2015) point out that while the highest value and highest corruption risk procurement tenders are in infrastructure provision in Hungary, the average corruption risk of the sector is not particularly high. Their findings suggest that at least some of our focus on corruption in infrastructure may be driven by salient cases rather than a solid understanding of the overall risk profile of sectors. Given these considerations and data limitations, *it has to be investigated to what degree corruption biases transport infrastructure spending towards high value projects in Europe*.

Second, corruption in transport infrastructure provision is likely to **increase prices** even when taking project specifications as given (summary of research findings are in Table 1). Price inflation can manifest itself in for example wages or material costs in the awarded contract or only later during contract implementation (European Court of Auditors, 2013). Although without connecting it to corruption risks per se, Flyvbjerg, Holm, & Buhl (2003) shows that price escalation (i.e. cost overrun) is indeed a common phenomenon in case of transport infrastructure projects both for road, rail and fixed links (bridges and tunnels), and although in a varying magnitude it is evidenced in developed and developing countries as well. Furthermore, Flyvbjerg, Holm, & Buhl (2004) also shows that price escalation is strongly affected by the length of the implementation phase, hence underlying the connection between costs and implementation period (see below).

Evidence from Italy contrasting data on cumulative investment into infrastructure and its available stock shows how region-level corruption in infrastructure positively correlates with the price of infrastructure even after controlling for input costs such as labour costs or construction material prices (Golden & Picci, 2005). Analysis of road prices in low and middle income countries using the World Bank's ROCKS dataset show that country-level corruption risks as measured by perception indices such as Transparency International's Corruption Perception Index or the World Governance Indicators Control of Corruption indicator increase prices between 7 and 11% (Collier, Kirchberger, & Söderbrom, 2015). While percentage price differences might appear as small, given the high value of many infrastructure projects, the absolute costs are high which is also demonstrated by research looking at absolute unit prices such as Alexeeva, Queiroz, & Ishihara (2011). In their approach which is more directly related to ours, they shows that an additional procurement red flag is associated with an average 91,000 – 100,000 USD/KM price increase on a unique sample of European and Central Asian road projects. These results are reached on a contract-level analysis controlling for geographical differences and economic development. This literature suggests that the *effect of corruption risk on infrastructure prices in terms of % increase or total EUR increase should be explored empirically in Europe*.

**Table 1.** Summary table of estimations for the cost of corruption in infrastructure

| Studies | Countries | Years | Products investigated | Corruption measure used | Cost estimate |
|---|---|---|---|---|---|
| (Alexeeva et al., 2011) | Albania, Armenia, Bosnia and Herzegovina, Bulgaria, Croatia, Estonia, Georgia, Kazakhstan, Macedonia, Poland, Romania, Serbia, Ukraine | 2000-2010 | road rehabilitation and reconstruction | Red flags, such as cost increases during implementation, time overruns, high proportion of prequalified companies not bidding etc. | One additional red flag (from 10) is associated with 91-100 thousand USD cost increase per km |
| (Pricewaterhouse Coopers, 2013) | 8 EU Member States | 2010 | road and rail construction; water and waste; urban and utility construction | Binary corruption indicator, based on a model prediction | In case of road and rail construction 1.9-2.9%, water and waste 1.8-2.5%, urban and utility construction 4.8-6.6% of the total procured value |
| (Collier et al., 2015) | 99 low and middle income countries included in the World Bank's ROCKS database | 1984-2008 | road development and road preservation works | Worldwide Governance Indicators – control of corruption; Transparency International's 2008 Corruption Perception Index | Countries above the median corruption level (WGI) have 15% higher unit costs; One point increase on a ten-point scale of the TI corruption index increases costs by 7% |
| (Olken, 2007) | Indonesia | 2003-2004 | road construction | Difference between the claimed expenditures and the independent engineer's estimate | Increasing monitoring from 4% to 100% of projects reduced average missing expenditure from 27.7% to 19.2% |
| (Kenny, 2007) | 22 transition countries | 1999-2000 | construction | Survey questions | The cost of securing government contracts is approximately 7% of the contract value |
| (Meduri & Annamalai, 2012) | India | 1996-2010 | road construction | Petty corruption index constructed by (TI, 2005) | 470 thousand USD increase per lane km due to one standard deviation increase of petty corruption |

Third, corruption in transport infrastructure can contribute to **delayed and low quality provision**, in extreme cases failed completion altogether (Lewis-Faupel, Neggers, Olken, & Pande, 2014). In EU funded road projects in 2000-2013, the average delay was 41% or 9 months of the initial estimate (European Court of Auditors, 2013). Significant time overruns (6 to 12 month) were also apparent in a range of Eastern European and Central Asian countries' road projects financed by the World Bank (Alexeeva, Queiroz, & Ishihara 2011). Similarly, in a wide set of African developing countries also experienced more than a year delay in contract implementation in their World Bank funded road projects (Alexeeva et al., 2008). Delayed provision and long implementation is also fertile ground for inflating costs as it is pointed out in Flyvbjerg et al. (2004). Although, time overruns are not straightforward indications of corruption (complex projects can have unforeseen complications), weak supervision and enforcement of the initial contracts give rise to corruption risks. While construction delays are easy to detect, assessing implementation quality is less straightforward (e.g. the effects are only visible after years). Gillanders (2013) – using World Bank survey data – shows that perceived corruption is significantly and positively related to the percentage of firms identifying transportation (i.e. its quality) as a major constraint both at the country and regional level.

Furthermore, the existing literature also shows, that there can be anomalies without rigorous monitoring of project implementation. Based on independent engineers' ex post estimates, an approximately 8% decrease was found in unexplained material costs due to increased monitoring in road projects in Indonesian villages (Olken, 2007). Despite the clearly relevant corruption risks at the implementation stage, *due to the lack of adequate data and the complexity of the quality-corruption relationship, the connection between corruption risks and delayed or low quality provision cannot be tested here*. While there is probably the scantest thorough research on project non-completions, they are likely to represent the most sever corruption harm and can become systemic at certain places[6]. The tragedy of unfinished infrastructure projects is that without full completion, most projects cannot be used at all.

While these different forms of direct corruption costs in transport infrastructure provision may occur jointly or substitute for each other, they are likely to carry different total social costs. If corruption only increases the price of infrastructure without impacting on project design, infrastructure quality, delivery time, or overall completion, total social cost is close to the direct cost. However, if corruption's direct impact goes beyond prices, additional indirect costs are likely inflicted on the society such as non-available infrastructure or unreliable provision. These issues we cannot discuss in details as their measurement is beyond the framework of this article.

### *2.2 Measuring corruption risks*

While there has been a lot of controversy over how to best define corruption (Johnston, 1996), we have adopted a definition which is specific to the domain of public procurement, backed up by legal principles and deeply rooted in the governance literature going back to the Weberian notion of impartial government (Rothstein & Teorell, 2008). In addition, this definition is also widely used by practitioners around the globe: In public procurement, the aim of institutionalised corruption is to steer the contract to the favoured bidder without detection in a recurrent and organised fashion (Fazekas & Tóth, 2014; World Bank, 2009). This requires at least two violations of principles of impartial distribution of public resources: 1) avoiding competition, by for example using unjustified sole sourcing or direct contract awards; and 2) favouring a certain bidder, by for example tailoring specifications, or sharing inside information. This definition of corruption focuses attention on restricted access to and unfair competition for public resources (Mungiu-Pippidi, 2015; North, Wallis, & Weingast, 2009). Restricted and unfair access then translates into higher prices, lower quality and quantity in order to generate corruption rents. Such rents may be extracted in the form of bribes, but it is more typical to channel rents through broker firms, subcontracts, offshore companies, and bogus consultancy contracts to name a few typical instruments. As public procurement and especially infrastructure delivery involves huge sums, the typical institutionalised corruption scenario involves elites from both the public and private sectors such as elected officials, high-level bureaucrats, and wealthy businessmen.

The measures of corruption risk used in this article directly stem from the above definition and follow work of the authors elsewhere discussed extensively both single country and cross-country analyses (Charron et al., 2017; Fazekas, Chvalkovská, Skuhrovec, Tóth, & King, 2014; Fazekas, Tóth, & King, 2016). We understand corruption risk indicators as metrics ranging between 0 and 1 whose values suggest the risk of corruption to occur, but without an exact empirically verified relationship between corrupt acts and the risk scores (hence we don't use the term corruption probability).

The measurement approach exploits the fact that for institutionalised corruption to work, procurement contracts have to be awarded recurrently to companies belonging to the corrupt network. This can only be

[6] E.g. as a recent documentary put it, "unfinished Italy" is probably the most important Italian architectural style since the 2nd World War till now". (http://www.unfinished-italy.com/)

achieved if legally prescribed rules of competition and openness are circumvented. By implication, it is possible to identify the input side of the corruption process, that is fixing the procedural rules for limiting competition, and also the output side of corruption, that is signs of limited competition. By measuring the degree of unfair restriction of competition in public procurement, a proxy indicator of corruption can be obtained.

First, the simplest and most conservative indication of restricted competition in line with the above definition is when only one bid is submitted to a tender on an otherwise competitive market (output side). While single bidding might also reflect non-corrupt behaviour such as contract renewal, its widespread presence over longer periods across many procuring bodies is more likely to signal systematic deviations from competitive norms. Hence, the **percentage of single-bidder contracts** awarded with all the contracts is the most straightforward corruption proxy we use.

Second, a more complex indicator of institutionalised corruption also incorporates characteristics of the tendering process that are in the hands of public officials who conduct the tender and suggests deliberate competition restriction (input side) (Fazekas, Tóth, & King, 2013). This composite indicator, which we call the **Corruption Risk Index (CRI)**, represents the risk of corrupt contract award in public procurement, and is constructed as a simple weighted average of individual red flags. CRI = 0 indicates minimum corruption risk while CRI=1 denotes maximum corruption risk observed. Based on qualitative interviews of corruption in the public procurement process, a review of the literature (OECD, 2007; Pricewaterhouse Coopers, 2013; World Bank, 2009), and regression analysis, we identified the components of the CRI in addition to single bidding:

1. A simple way to fix tenders is to **avoid the publication of the call for tenders** in the official public procurement journal as this would make it harder for competitors to prepare a bid. This is only considered in non-open procedures as in open procedures publication is mandatory (Lengwiler & Wolfstetter, 2006).
2. While open competition is relatively hard to avoid in some tendering procedure types such as open procedures, others such as accelerated negotiated or negotiated without competition procedures are by default much less competitive; hence, using **less open and transparent procedure types** can indicate the deliberate limitation of competition, that is corruption risks (Chong, Klien, & Saussier, 2015).
3. If the **advertisement period**, i.e. the number of days between publishing a tender and the submission deadline, is too short for preparing an adequate bid, especially for large tenders considered in this study, it can serve corrupt purposes; whereby the issuer informally tells the well-connected company about the opportunity well ahead (Piga, 2011).
4. Different types of **evaluation criteria** are prone to manipulation to different degrees, subjective, hard-to-quantify criteria often accompany rigged assessment procedures as it creates room for discretion and limits accountability mechanisms (OECD, 2007).
5. If the **time used for deciding on the submitted bids** is excessively short or lengthened by legal challenge, it can also signal corruption risks. Snap decisions on large value tenders may reflect premediated assessment, while legal challenge and the corresponding long decision period suggests outright violation of laws (Heggstad & Froystad, 2011).

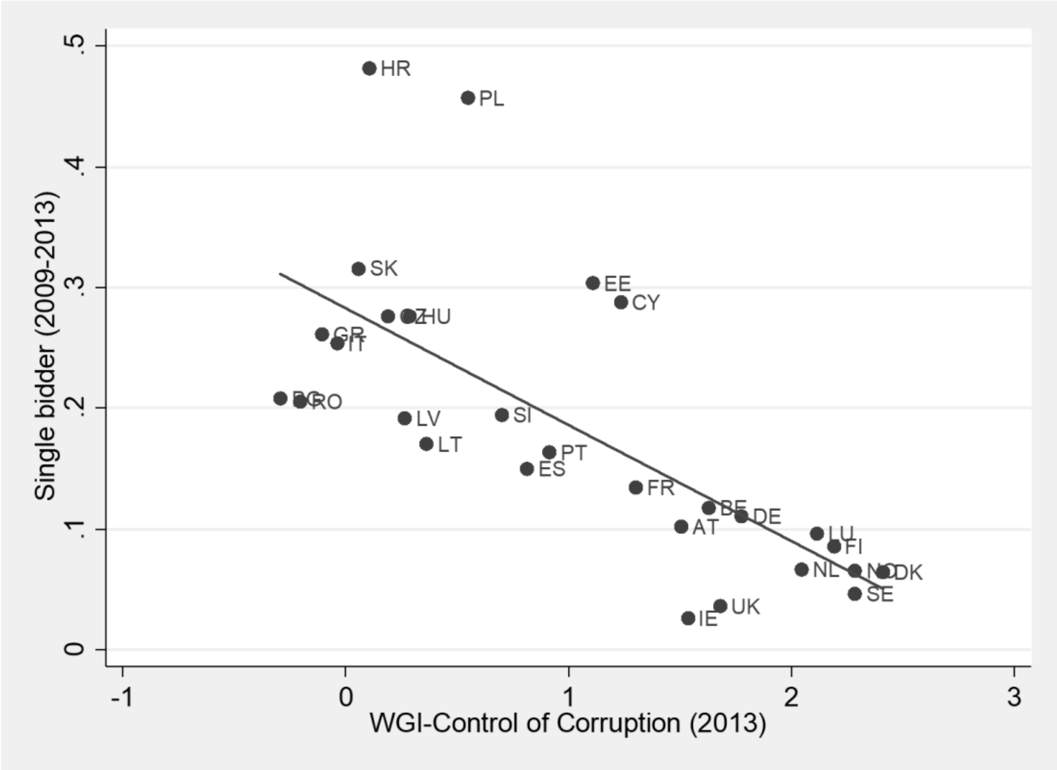
Regression analysis was used to identify ‘red flags’ which are most likely to signal corruption rather than any other phenomena such as low administrative capacity (full regression details can be found in Appendix A. Ultimately, those variables were selected which were large and significant predictors of single bidder contracts. The regression set-up controlled for a number of likely confounders of bidder numbers: (1) institutional endowments measured by type of issuer (e.g. municipal, national), (2) product market and technological specificities measured by CPV division of products procured, (3) contract size (log contract value in EUR), and (4) regulatory changes as proxied by year of contract award. Regressions were run for each country separately in order to best capture national specificities of corruption technologies and institutional endowments. We restricted the sample used in the regression analysis and eventually for defining corruption proxies in two ways: 1) Competitive markets: we only examine tenders in markets with at least 10 contracts awarded throughout 2009-2014, where markets are defined by product type (CPV level 3) and location (NUTS level 1). 2) Regulated tenders: we only used those tenders which are above EU thresholds in order to avoid the noise of too small contracts and voluntary reporting which follows erratic patterns across countries and over time. These together excluded 17% of the original sample.

The logic of regression analysis is the following: if in a certain country, not publishing the call for tenders in the official journal for open procedures is associated with a higher probability of a single bidder contract award, it is likely that avoiding the transparent and easily accessible publication of a tender is typically used for limiting competition. This would imply that call for tenders not published in the official journal becomes part of the analysed country’s CRI. Taking another example, if we found that leaving only 5 or fewer days for bidders to submit their bids is associated with a higher probability of a single bidder contract compared to periods longer than 20 days (a category mandated by many procurement regulations as golden standard), this would indicate that extremely short advertisement periods are often used to limiting competition. Then this would provide sufficient grounds to include the ‘5 or fewer days‘ category of the advertisement period variable in the CRI of the country in question. Following this logic, in addition to the outcome variable of these regressions (single bidding) only those variables are included in CRI which are in line with our corrupt contracting model, that is which are significant and large predictors of single bidding. Once the list of elementary corruption risk indicators is determined with the help of the above regressions, each of the variables and their categories receive a component weight in order to scale CRI to the 0 and 1 band.

Each of the two corruption risk indicators have its pros and cons. They are both preferable over widely used perception indicators as they directly point at specific corruption risks reflecting actor behaviour as captured by objective data. The strength of the single bidder indicator is that it is very simple and straightforward to interpret. However, it is also more prone to gaming by corrupt actors due to its simplicity (e.g. faking competition by submitting another, deliberately losing bid). The strength of the composite indicator approach (CRI) is that while individual strategies of corruption may change as the environment changes, they are likely to be replaced by other techniques. Therefore, the composite indicator is a more robust proxy of corruption over time than a single variable approach. In an international comparative perspective, a further strength of CRI is that it balances national specificities with international comparability by allowing for the exact list and formulation of the components to vary reflecting differences in local market conditions and corrupt practices (e.g. in some countries short advertisement periods are rare and are not systematically associated with a heightened probability of single bidding, hence the national CRI does not include this component). Hence, CRI in an international comparative perspective captures the frequency of likely corrupt behaviours in each context while abstracting from the specific procedural details. The main weakness of CRI is that it can only capture a subset of corruption strategies used, hence it may be a biased measure of corruption.

It can be argued that both corruption proxies capture the simplest and most straightforward signs of competition restriction, hence they miss out on sophisticated types of corruption such as corruption combined with inter-bidder collusion. This is problematic in as much as corruption in infrastructure is frequently combined with collusive bidding as for example extensive market transparency and high entry costs encourage cartels (Klemperer, 2007; Tóth, Fazekas, Czibik, & Tóth, 2014).

As each of the above corruption proxies can be associated with non-corrupt phenomena, we need to look at external validity tests to demonstrate that they are more often linked to corrupt phenomena than other behaviours. The simplest such test is comparing objective corruption proxies with corruption perceptions which provides a strong support to our interpretation of the data (Figure 1). As country-level perceptions might be biased in a number of ways, we also need to conduct micro-level validity tests using other objective corruption risk proxies. One such widely accepted proxy is the registration of companies in tax haven jurisdictions such as Panama (Shaxson & Christensen, 2014). Hence, it is expected that a contract represents a higher corruption risk if it is awarded to a company registered in a tax haven as its secrecy allows for hiding illicit money flows. In line with our expectations, all across the EU27 plus Norway there is a marked and significant difference in the percentage of single bidder contracts won by foreign companies registered in tax havens versus those which are not: 0.28 versus 0.26; similarly for CRI: 0.34 versus 0.31 respectively ($N_{contract}$=28,642). For further validity tests and validating applications, see (Charron et al., 2017; Fazekas & Kocsis, 2017; Klasnja, 2016).




Panel II

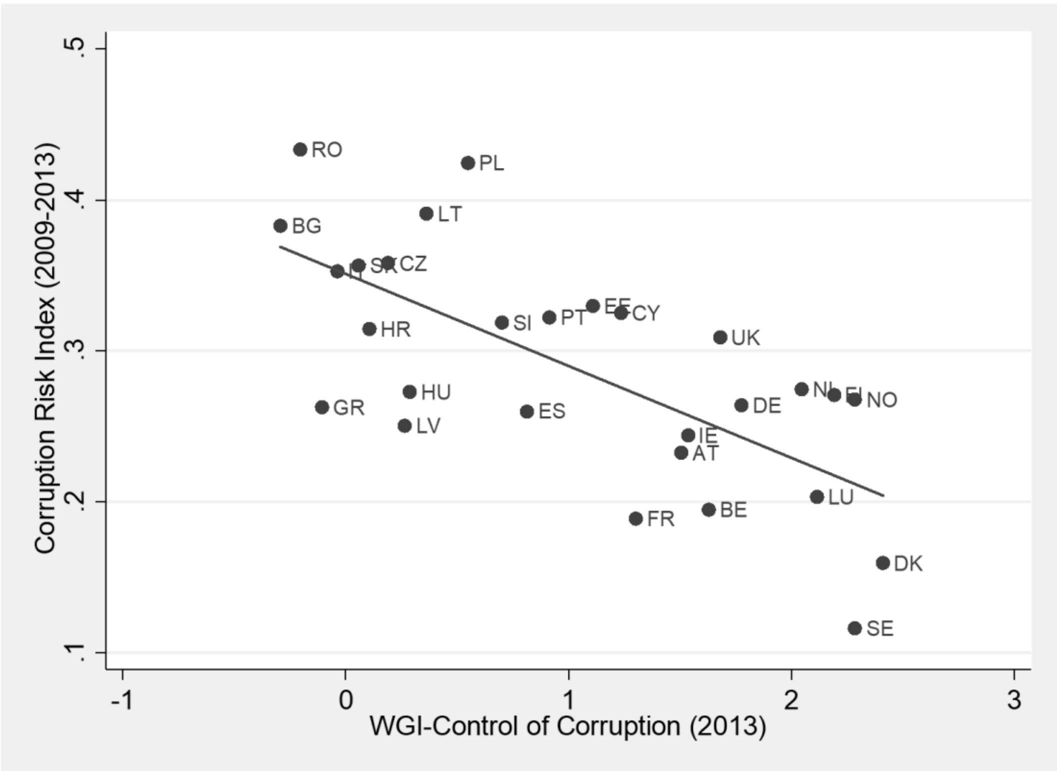



**Figure 1.** Bivariate relationship between WGI-Control of Corruption (2013) and corruption proxies: single bidder ratio and average CRI (period averages for 2009-2013), EU-27+Norway

## 3. Public procurement data

The database we use derives from public procurement announcements of 2009-2014 in the EU27 (i.e. EU28 minus Malta) (this database is called Tenders Electronic Daily (TED)[7], which is the online version of the 'Supplement to the Official Journal of the EU, dedicated to European public procurement.) (DG GROWTH, 2015). The data represents a complete database of all public procurement procedures conducted under the EU Public Procurement Directives. The database was released by the European Commission - DG GROWTH which also has conducted a series of data quality checks and enhancements. TED contains variables appearing in 1) calls for tenders, and 2) contract award notices. All the countries' public procurement legislation is within the framework of the EU Public Procurement Directive and are therefore, by and large, comparable. The TED database contains over 2.8 million contracts for the 27 EU Member States considered.

For most calculations, we restricted the sample to contracts directly linked to transport infrastructure development types readily identifiable in the CPV nomenclature[8]: roads, railways, airports and water transport (Table 2).[9]

**Table 2.** Descriptive statistics of spending on main infrastructure types, TED database, EU27, 2009-2014

| Infrastructure type | mean (million EUR) | total (million EUR) | $N_{contracts}$ |
|---|---|---|---|
| Roads | 6.3 | 134,652 | 31,139 |
| Railway | 19.7 | 57,220 | 4,027 |
| Airports | 5.2 | 2,948 | 751 |
| Water transport | 6.7 | 20,668 | 4,124 |

## 4. Corruption risks, prices, spending structure in European infrastructure sector

### *4.1 Mapping corruption risks in infrastructure delivery across Europe*

Looking at corruption risks in transport infrastructure related public procurement tenders across Europe as measured by objective indicators partially confirms widely held perceptions while partially contradicting them too. On the one hand, the simplest indicator of corruption risk we employ – single bidding - suggests that countries which are perceived to be most corrupt in general are the ones having the highest risk of corruption in transport related investments (Figure 2). These countries are mainly Central- and Eastern European countries, and Italy which have about 2-3 times higher share of contracts with a single bidder than their Western European counterparts. Given that single bidder contracts are on average 20% more expensive[10] than contracts with multiple bidders, such macro differences translate into tangible and large costs to public budgets.

[7] http://ted.europa.eu/
[8] Detailed CPV codes can be found in Appendix C.
[9] Contract-level data on infrastructure provision including the risk scores used in this article are accessible at www.digiwhist.eu and www.governancereport.org.
[10] The relative contract value – i.e. the ratio of the final price and the estimated price – is around 1 in case of single bidding, while only 0.83 in case of multiple bidders in case of transport related infrastructure tenders. For more on relative contract values see section 4.3.

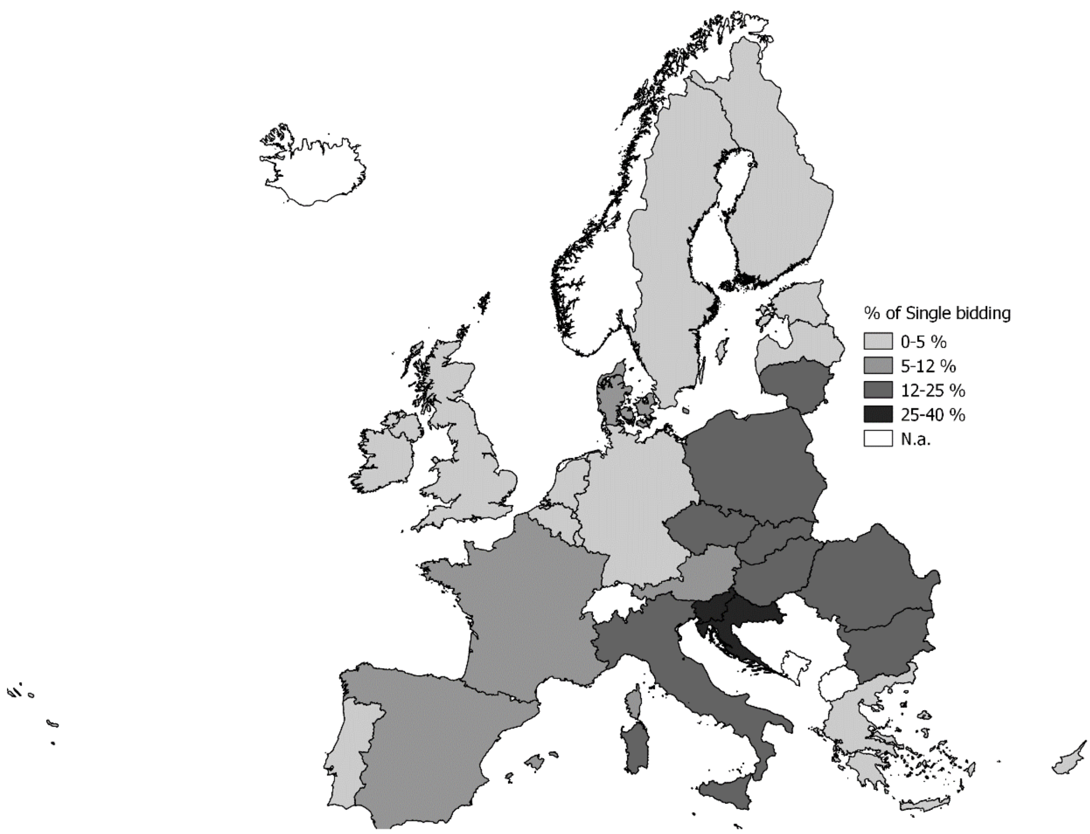



**Figure 2.** Average share of single-bidder tenders of transport infrastructure related public procurement tenders by countries (2009-2014)

On the other hand, looking at the regional level unveils significantly more diverse patterns (Figure 3). The share of single bidder transport infrastructure contracts varies greatly within countries and in some cases such as the UK, Spain or France, a few regions greatly underperform compared to the national average making them more similar to Central and Eastern European regions rather than the rest of their own countries. These large within-country differences at least partially reflect decentralized infrastructure decision making (Bardhan & Mookherjee, 2006).

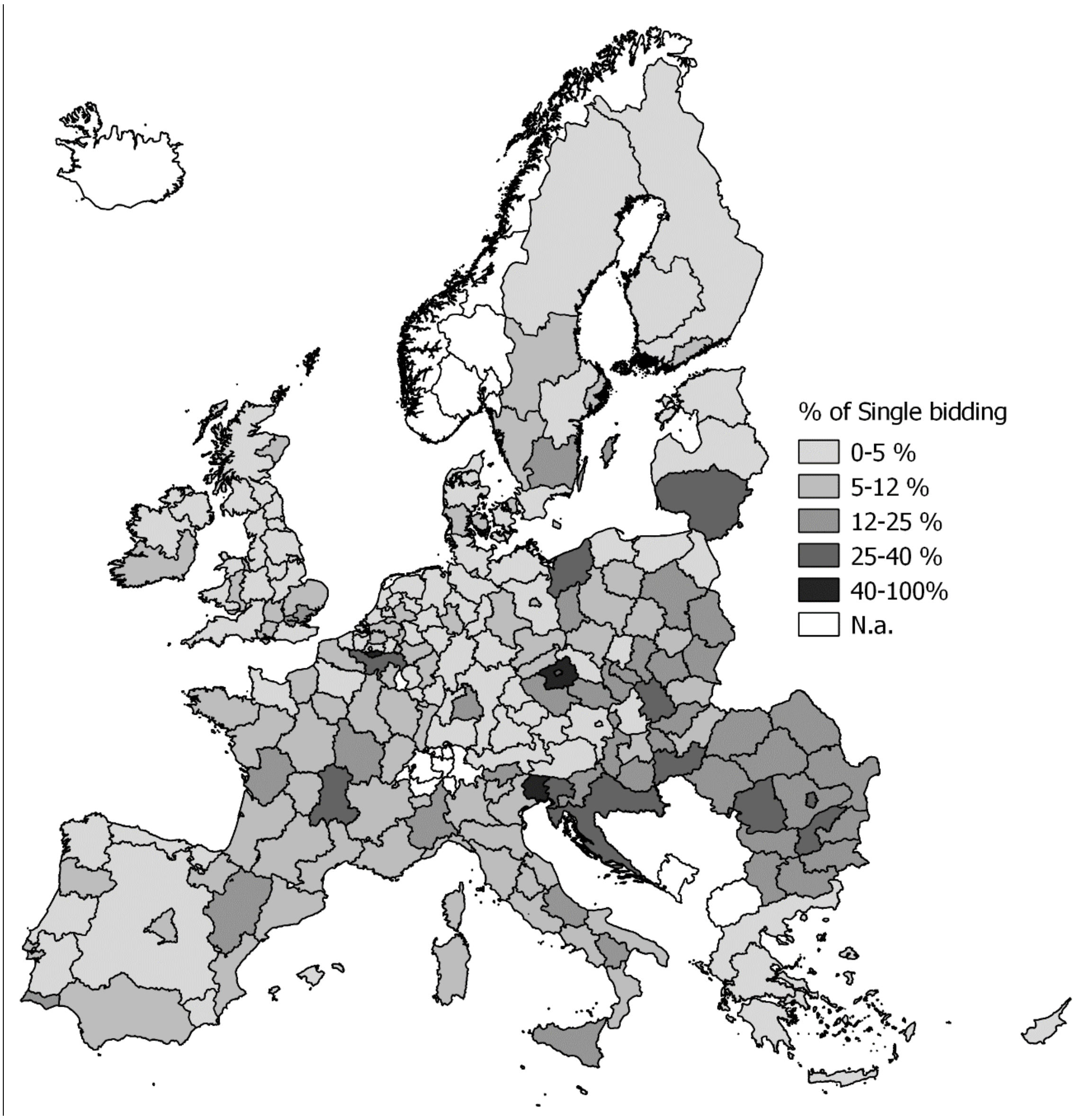



**Figure 3.** Average share of single-bidder tenders of transport infrastructure related public procurement tenders by NUTS2 regions[11]

This diversity of corruption risks within countries is paralleled by sectorial differences all across the EU. When comparing transport infrastructure tenders to other construction and non-construction related tenders the findings are surprising (Figure 4). The prevalence of single bidding outside the construction sector is roughly 2-3 times higher in every major EU region (for full definition of macro regions see Appendix B). However, the picture remains unaltered when we look at transport related investments only.

[11] Figure 3 shows the average share of single-bidder tenders by NUTS-2 regions. Where there were no observations on the NUTS-2 level, the average of the NUTS-1 level single-bidder share was used.

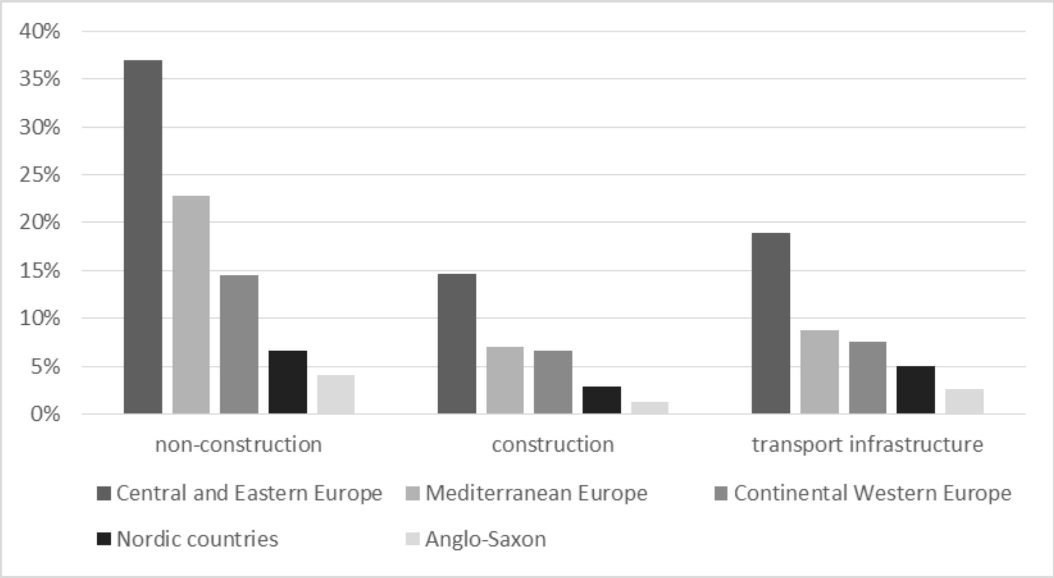



**Figure 4.** % of single bidder contracts in major European regions according to sectors[12], 2009-2014

Finally, the transport sector in itself is highly diverse, depending on the type of transport infrastructure built (Table 3). According to both the simple and complex measures of corruption risks, airport related contracts are the highest risk. Surprisingly, road construction scores are the best among transportation types analysed which may contradict some perceptions. From a theoretical perspective these results are hardly surprising, more specific large value tenders with relatively few providers are more prone to corruption. Looking at the time trends reveals that while corruption risks of road and water transport related construction decreased slightly in our period, rail and air transport related risks increased (Figure 5).

While the objective corruption proxies are the best which are available for such analysis, they may understate corruption risks in a number of ways. For example, sophisticated tactics of limiting competition such as companies taking turns in winning in a cartel (see e.g. (Tóth et al., 2014)) or non-observed corruption risks such as price increases during implementation are potential reasons for lower than expected risks in some sectors and regions.

**Table 3.** Descriptive statistics of different infrastructure types according to % single bidder and CRI, TED database, EU27, 2009-2014

| | Mean single bidder share | SD of single bidder share | Mean CRI | SD of CRI | Number of contracts |
|---|---|---|---|---|---|
| Road construction | 9% | 0.28 | 0.18 | 0.15 | 25,581 |
| Railway construction | 12% | 0.32 | 0.23 | 0.21 | 2,822 |
| Airport construction | 24% | 0.43 | 0.26 | 0.21 | 635 |
| Water transport | 12% | 0.33 | 0.22 | 0.19 | 3,328 |

[12] 'Non-construction (public works)' contains all tenders referred as „works" by contract type excluding construction works based on CPV code (i.e. excluding all contracts under 45000000-7 (Construction works) CPV category), while 'construction (public works)' only contains tenders under 45000000-7 CPV category within "works".

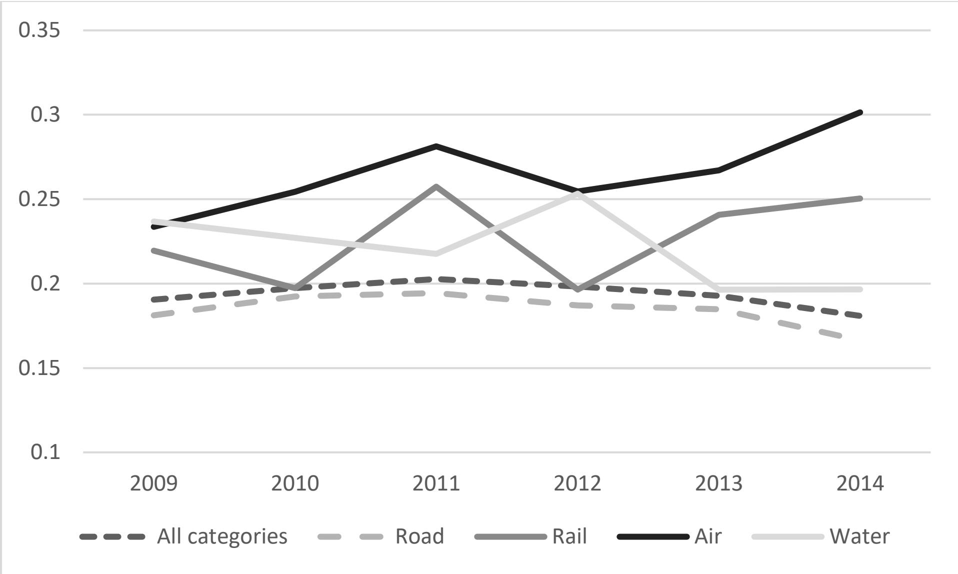



**Figure 5.** CRI values across years by infrastructure categories, TED database, EU27, 2009-2014

### *4.2 Impact of corruption on infrastructure spending structure*

As discussed above, corruption can distort government spending structure in many ways such as leading to over-investment into new infrastructure projects to the detriment of maintaining existing infrastructure. In this section, we explore how corruption risks are associated with high value hence more complex projects across Europe.

We use the same narrowly defined transport infrastructure sample as before. Spending structure is proxied by a simple measure: the share of the three largest projects within total spending[13]. If this indicator takes a high value, i.e. close to 100%, it suggests that government spending on infrastructure is dominated by a few high value, hence complex projects. In order to best capture the variation in corruption risks demonstrated above, the analysis is conducted on the regional level by transportation infrastructure types (i.e. each observation corresponding to a NUTS-2 region-transportation type-year combination). The regional categorization is based on the place of project implementation, hence centrally purchased projects are assigned to regions using the location of contract performance which reflects the focus on implementation risks. Corruption risks are measured by the more comprehensive Corruption Risk Index (CRI) variable for contracts awarded.

Although, this narrowly defined calculation leads to a relatively high share of the three largest projects (approximately 80%), there is large enough variance to investigate the relationship between corruption risks and the share of the largest projects by region[14]. Figure 6 depicts the relationship between the average CRI and the share of the three largest projects across Europe. We find a rather weak positive relationship between the two variables which is particularly pronounced for high corruption risk regions (i.e. regions with higher than 0.4 CRI). This suggests that it is especially high degrees of corruption amounting to systematic corruption in a region which is most influential on project size distributions.

[13] This approach is similar to the simple concentration measures used in competition policy (i.e. market share of the top 4 companies). Using similar measures of concentration (e.g.$concentration_m = \sum_{i=1}^{n}(share\ of\ contract_i)^2$, where $m$ stands for a given 'market' corresponding to a NUTS-2 region-infrastructure sector-year combination), leads to analogous results as Figure 6 and Table 4 presents, i.e. high average corruption risks are associated with more concentrated market structures. A further alternative is using the top 5 contracts rather than top 3 contracts. Again the below results are analogous. Full calculations can be obtained from the authors.

[14] Note: we include only those observations having at least six tenders.

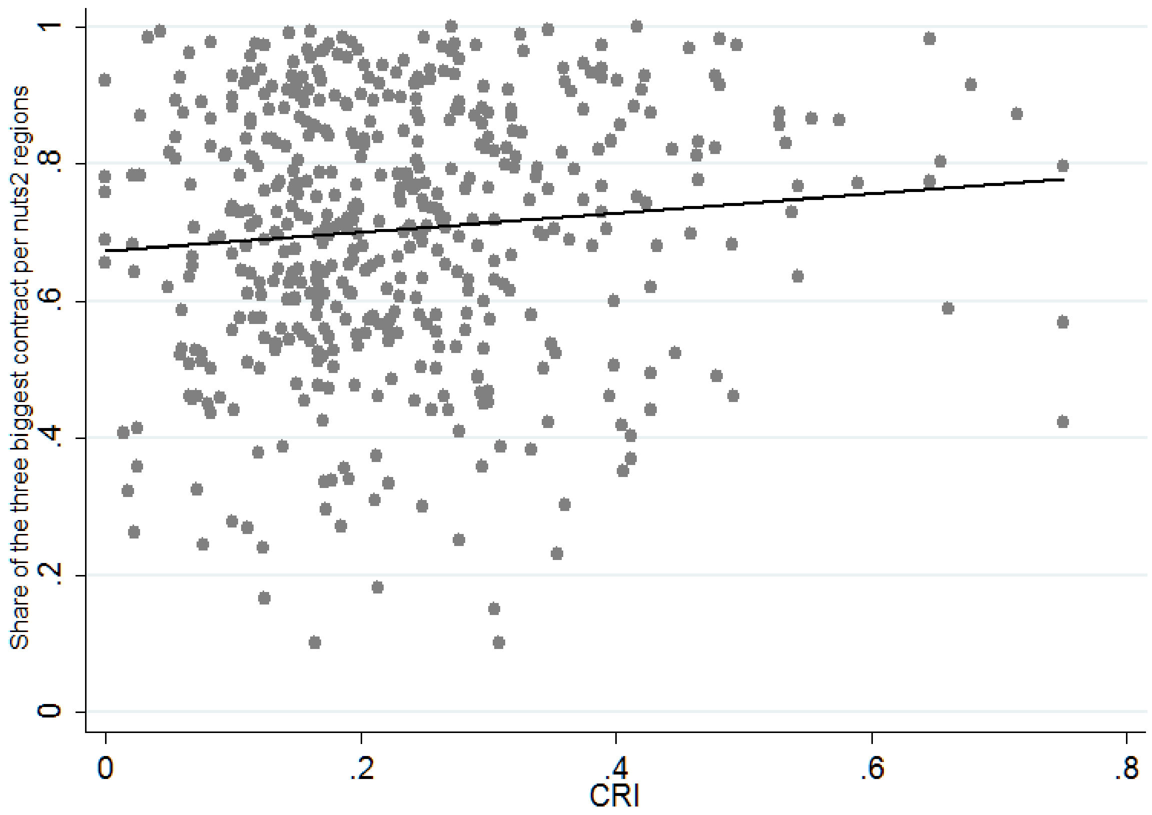



**Figure 6.** Share of 3 largest infrastructure contracts in total transport infrastructure spending and Corruption Risk Index, 2009-2014, $N_{region-year-sector}$=493

As both the share of the largest projects and its relationship with corruption risks can be driven by many unobserved factors, this connection is further investigated using linear regression analysis (OLS=ordinary least squares method). Since it is likely, that project size depends on the overall coverage and quality of the existing infrastructure, and also on the related policies and institutions (e.g. whether design or maintenance is procured together with construction works etc.), regression models include country and transportation infrastructure type dummy variables to control for these factors. Also, as the contract size of public works differs significantly from services and supplies, we include the share of public works contracts to control for its effect on the top three project's share. Furthermore, we also enter a regional GDP[15] measure to control for the level of development, the number of contracts to remove market size distortions, and year dummies to control for any time trends.

After including this wide range of control factors, the size of the CRI effect on spending concentration becomes significant and remains positive, indicating that corruption risks have an independent impact on spending structure even after taking into account major institutional and contract-related factors (Table 4). The fact that the estimated effect becomes significant in the more complex models shows that the above seemingly weak relationship in Figure 6 was confounded by these factors. In the most complete model with highest explanatory factor, increasing CRI for example from a low corruption risk Danish region (DK04) to a high corruption risk Czech region (CZ07) (increasing CRI by 0.36) is associated with an approximately 6.3% point increase in the largest three projects' share in total spending. All other control variables have the expected sign, e.g. higher share of public works contracts leads to a higher share of the largest infrastructure projects, while higher number of distinct contracts is associated with a smaller share.

In order to test the robustness of these relationships we considered two alternative specifications in Appendix D. First, considering that corruption risks may take a specific form in larger value infrastructure projects, we calculated a transport infrastructure specific corruption risk index (recall, the CRI used in the analysis has been developed using the economy-wide sample of contracts awarded which makes it suitable for cross-sectoral comparisons). Confirming our main model, the sector specific CRI effects on spending concentration are very similar to the ones estimated with the economy-wide CRI (Table D1). We continue reporting the broader economy-wide CRI results throughout the main text and delegate the alternative

[15] The logarithm of the average NUTS-2 level GDP (EUR per inhabitant) was used (Eurostat), as a control for regional development.

calculations to Appendix D in order to retain comparability with other research using the same CRI formulation. Second, in order to control for potential regional confounders not captured by regional GDP, we also run the regression models with region-sector fixed-effects which also show positive and significant relationship with spending concentration, while the CRI coefficients are even larger (Table D2).

**Table 4.** OLS regressions explaining the share of 3 largest infrastructure contracts in total transport infrastructure spending, 2009-2014, region-year-sector observations

| Model Nr. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: share of 3 largest contracts | | | | |
| avg. regional CRI | 0.182 | 0.164 | 0.200* | 0.176* |
| | (0.169) | (0.230) | (0.015) | (0.042) |
| Public works share | 0.177** | 0.173** | 0.182*** | 0.179*** |
| | (0.003) | (0.007) | (0.000) | (0.000) |
| Ref.cat.: Road | | | | |
| Rail | | 0.0403 | 0.0171 | -0.00886 |
| | | (0.317) | (0.631) | (0.809) |
| Airport | | 0.0399 | 0.00151 | -0.0138 |
| | | (0.719) | (0.988) | (0.892) |
| Water transport | | 0.0251 | 0.0228 | 0.00721 |
| | | (0.435) | (0.483) | (0.856) |
| # of contracts | | | | -0.00543* |
| | | | | (0.011) |
| Country control | | | X | X |
| Year control | | | X | X |
| GDP control | | | | X |
| Observations | 493 | 493 | 493 | 454 |
| R-squared | 0.065 | 0.071 | 0.255 | 0.343 |

Notes: p-values in parentheses; * p<0.05, ** p<0.01, *** p<0.001; standard errors clustered by country

### *4.3 Impact of corruption on infrastructure prices*

Transportation infrastructure development typically involves the procurement of complex public works making it difficult to measure prices. Given suitable resources, complex projects can be broken down into constitutive standard components such as labour costs, price of electricity used, or price of concrete built-in (European Court of Auditors, 2013). In the absence of such detailed information this article adopted two complementary strategies for estimating the impact of corruption risks on transport infrastructure prices. First, relative contract value, the ratio of actual contract value at contract award divided by originally estimated contract value at the call for tender stage (i.e. consultant’s estimate), was used, which is a standard, widely available measure of prices (Coviello & Mariniello, 2014). This ratio gives a rough estimate of price savings a tender achieved compared to the initial estimate. Second, unit prices could be calculated for a small set of new road construction contracts across Europe. This approach relies on the publicly available records in TED announcements which are often too imprecise for unit cost calculations hence the available sample is tiny compared to the whole infrastructure sector. The relative precision of calculations were traded off against scope of the sample, while our aim remained to get a better understanding about the magnitude of cost effects of corruption.

First, relative contract values increase in almost every EU Member State as a function of both corruption proxies – single bidding and CRI - in transport infrastructure contracts after controlling for usual determinants of prices such as transport infrastructure sub-sector, year, contract size, and contracting body characteristics such as main sector and type (full multilevel modelling regression outputs can be found in Table D3, with a confirmatory robustness test using the transport investment specific CRI in Table D4). Overall, all across the EU, CRI's total effect size amounts to 17%, that is increasing corruption risks from CRI=0 to CRI=1 is associated with about 17% points higher contract value compared to the original estimate. More realistically, an additional red flag (1/6 CRI score ~ 0.166) is associated with a 3% point higher prices compared to the original estimate. The estimated effect of CRI on relative prices per country is depicted in Figure 7. Interestingly, the strength of the relationship between corruption risks and prices greatly varies across countries with high corruption risk countries such as Poland, Slovakia and Italy displaying the strongest relationship with some low corruption risk countries also displaying a relatively strong relationship such as Finland[16]. In Poland for example, adding one red flag is associated with roughly 6% point higher transportation infrastructure prices. In order words, foregone savings amount to about 6% of the original estimated contract value.

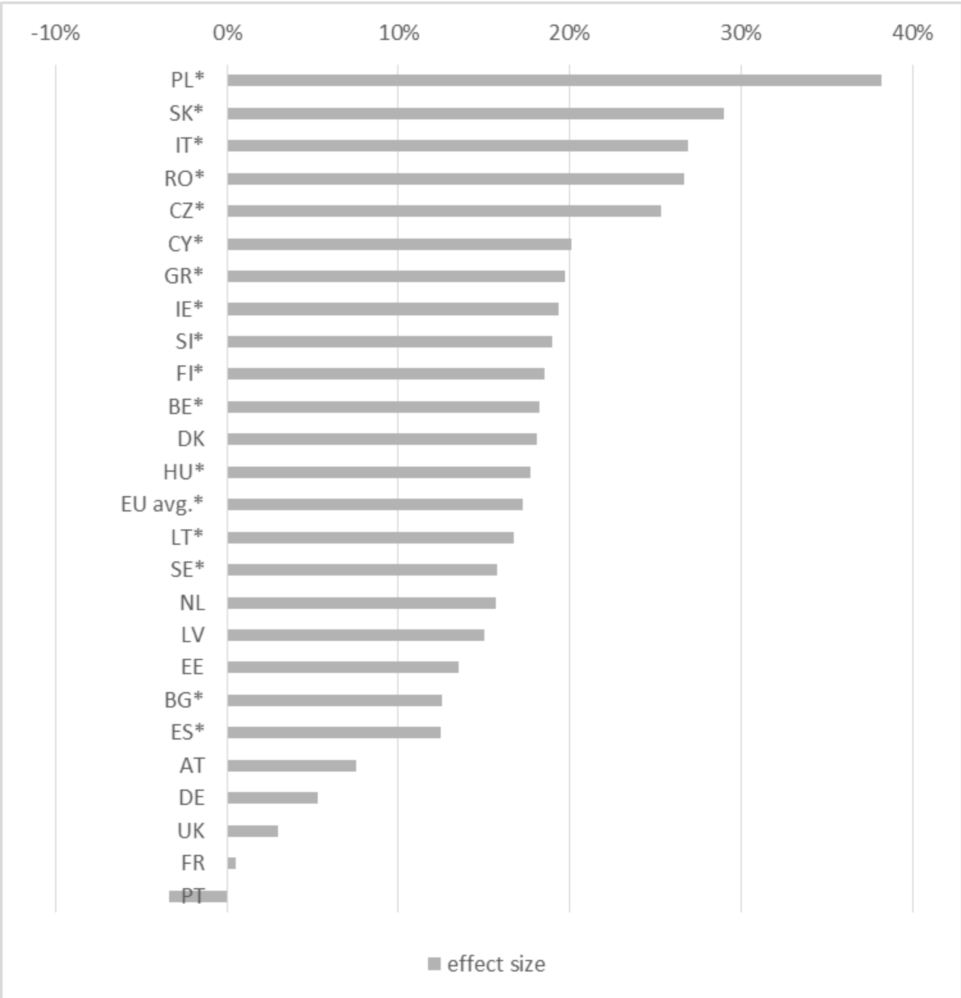



**Figure 7.** Average price increase as a function of Corruption Risk Index (regression coefficient), %, transport infrastructure, 2009-2014, $N_{contract}$= 10,600

*Note: *=country specific effect different from 0 at the 10% level.*

Second, the price of constructing 1 kilometre of new motorway differs greatly in Europe, apparently quite in the opposite direction than input prices such as labour or electricity would suggest making Spanish motorways

[16] Recall that Figure 7 only depicts the strength of relationship between corruption risks and relative contract values, which may or may not correlate with the prevalence of corruption risks in the country. For example, corruption risks are high in Bulgaria, but their estimated price impact is relatively muted, while there are very low corruption risks in Ireland, but their impact on prices is well above the EU average.

over 70% more expensive than German motorways (European Court of Auditors, 2013). At least some of these huge price differences can be explained by corruption risks according to our regression analysis. We regressed the contract-level average unit price for road and express road construction on CRI in 2010-2014 across the whole EU in a hand-collected sample of new express road and road construction projects neither of which included maintenance or design costs[17]. The regressions control for macro region (e.g. Western Europe), road length (km), road type (express road or road), development complexity (e.g. including bridge), geographical features (Shannon Evenness Index[18]), and civil engineering price level. According to such a simple regression, moving from the CRI of an average Bulgarian to an average Swedish road construction project (i.e. ~1.5 standard deviation increase or ~1 additional red flag) is associated with a EUR 1.2 million or ~35%[19] increase in the cost of 1 kilometre new road construction compared to the average road price in the sample (Table 5). The exact size of the estimated effect of corruption risks differs somewhat according to the control factors included in the regressions, but the overall relationship and its magnitude remains by and large the same, in the order of 30-35%[20].

Our estimations using newly collected data confirm prior research using independent samples and somewhat different methods. In the most directly comparable article using World Bank financed projects' data, one additional red flag is associated with a price increase of 30-40% compared to the mean km price (Alexeeva et al., 2011). Furthermore, key informant interviews in Hungary suggest about a 50% price increase in highway construction projects in the late 2000s (Fazekas et al., 2013). In some cases, the cost of corruption could be even higher as suggested by a recent 29 km highway construction project in Hungary (M4 motorway between Budapest-Szolnok). It was initially won by a consortium of 5 construction companies, with a stunning 12.6 million EUR/km cost level on probably the flattest part of the country. A comparably low-lying and flat terrain motorway construction project (Motorway A20, Grimmen-East to Strasburg) was audited by the European Court of Auditors calculating a 4.3 million EUR/km price, roughly one third of the Hungarian M4 price (European Court of Auditors, 2013). Most interestingly, the faith of the project seems to follow political-personal conflicts rather than engineering or economic calculation: initially a company called Közgép plc. was also among the winners which saw its public procurement revenue skyrocket after the 2010 government change and is owned by the high school roommate of the Prime Minister who also happened to be the ex-treasurer of the governing party. However, soon after the Prime Minister and the former roommate had a public fall-out, the project was halted, no further payment was made, and allegations of bid rigging were investigated[21].

[17] The exclusion of tenders also containing maintenance or design costs was based on the publicly available tender descriptions published on the TED website which was not detailed in every case making sample selection uncertain at places.

[18] http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Shannon_evenness_index_%28SEI%29

[19] Note that effect magnitudes between road unit-price and relative price models are not directly comparable as they relate to different samples and their baselines are different: relative prices use the initial cost estimate to measure changes in percentage points, whereas road unit prices use million EUR/km.

[20] This effect is likely to include both the direct effect of corruption risks on prices and the indirect effects going through unobserved cost drivers such as detailed technical specification as long as they are correlated with corruption risks. As section 4.2 suggests, corruption risks also influence project design.

[21] For more information (in Hungarian) see :
http://index.hu/gazdasag/2015/06/19/nem_epul_meg_az_m4-es_autopalya/;
http://nol.hu/gazdasag/tudtak-mertek-tettek-1531823;
http://mandiner.hu/cikk/20150404_nepszabadsag_kirugja_a_kozgepet_a_kormany;
http://index.hu/gazdasag/2015/04/07/m4_autopalya_leallitas_miniszterelnokseg_nemeth_laszlone_lazar_unio_kartellgyanu/

**Table 5.** Average unit price increase as a function of Corruption Risk Index (regression coefficient), %, transport infrastructure, 2010-2014[22]

| Model description | Region controls | Region, geography, price level controls |
|---|---|---|
| Dependent variable: unit price | | |
| CRI | 5770353.2* | 6431910.6* |
| | (0.021) | (0.032) |
| Km | -155132.8** | -144157.8** |
| | (0.002) | (0.002) |
| Express road | 1648830.8 | 2057738.1 |
| | (0.171) | (0.081) |
| Complex | 1152491.7 | 1156695.1 |
| | (0.052) | (0.055) |
| Shannon-index | | 10933458.1 |
| | | (0.118) |
| Civil engineering price level | | 3602.6 |
| | | (0.805) |
| Observations | 97 | 96 |
| R-squared | 0.257 | 0.303 |

*Notes: p-values in parentheses; * p<0.05, ** p<0.01, *** p<0.001, standard errors clustered by region*

## 5. Conclusions and policy lessons

This article presented a unique summary of the existing literature of transportation investment related corruption risks and the associated costs. It also provided a novel measurement approach to corruption and its effects on spending structure and prices by using newly collected, high resolution, EU-wide government contracting data on transport-related public procurement tenders. Countries in greatest need of investment in transportation infrastructure and those receiving the bulk of EU Funds are those which have the highest corruption risks. Central- and Eastern European and Mediterranean countries procure transportation infrastructure development at about 2-3 times higher corruption risks than Western European countries.

The review of international evidence on the effects of corruption in transport infrastructure showed three direct ways through which public goals can be compromised: 1) distortion of spending structure and project design; 2) inflation of prices; and 3) contribution to delayed, low quality, or incomplete provision. This article investigated the first two effects in detail by using novel data and indicators. Findings suggest that corruption can steer infrastructure spending towards high value investments, especially in high corruption risk regions of Czech Republic, Romania, Hungary and Italy. In our regression models, increasing Corruption Risk Index (CRI) for example from a low corruption risk Danish region (DK04) to a high corruption risk Czech region (CZ07) (increasing CRI by 0.36) is associated with an approximately 6.3% point increase in the largest three projects' share in total spending. Corruption is also likely to inflate prices. For example, in a simple regression model controlling for macro region, road length, road type, project complexity, geographical features, and civil

[22] The sample partly follows from a manual data collection. The road length data was collected from tender announcement documents, which causes some degree of selection bias (less transparency considered is a good way to hinder competitive bidding). The variable complex is a dummy variable, controlling for additional investments in the projects (bridges, several road crossings, viaducts etc.). Shannon-index is a control for land cover variability, while civil engineering costs are used as a proxy for labour costs (both variables come from Eurostat data). Note, that we also estimated these effects with the transport investment specific CRI and the results remain positive and significant. For the sake of brevity, we do not report these results separately.

engineering price level, moving from the CRI of an average Bulgarian to an average Swedish road construction project (i.e. ~1.5 standard deviation increase or ~1 additional red flag) is associated with a EUR 1.2 million or ~35% increase in the cost of 1 kilometre new road construction compared to the average road price in the sample. Such effect size is in line with prior research using independent samples, but comparable methodology.

Our findings show the potential for future analysis using Big Data on transport infrastructure projects providing new insights in addition to existing case study research and analysis using country-level indicators. Nevertheless, our work only represents the first step in this direction with a number of limitations. Crucially, our evidence is only suggestive of causal relationships calling for further work precisely identifying causal mechanisms. Furthermore, detailed data on projects as well as contract implementation will over time complement our dataset on the tendering process and contract award in order to gain a more complete picture of when and how corruption can influence infrastructure provision. Detailed case studies carefully selected based on the here presented Big Data analysis could provide further crucial cues as to how precisely corruption impacts accrue.

The evidence presented here demonstrates the detrimental effects of corruption in transport infrastructure development and adds to the growing concern that many countries heavily investing in infrastructure achieve value for money and the desired outcomes. Tighter, real-time monitoring of corruption risks on the contract level would allow for early intervention and preventive measures curbing the cost of corruption in infrastructure provision.

## References

Alexeeva, V., Padam, G., & Queiroz, C. (2008). *Monitoring Road Works Contracts and Unit Costs for Enhanced Governance in Sub-Saharan Africa*. Washington, DC.

Alexeeva, V., Queiroz, C., & Ishihara, S. (2011). *Monitoring Road Works Contracts and Unit Costs for Enhanced Governance in Europe and Central Asia* (Transport Papers No. TP-33). Washington, DC.

Bardhan, P., & Mookherjee, D. (2006). Decentralisation and Accountability in Infrastructure Delivery in Developing Countries*. *The Economic Journal*, *116*(508), 101–127. JOUR. http://doi.org/10.1111/j.1468-0297.2006.01049.x

Becker, S. O., Egger, P. H., & Ehrlich, M. von. (2012). *Absorptive Capacity and the Growth and Investment Effects of Regional Transfers: Regression Discontinuity Design with Heterogeneous Treatment Effects* (Working Paper Series No. No. 89). Warwick.

Benitez, D., Estache, A., & Soreide, T. (2010). *Dealing with politics for money and power in infrastructure* (Policy Research Working Paper No. 5455). *Policy Research Working Papers*. Washington, DC: The World Bank.

Charron, N., Dahlström, C., Fazekas, M., & Lapuente, V. (2017). Careers, Connections, and Corruption Risks: Investigating the impact of bureaucratic meritocracy on public procurement processes. *Journal of Politics*, 79(1), 89-104.

Chong, E., Klien, M., & Saussier, S. (2015). *The Quality of Governance and the Use of Negotiated Procurement Procedures: Evidence from the European Union* (EPPP DP No. 2015–3). Paris.

Collier, P., Kirchberger, M., & Söderbrom, M. (2015). The cost of road infrastructure in low and middle income countries. Policy Research working paper; no. WPS 7408. Washington, D.C. : World Bank Group.

Coviello, D., & Mariniello, M. (2014). Publicity requirements in public procurement: Evidence from a regression discontinuity design. *Journal of Public Economics*, *109*, 76–100.

DG GROWTH. (2015). *TED Structured Dataset (2009-2014), Tenders Electronic Daily, supplement to the Official Journal of the European Union*. Brussels.

Dixit, A. K. (1998). *The Making of Economic Policy: A Transaction-cost Politics Perspective*. MIT Press.

European Court of Auditors. (2013). *Are EU Cohesion Policy funds well spent on roads?* Luxembourg: European Court of Auditors.

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

Fazekas, M., & Kocsis, G. (2017). *Uncovering High-Level Corruption: Cross-National Corruption Proxies Using Government Contracting Data*. British Journal of Political Science, available online: https://doi.org/10.1017/S0007123417000461

Fazekas, M., & Tóth, I. J. (2014). *New ways to measure institutionalised grand corruption in public procurement* (U4 Brief No. 2014:9). U4 Anti-Corruption Resource Centre, Bergen, Norway.

Fazekas, M., Tóth, I. J., & King, L. P. (2013). *Corruption manual for beginners: Inventory of elementary "corruption techniques" in public procurement using the case of Hungary* (Working Paper Series No. GTI-WP/2013:01). Government Transparency Institute, Budapest.

Fazekas, M., Tóth, I. J., & King, L. P. (2016). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal of Criminal Policy and Research*, *22*(3), 369–397.

Flyvbjerg, B., Holm, M. S., & Buhl, S. L. (2003). How common and how large are cost overruns in transport infrastructure projects? *Transport Reviews*, *23*(1), 71–88. http://doi.org/10.1080/01441640309904

Flyvbjerg, B., Holm, M. S., & Buhl, S. L. (2004). What Causes Cost Overrun in Transport Infrastructure Projects? *Transport Reviews*, *24*(1), 3–18. http://doi.org/10.1080/0144164032000080494a

Gillanders, R. (2013). Corruption and Infrastructure at the Country and Regional Level. *The Journal of Development Studies*, *50*(6), 803–819. http://doi.org/10.1080/00220388.2013.858126

Glazer, A. (1989). Politics and the Choice of Durability. *American Economic Review*, *79*(5), 1207–1213. http://doi.org/10.2307/1831446

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

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

Johnston, M. (1996). The search for definitions: the vitality of politics and the issue of corruption. *International Social Science Journal*, *48*(149), 321–335. JOUR.

Kenny, C. (2006). *Measuring and Reducing the Impact of Corruption in Infrastructure* (World Bank Policy Research Working Paper No. 4099). Washington DC.

Kenny, C. (2007). *Construction, corruption, and developing countries* (World Bank Policy Research Working Paper No. 4271). Washington, DC.

Kenny, C. (2009a). Measuring Corruption in Infrastructure: Evidence from Transition and Developing Countries. *The Journal of Development Studies*, *45*(3), 314–332. JOUR. http://doi.org/10.1080/00220380802265066

Kenny, C. (2009b). Transport Construction, Corruption and Developing Countries. *Transport Reviews*, *29*(1), 21–41. JOUR. http://doi.org/10.1080/01441640802075760

Klasnja, M. (2016). Corruption and the Incumbency Disadvantage: Theory and Evidence. *Journal of Politics*, *77*(4), 928–942.

Klemperer, P. (2007). Bidding markets. *Journal of Competition Law and Economics*, *3*(1), 1–47.

Le, Y., Shan, M., Chan, A. P. C., & Hu, Y. (2014). Overview of Corruption Research in Construction. *Journal of Management in Engineering*, *30*(4), 2514001. http://doi.org/10.1061/(ASCE)ME.1943-5479.0000300

Lengwiler, Y., & Wolfstetter, E. (2006). Corruption in procurement auctions. In N. Dimitri, G. Piga, & G. Spagnolo (Eds.), *Handbook of Procurement* (pp. 412–429). Cambridge, UK: Cambridge University Press.

Lewis-Faupel, S., Neggers, Y., Olken, B. A., & Pande, R. (2014). *Can Electronic Procurement Improve Infrastructure Provision? Evidence From Public Works in India and Indonesia* (NBER Working Paper Series No. 20344).

Lukács, P. A., & Fazekas, M. (2015). *The political economy of grand corruption in public procurement in the construction sector of Hungary*. ANTICORRP: Work Package 8 - Corruption, assistance and development, Berlin.

Mauro, P. (1997). The Effects of Corruption on Growth, Investment, and Government Expenditure: A Cross-Country Analysis. In K. A. Elliott (Ed.), *Corruption and the Global Economy* (pp. 83–108). Washington, DC: Institute for International Economics.

Mauro, P. (1998). Corruption and the composition of government expenditure. *Journal of Public Economics*, *69*, 263–279.

Meduri, S. S., & Annamalai, T. R. (2012). Unit Costs of Public and PPP Road Projects: Evidence from India. *Journal of Construction Engineering and Management*, (January), 409. http://doi.org/10.1061/(ASCE)CO.1943-7862.0000546

Mungiu-Pippidi, A. (2015). *The Quest for Good Governance. How Societies Develop Control of Corruption*. Cambridge, UK: Cambridge University Press.

North, D. C., Wallis, J. J., & Weingast, B. R. (2009). *Violence and Social Orders. A Conceptual Framework for Interpreting Recorded Human History*. Book, Cambridge, UK: Cambridge University Press.

OECD. (2007). *Integrity in Public Procurement. Good Practice from A to Z*. Paris: OECD.

Olken, B. A. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, *115*(2), 200–249. JOUR. http://doi.org/10.1086/517935

Olken, B. A. (2009). Corruption perceptions vs. corruption reality. *Journal of Public Economics*, *93*(7–8), 950–964. http://doi.org/10.1016/j.jpubeco.2009.03.001

Piga, G. (2011). A fighting chance against corruption in public procurement? In S. Rose-Ackerman & T. Soreide (Eds.), *International Handbook on the Economics of Corruption, Volume Two* (pp. 141–181). Cheltenham, UK: Edward Elgar.

Pricewaterhouse Coopers. (2013). *Identifying and reducing corruption in public procurement in the EU*. Brussels: PricewaterhouseCoopers and Ecorys.

Rose-Ackerman, S. (1999). *Corruption and Government: Causes, Consequences, and Reform*. Cambridge, UK: Cambridge University Press.

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

Shaxson, N., & Christensen, J. (2014). *The Finance Curse. How Oversized Financial Centres Attack Democracy and Corrupt Economies*. Chesham, UK: Tax Justice Network.

Tanzi, V., & Davoodi, H. R. (1997). *Corruption, Public Investment, and Growth* (IMF Working Paper No. WP/97/139). Washington, DC.

Tóth, B., Fazekas, M., Czibik, Á., & Tóth, I. J. (2014). *Toolkit for detecting collusive bidding in public procurement. With examples from Hungary* (Working Paper Series No. CRCB-WP/2014:02). Budapest: Corruption Research Center Budapest.

Transparency International. (2005). India Corruption Study 2005.

World Bank. (2009). *Fraud and Corruption. Awareness Handbook*. Washington DC: World Bank.

## Appendix A: Regressions underlying the Corruption Risk Index (CRI)

**Table A1.** Binary logistic regression results on contract level, 2009-2014, EU27+Norway, average marginal effects reported, N=1,306,025, all regressions contain control variables: buyer sector, buyer type, year, product market, contract value, country

| Dependent variable | single bid=1 | | | | | |
|---|---|---|---|---|---|---|
| Model | (1) | (2) | (3) | (4) | (5) | (7) |
| no call for tender published | 0.182** | | | | | 0.120** |
| | (0.000) | | | | | (0.000) |
| restricted procedure | | 0.188** | | | | 0.141** |
| | | (0.000) | | | | (0.000) |
| risky evaluation criteria | | | 0.038** | | | 0.039** |
| | | | (0.000) | | | (0.000) |
| extreme submission period | | | | 0.008** | | 0.014** |
| | | | | (0.000) | | (0.000) |
| extreme decision period | | | | | 0.034** | 0.057** |
| | | | | | (0.000) | (0.000) |
| R-squared | 0.143 | 0.145 | 0.135 | 0.135 | 0.136 | 0.151 |

*Note: p-values in parentheses;* *$p < 0.05$, **$p < 0.01$
Source: Fazekas & Kocsis (2017), Table 1.

## Appendix B: Macro-region definitions

**Table B1.** Definition of EU macro-regions

| **Macro region** | **Countries** |
|---|---|
| Anglo-Saxon countries | IE, UK |
| Central and Eastern Europe (countries of the 2004 and 2007 EU accession rounds) | BG, HR, CZ, CY, EE, HU, LV, LT, MT, PL, RO, SK, SI |
| Continental Western Europe | AT, BE, DE, FR, LU, NL |
| Mediterranean Europe | GR, IT, PT, ES |
| Nordic countries | DK, FI, SE |

## Appendix C: Infrastructure definitions

Generally, we refined the CPV code classification based on 'additional CPV' codes. In case of having a more precise additional CPV code in comparison to the main CPV code, and if no contradicting additional CPV codes were used in the given contract, then the more precise additional CPV code was used for classification. E.g. in case of having 45000000 (Construction works) as the main CPV code, and only one additional CPV code was given, which is 45233100 (Construction work for highways, roads), then this latter code was used in order to have a better sample.

The (refined) CPV codes defining the different types of infrastructure categories are enlisted below. In certain cases, the categorization also depends on additional CPV codes, e.g. if the main CPV is construction material, but the additional CPV codes specify whether they belong to road, rail etc. construction, then the items are categorized based on these additional codes.

**Table C1.** CPV codes used for defining different infrastructure sectors

| Sector | CPV codes used |
|---|---|
| Road construction | ▪ Between 45233000 and 45234000: Construction, foundation and surface works for highways, roads<br>▪ 45221111: Road bridge construction work<br>▪ 44113000: Road-construction materials<br>▪ 44113700: Road-repair materials<br>▪ 44113900: Road-maintenance materials<br>▪ 44113310: Coated road materials<br>▪ 44811000: Road paint<br>▪ 44113800: Road-surfacing materials<br>▪ All material related CPV (beyond 44000000) if the additional CPV code is related to road construction (beyond 45000000)<br>▪ Road equipments: between 34920000 and 34930000<br>▪ 50232000 : Maintenance services of public-lighting installations and traffic lights<br>▪ 50232100 : Street-lighting maintenance services<br>▪ 50232110 : Commissioning of public lighting installations<br>▪ 50232200 : Traffic-signal maintenance services<br>▪ Between 71311210 and 71311220: Architectural works related to road construction<br>▪ Any architectural works (beyond 71000000) if additional CPV code is related to road construction (between 45233000 and 45234000) |
| Railway construction | ▪ Between 45234100 and 45234240: Construction work for railways and cable transport systems<br>▪ 45221112: Railway bridge construction work<br>▪ 45213321: Railway station construction work<br>▪ 45213320: Construction work for buildings relating to railway transport<br>▪ 45221242: Railway tunnel construction work<br>▪ **Equipment:**<br>▪ 34940000 : Railway equipment<br>▪ 34941000 : Rails and accessories<br>▪ 34942000 : Signalling equipment<br>▪ 34943000 : Train-monitoring system<br>▪ 34944000 : Points heating system<br>▪ 34945000 : Track-alignment machinery<br>▪ 34946000 : Railway-track construction materials and supplies<br>▪ 50225000 : Railway-track maintenance services<br>▪ 71311230 : Railway engineering services<br>▪ Any architectural works (beyond 71000000) if additional CPV code is related to rail construction (between 45234100 and 45234240) |
| Airport construction | ▪ Between 45235000 and 45235320: Construction work for airfields, runways and manoeuvring surfaces |

| | |
|---|---|
| | ▪ 45213331: Airport buildings construction work<br>▪ 45213332: Airport control tower construction work<br>▪ **Equipment:**<br>▪ 34960000 : Airport equipment<br>▪ 34961000 : Baggage-handling system<br>▪ 34962000 : Air-traffic control equipment<br>▪ 34963000 : Instrument Landing System (ILS)<br>▪ 34964000 : Doppler VHF Omni direction Range (DVOR)<br>▪ 34965000 : Distance Measuring Equipment (DME)<br>▪ 34966000 : Radio Direction Finder and Non Directional Beacon<br>▪ 34967000 : Airport Communication System (COM)<br>▪ 34968000 : Airport Surveillance System and Lighting System<br>▪ 34969000 : Passenger boarding bridges and stairs for aircraft<br>▪ 34995000 : Lighting for aircraft guidance and illumination<br>▪ 34997000 : Control, safety or signalling equipment for airport<br>▪ 34997100 : Flight recorders<br>▪ 34997200 : Airport lighting<br>▪ 34997210 : Runway lights<br>▪ 34999200 : Aerial signal splitters<br>▪ **Engineering:**<br>▪ 71311240: Airport engineering services<br>▪ 72212120: Flight control software development services<br>▪ 72212121: Air traffic control software development services<br>▪ 72212130: Aviation ground support and test software development services<br>▪ 72212131: Aviation ground support software development services<br>▪ 72212132: Aviation test software development services |
| Water transport | ▪ Between 45240000 and 45248500: Construction work for water projects<br>▪ **Equipment:**<br>▪ 34930000 : Marine equipment<br>▪ 34931000 : Harbour equipment<br>▪ 34932000 : Radar sets<br>▪ 34933000 : Navigation equipment<br>▪ 34934000 : Propeller blades<br>▪ 34994000 : Lighting for ship guidance and illumination<br>▪ 34994100 : Lighting for river guidance and illumination<br>▪ 34998000 : Control, safety or signalling equipment for port installations<br>▪ **Repair**<br>▪ 50240000 : Repair, maintenance and associated services related to marine and other equipment<br>▪ 50241000 : Repair and maintenance services of ships<br>▪ 50242000 : Conversion services of ships<br>▪ 50243000 : Demolition services of ships<br>▪ 50244000 : Reconditioning services of ships or boats<br>▪ 50245000 : Upgrading services of ships<br>▪ 50246000 : Harbour equipment maintenance services<br>▪ **Architectural/engineering**<br>▪ 71631460 : Dam-inspection services |

## Appendix D: Additional regression tables

**Table D1.** OLS regressions explaining the share of 3 largest infrastructure contracts in total transport infrastructure spending using the transport investment specific CRI indicator, 2009-2014, region-year-sector observations

| Model Nr. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Model descr. | no controls | infr. type control | infr. type, country & year control | infr. type, country, year, gdp, # of contracts control |
| Dependent variable: share of largest 3 contracts | | | | |
| avg. regional CRI | 0.174** | 0.156** | 0.181** | 0.156* |
| | (0.017) | (0.034) | (0.050) | (0.099) |
| Public works share | 0.170*** | 0.165*** | 0.169*** | 0.165*** |
| | (0.002) | (0.005) | (0.000) | (0.000) |
| Ref.cat.: Road | | | | |
| Rail | | 0.0309 | 0.0109 | -0.0142 |
| | | (0.400) | (0.733) | (0.679) |
| Airport | | 0.00893 | -0.0227 | -0.0416 |
| | | (0.924) | (0.803) | (0.672) |
| Water transport | | 0.0216 | 0.0168 | 0.00190 |
| | | (0.481) | (0.543) | (0.957) |
| # of contracts | | | | -0.00560*** |
| | | | | (0.005) |
| Observations | 497 | 497 | 497 | 457 |
| R-squared | 0.069 | 0.072 | 0.259 | 0.350 |

Notes: p-values in parentheses; * p<0.1, ** p<0.05, *** p<0.01; standard errors clustered by country

**Table D2.** Fixed-effects panel regressions explaining the share of 3 largest infrastructure contracts in total transport infrastructure spending using the economy-wide CRI indicator, 2009-2014, region-year-sector observations, region-sector fixed-effects applied

| Model Nr. | (1) | (2) | (3) |
|---|---|---|---|
| Dependent variable: share of largest 3 contracts | | | |
| avg. regional CRI | 0.478** | 0.488** | 0.389* |
| | (0.007) | (0.010) | (0.025) |
| Public works share | 0.163 | 0.173 | 0.121 |
| | (0.083) | (0.063) | (0.195) |
| # of contracts | | | -0.00559 |
| | | | (0.051) |
| Year control | | x | x |
| # of contracts | | | x |
| Observations | 493 | 493 | 493 |
| R-squared | 0.071 | 0.086 | 0.177 |

Notes: p-values in parentheses; * p<0.05, ** p<0.01, *** p<0.001; standard errors clustered by country

**Table D3.** Multilevel regression model explaining relative price (final/estimated price) using the economy-wide CRI indicator, 2009-2014

| Model Nr. | (1) |
|---|---|
| Dependent variable: relative price | |
| CRI | 0.162*** |
| | (0.000) |
| Ref.cat.: Road | |
| Rail | 0.017 |
| | (0.092) |
| Airport | 0.016 |
| | (0.339) |
| Water_transport | -0.038*** |
| | (0.000) |
| Contract value | X |
| Buyer sector | X |
| Buyer type | X |
| Year | X |
| *Random-effects parameters* | |
| CRI | 0.0167 |
| | se: 0.0073 |
| Constant | 0.0067 |
| | se: 0.0021 |
| Residual | 0.0266 |
| | se: 0.004 |
| N | 7774 |
| Log likelihood | 2955.61 |
| R-squared | 0.038 |

Notes: p-values in parentheses; * p<0.05, ** p<0.01, *** p<0.001; standard errors clustered by country

**Table D4.** Multilevel regression model explaining relative price (final/estimated price) using the transport investment specific CRI indicator, 2009-2014

| Model Nr. | (1) |
|---|---|
| Dependent variable: relative price | |
| CRI (transport specific) | 0.136*** |
| | (0.000) |
| Ref.cat.: Road | |
| Rail | 0.013 |
| | (0.175) |
| Airport | 0.012 |
| | (0.481) |
| Water_transport | -0.036*** |
| | (0.000) |
| Contract value | X |
| Buyer sector | X |
| Buyer type | X |
| Year | X |
| *Random-effects parameters* | |
| CRI | 0.0167 |
| | se: 0.0073 |
| Constant | 0.0067 |
| | se: 0.0021 |
| Residual | 0.0266 |
| | se: 0.004 |
| N | 7774 |
| Log likelihood | 2996.65 |
| R-squared | 0.047 |

Notes: p-values in parentheses; * p<0.05, ** p<0.01, *** p<0.001; standard errors clustered by country