INSTITUTE OF ECONOMICS
Centre for Economic and Regional Studies

HUN-REN
Hungarian Research Network

Közgazdaság- és Regionális Tudományi Kutatóközpont
KÖZGAZDASÁG-TUDOMÁNYI INTÉZET

KRTK-KTI WORKING PAPERS | KRTK-KTI MŰHELYTANULMÁNYOK

Corruption Risk and Education at Regional Level

ISTVÁN JÁNOS TÓTH – MIKLÓS HAJDU

KRTK-KTI WP – 2023/35

December 2023

INSTITUTE OF ECONOMICS, CENTRE FOR ECONOMIC AND REGIONAL STUDIES
HUNGARIAN RESEARCH NETWORK (HUN-REN)
BUDAPEST

ABSTRACT
In this study, we investigate the correlation between corruption risk and the level of education in European sub-national regions (NUTS2 level) between 2006 and 2020 in 16 member countries. We use the data of Tenders Electronic Daily (TED) covering the parameters of 6,766,274 public procurement contracts in total and NUTS2 level Eurostat data. We found that higher educational attainment is associated with lower corruption risk and a higher level of control of corruption, indicating that better-educated locals may force authorities to limit corruption risk as they have less tolerance for corrupt behavior. In addition, the results point out that the increasing level of education is associated with a decreasing level of corruption risk. Our study contributes to corruption research by using objective indicators characterizing the NUTS2 regions of some European countries.

JEL codes: D73; H57; H75, I25, R11
Keywords: regions in Europe, education, corruption risk

István János Tóth
HUN-REN KRTK KTI, SGH Warszawa and CRCB
toth.istvanjanos@krtk.hun-ren.hu

Miklós Hajdu
CUB
miklos.hajdu@uni-corvinus.hu

Korrupciós kockázat és iskolázottság regionális szinten

TÓTH ISTVÁN JÁNOS – HAJDU MIKLÓS

ÖSSZEFOGLALÓ
A tanulmányban a korrupciós kockázat és az oktatási szint közötti összefüggést vizsgáljuk 2006 és 2020 között 16 tagállamban NUTS2 szintű adatok alapján. A korrupciós kockázat mérésénél a Tenders Electronic Daily (TED) adatait használjuk, amelyek összesen 6.766.274 közbeszerzési szerződés paramétereire terjednek ki, valamint az Eurostat NUTS2 szintű adatait. Eredményül azt találtuk, hogy a magasabb iskolai végzettség alacsonyabb korrupciós kockázattal és a korrupció magasabb szintű ellenőrzésével jár együtt, ami arra utal, hogy a jobban képzett helyiek erősebben rá tudják kényszeríteni a hatóságokat a korrupciós kockázat korlátozására, mivel kevésbé tolerálják a korrupt magatartást. Emellett az eredmények rámutatnak arra is, hogy a növekvő iskolai végzettség csökkenő korrupciós kockázattal jár együtt. Tanulmányunk európai országok NUTS2 régióira vonatkozó objektív mutatók alkalmazásával járul hozzá a korrupció jelenségének kutatáshoz.

JEL: D73; H57; H75, I25, R11
Kulcsszavak: korrupció, kartell, kerekített árak, Magyarország, EU támogatások

Introduction[1]

Individuals with high social status tend to be less tolerant of corruption (Bruner and Korchin, 1946; Lipset, 1960). They are more likely to reject corrupt politicians in elections. In contrast, individuals with lower social status and lower levels of education tend to be more tolerant, accepting the "*he robs, but he gets things done*" principle when they see the politicians competent (Winters and Weitz-Shapiro, 1996; Marínez Rosón, 2016). Consequently, increased educational attainment within the population is expected to be associated with less corruption. The paper by Glaeser and Saks based on US data supports this correlation (Glaeser and Saks, 2005).

This paper aims to examine the relationships between educational attainment and corruption. We use regional (NUTS2) data from 16 EU countries. The paper investigates whether regions with a higher percentage of university graduates produce lower levels of corruption. Additionally, the analysis explores whether an increase in the proportion of university graduates in a region is associated with lower levels of corruption.

We analyze the impact of education on the level of corruption risk using public contract, social and economic data. We use hard, objective data instead of perception indices of corruption, and we focus on subnational differences rather than the cross-national comparison of corruption risk. Furthermore, as we deal indicators of or proxies for institutional quality in terms of integrity or the effectiveness of fighting corruption, the present study may not solely contribute

[1] This research was made possible by a grant from the Hungarian Scientific Fund (OTKA, K-135598). We gratefully acknowledge the comments by Mihály Fazekas, Gábor Kertesi, György Molnár, Ádám Reiff, Balázs Váradi, and the participants of the 'Social attitudes' section of the 2022 annual conference organized by the Hungarian Society of Economics.

to the domain of public procurement corruption research, but also may also lead to important conclusions from the perspective of the institutional conditions of the economic convergence.

Public procurement corruption risk is present if conditions of the tender favor the rent extraction from public procurement in an institutionalized way, more precisely they allow winners of the tenders to be pre-selected (Fazekas et al, 2014; Fazekas et al., 2016b). This can be done in three primary forms related to different phases of procurement and can be certainly combined: limiting the set of applicants to the tenders during the submission phase (i); unfair assessment of the applications during the evaluation period (ii); ex-post modifications of the contracting conditions in the delivery phase (iii). For example, if there is only one bidder or the tender is not open for every potential applicant such endeavors are easier to implement, however, the presence of such circumstances does not indicate automatically that corruption happened, as it can be the result of chance or intentions but without the idea of corruption. Also, even if these conditions are present, because the actors would like to perform a corrupt transaction, it is possible, that finally the corrupt transaction does not happen. Nonetheless, the systematic occurrence of certain characteristics of public procurement contracts can indicate the risk of corruption.

The empirical analysis of control corruption in cross-national context was begun based on the corruption perception indices prepared by business risk analysts and polling companies (Treisman, 2000; Burguet, Ganuza and Garcia Montalvo, 2016). One of the main sources of data characterizing the level of corruption in different countries is the Corruption Perception Index (CPI) of Transparency International (Transparency International, 2017). The yearly publication of the CPI started in 1995, and its latest values were calculated for 176 countries. CPI is a widely-used tool by scholars, journalists, and policy-makers for assessing the extent of corruption, even though it has several weaknesses leading to controversial results and interpretations (Sik, 2002; Heywood and Rose, 2013; Barrington, 2014). Also, the methodology

of the CPI has been revised several times, which affects the comparability of its values over time (Rohwer, 2009).

The Control of Corruption Index of Worldwide Governance Indicators (WGI) reported by the World Bank also includes data concerning the perception of corruption (The World Bank, 2017). The project covered more than two hundred countries since 1996, and its indicators were also constructed based on multiple perception-based data sources, like surveys of firms and households, subjective assessments of commercial business information providers, NGOs, multilateral organizations and public-sector bodies (Kaufmann, Kraay and Mastruzzi, 2011). Certainly, this index was also affected by several methodological issues (Kaufmann, Kraay and Mastruzzi, 2006).

Some important points of the general criticism towards these subjective perception indices are that perceptions may or may not be linked to the experience (Sik, 2002; Thompson and Shah, 2005; Rose and Peiffer, 2012); they may be distorted by developments on more broader domains, for example, by economic growth (Kurtz and Schrank, 2007); or because respondents who are taking part in corruption may be motivated to underreport its extent, or those who are not involved lack accurate information (Golden and Picci, 2005); and also instead of relying on own experiences, the respondents may formulate their opinions based on the media coverage of corruption cases (Lambsdorff, 2007).

In the past decade, the need arose for alternative methods capturing the control of corruption based on objective, albeit indirect data resulting in composite national indices grasping several characteristics of countries that may be relevant from the perspective of integrity or corruption, like administrative burden, enabling competition, budget transparency, social accountability, press freedom and independence of the judiciary (Mungiu-Pippidi and Dadašov, 2016). Fazekas, Tóth, and King also discussed a new, objective method in assessing the presence of corruption, called corruption risk indicators (Fazekas, Tóth and King, 2014). Such indices are

constructed by identifying ‘red flags’ at certain points of a purchase procedure that restrict transparency (Kenny and Musatova, 2010). The methodology concerning the corruption risk and competition intensity indicators is proven to be a fruitful field for research on the domains of public procurement (Fazekas, Tóth and King, 2014, 2016; Dávid-Barrett and Fazekas, 2016; Fazekas and Tóth, 2016b; Broms, Dahlström and Fazekas, 2017; Palguta and Pertold, 2017; Szűcs, 2017; Tóth and Hajdu, 2017; Abdou et al., 2022). Analysing public procurement corruption risk also leads to abundant results on the level of municipalities, e.g., the local educational attainment and the transparency of the tenders seems to be correlated, however, risk indicators may signal different issues apart from corruption too (Hajdu, 2022).

Numerous studies have aimed to analyze the relationship between corruption and certain economic, social, and political indicators on macro level (Dimant and Tosato, 2018). Most of these papers focus on the consequences of corruption; the ones that study the causes of corruption based on empirical evidence seem to be rare. Although it has to be emphasized that these causal links are not always clear. Empirical research dealing with country-level data suggests that the level of corruption is lower in a country where the population is more educated (Treisman, 2000). Components of modernity are also correlated with the control of corruption: low life expectancy, increased rural population and low educational attainment all significantly predict increased likelihood of corruption (Mungiu-Pippidi, 2015). Other research analyzing the corruption at the local level concluded similar results: the strong social capital predicts a low level of corruption (Wachs *et al.*, 2019).

The problem of white-collar crime, however, may have an opposite effect on public procurement corruption as individual capabilities are needed to commit fraud (Rustiarini *et al.*, 2019), but empirical investigations on this question are scarce (Smith, 2022); nevertheless, it is observed that fraud (mostly with a large nominal value) does not occur if the committer does not have the right abilities (Wolfe and Hermanson, 2004). The tolerance of corruption, however,

decreases in better-educated groups of individuals (Mungiu-Pippidi, 2015), since general law abidance tends to decline with a higher level of education. These are individuals who prefer to decide what is right without the guidance of regulations (Orviska and Hudson, 2003).

On the one hand, concerning tax evasion, there seems to be a consensus among the majority of researchers that higher education enhances taxation knowledge, contributing to a better understanding of taxation in terms of laws and regulations and mitigating tax fraud (Alshira'h *et al.*, 2020). A potential reason for these findings is that wealthier taxpayers are likely to be more educated than the general population and may have more respect for the rule of law. On the other hand, wealthier people are taxed more than poor people and may resent paying so much in taxes, causing them to view tax evasion more approvingly and consciously (McGee, 2012).

Regarding the consequences of corruption, it has been pointed out that higher perceived corruption is linked to lower investments and economic growth (Mauro, 1995). Corruption also has a negative impact on the efficacy of public spending in education (Suryadarma, 2008) and on enrollment rates (Dridi, 2014). Furthermore, countries with higher levels of corruption are inclined to have a larger shadow economy, or in other terms, an unofficial economy (Johnson, Kaufmann and Zoido-Lobaton, 1998) and public debt (Cooray, Dzhumashev and Schneider, 2016). Additionally, there is a distinctive impact on post-socialist countries following the transition: corruption had the potential consequence of inhibiting the consolidation of democratic institutions and open market economies (Shleifer, 1997).

Important limitations in studies investigating the relationship between perception or the risk of corruption and socioeconomic features of territorial units stem from unclear causal relations during quantitative analyses. Furthermore, the majority of papers aiming to establish a correlation between corruption and specific macro indicators typically concentrate on economic characteristics rather than socio-demographic features. In an article by Treisman that assesses

the causes of corruption in a cross-national context (Treisman, 2000), the issue of omitted variable bias and endogeneity is raised. Treisman also considers Ordinary Least Squares (OLS) regression as an essential starting point that needs enhancement through techniques aimed at exploring the direction of causation. Due to the absence of proper instruments, the instrumental variable approach was only partially applicable in testing hypotheses; the distance from the Equator appeared to be a suitable instrument for log per capita GDP, allowing the assessment of the link between economic development and corruption through an Instrumental Variables (IV) estimation.

Cooray, Dzhumashev, and Schneider employed latitude and settler mortality rate as instruments for corruption in order to correct their results for endogeneity (Cooray, Dzhumashev and Schneider, 2016). These are commonly used variables for Instrumental Variables (IV) estimations in corruption research. It was necessary for the instruments to be correlated with corruption and not influence public debt through other channels.

Treisman also run a series of nested regressions beginning with the most plausibly exogenous variables and attempting to move down the causal chain by including more and more variables (Treisman, 2000). More practically, this means that he began with the inclusion of long-predetermined historical, cultural or ethnic parameters, like the legal system, colonial heritage, religious affiliation, ethnolinguistic fragmentation and natural resource endowments. Then, he created four further estimations with the involvement of more and more explanatory variables, which are increasingly endogenous (for example, the frequency of turnover in government leadership). In addition, Treisman repeated the estimations for several different corruption perception indices (which are strongly correlated with each other) and he developed and tested several alternative indicators for the independent variables to check the robustness of his results. Furthermore, weighted least squares estimations were carried out, weighting cases by the inverse of the variance of ratings for that country on the corruption perception indices to be

more focused on those countries which obtained more similar (and thus presumably more reliable) ratings.

The article by Paolo Mauro cited earlier (Mauro, 1995) also raises the issue of endogeneity, however, it considers the level of corruption as an explanatory variable for the economic growth of countries. The ethnolinguistic fractionalization that Treisman used as an exogenous independent variable in his nested regressions is applied as an instrument in the two-stage least squares (2SLS) estimations conducted by Mauro. The reason why he found it a good instrument is that it is in negative and significant correlation with institutional efficiency and corruption but is unrelated to the economic characteristics of the investigated countries other than through its effects on the explanatory variables.

Overall, it is a challenging methodological issue to find a proper way in handling the uncertain causal directions and the endogeneity; the review of the papers assessing the causes and the effects of corruption suggests that there is no obvious solution for the problem, but there are several possible approaches with different advantages and disadvantages.

Unfortunately, some of the independent variables of the present research are not available even on a yearly basis, thereby the methods requiring panel data cannot be implemented. The strategy of Treisman, however, can be at least partly followed, certain kinds of panel regressions can be run with time-invariant variables also. Furthermore, as the observations can be assigned to distinct locations, regressions on a dataset aggregated to the level of NUTS2 regions can also be run. The results deriving from different approaches can verify and reinforce each other, but at the same time, if ambivalent outcomes turn out, then it may raise uncertainties concerning the findings.

Our paper is based on the EU Tender Electronic Daily (https://data.europa.eu/data/datasets/ted-csv?locale=en) dataset of public procurement contracts from the period between 2006 and 2020

and the regional socioeconomic data gathered by the Eurostat. The paper is structured as follows. Section 2 and 3 describe the dataset used for the empirical analysis and the empirical strategy. Section 4 presents the results and Section 5 concludes the paper. In the Appendix we demonstrate the relationship between the corruption risk indicators and the corruption perception indicators in order to demonstrate the validity of the approach used in the study.

2 Data

In the present analysis we take into consideration the NUTS2 regions of 16 member-countries of the European Union: Austria, Bulgaria, Cyprus, Czechia, Estonia, France, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia. The data covers the period between 2006 and 2020. Firstly, data on public procurement corruption risk is obtained from the Tenders Electronic Daily (TED) covering the parameters of 6,766,274 public procurement contracts in total. Secondly, regional data on educational attainment is obtained from the Eurostat. We take into consideration the proportion of population aged 25 and 64 completed tertiary education (levels 5-8) according to the EDAT_LFSE_04 dataset, as some existing empirical research found that this level of education may account for a considerable level of variation observed in fraudulent activities (Babic and Zarić, 2022) and there is a great variation between the European countries regarding their education systems on the lower levels, e.g. in terms of the organizational models and length of the compulsory education (European Commission et al., 2021).

In the analysis we use indicators to characterize the intensity of competition and the corruption risk of tenders as dependent variables. The first indicator refers to competition: it presents single-bidder contracts, which is an essential indicator of corruption risk or, in other words, of the conditions facilitating corruption. The World Bank and the EU Commission consider the

occurrence of a single bidder in public procurement as a red flag (European Commission, 2014, 2020; The World Bank, 2017). In addition, several studies have analyzed the share of contracts with a single bidder as an objective indicator of corruption risk (Gagliarducci and Coviello, 2010; Fazekas *et al.*, 2013; Fazekas and Tóth, 2016a; Tóth and Hajdu, 2016).

For measuring the prevalence of single-bidder contracts, we use an indicator called 'Corruption Risk' (CR) using the following rule:

CR = 0, if the tender was conducted with more than one bid,

= 1, if there was only one bid.

This approach considers public procurement corruption as the result of dyadic connections, however, in many cases, the CR variable does not adequately reflect corruption risk as the European Union and the Hungarian Public Procurement Act ([Act CXLIII of 2015 on public procurement)] also prescribes in certain cases that the contracting authority must invite at least three tenderers to submit a tender, e.g., if a negotiated procedure without prior publication is applied:

"(3*), If the negotiated procedure is conducted under sections 98 (2) b) and 98 (4) a) or if it is reasonably possible under the circumstances causing extreme urgency, in the cases specified in section 98 (2) e) contracting authorities shall invite at least three tenderers to tender as possible."*

Due to these rules, corrupt actors could potentially organize two fraudulent (losing) bids, thereby meeting the legal requirements: *'You will win, just bring two losers!*' (Tóth and Hajdu, 2022). In cases of institutionalized corruption, contracting authorities themselves could organize fraudulent bids or imitate competition by meeting formal criteria, such as having at least three bidders in the tender.

This form of corruption could also involve companies acting as 'losing' bidders in tenders, allowing the contracting authority's pre-selected company to win. In such situations, the corruption risk of contracts with three bidders does not differ significantly from tenders with a single bidder. Therefore, it is essential to observe the proportion of tenders carried out with at least four bidders. Consequently, we created an indicator based on the number of bidders to distinguish the contracts with more than three bidders (CoCR).

CoCR = 0, if the tender was conducted with no more than three bids,

=1, if there were at least four bidders for the contract.

We interpret the CoCR as a proxy for control of corruption risk. If the share of tenders with more than three bidders is high, the proportion of independent competitors is also high, which means more robust control of corruption. There is a minimal incentive for corrupt actors to organize 3 three or more losing ‘bidders’ when organizing three formally independent bidders is enough to meet the formal requirements.

3 Empirical Strategy

The empirical strategy of our paper is based on the correlation shown by previous empirical research that higher levels of education are associated with lower levels of corruption (Glaeser and Saks, 2006; Mungiu-Pippidi, 2015). We have seen from our country-level data that this relationship holds even when considering the impact of economic development, GDP per capita (see Annex 1.). We have also seen that this correlation holds not only for perceptual corruption indicators (CPI_INV or CoC) but also for the correlation between corruption risk indicators based on objective data (CR or CoCR) and the level of education (see Annex 2.).

We have separately analyzed the relationship between (i) corruption risk and levels of education and (ii) the relationship between changes in the levels of these two factors. Previous studies

have focused only on the relation of level of corruption and education. However, it is also worth investigating whether an increase in the educational attainment of the population in a region is associated with a more robust control of corruption risk in public procurement in that region, i.e., a reduction in corruption measured by the risk of corruption in public procurement. Such a dynamic effect would complement the results of existing cross-sectional studies.

For the model specifications, it is essential to consider that the composition of public procurement may differ across regions according to the type of funding (whether EU or national sources finance it). The results of previous studies (Fazekas et al. 2013; Fazekas, Tóth, 2017; Palócz, Tóth, 2022) demonstrate that EU-funded tenders have a higher corruption risk than those financed by national sources. The effect of the size of contracts is also worth considering: smaller tenders have a higher corruption risk (Tóth and Hajdu, 2017; Tóth and Hajdu, 2022).

The different sizes of public procurement markets across regions should also be considered in the estimates. We measure this effect by the number of tenders conducted by year in a region, which varies considerably across regions (see Figure A1.2g). Given the same institutional conditions, it is more difficult to control corruption risk where more tenders are conducted in a period.

In addition to taking into account the characteristics of the public procurement market, it is also essential to consider the region's economic development to observe the independent effect of educational attainment. We measure it with GDP per capita by NUTS2 regions. Furthermore, since we are looking at data for the regions from 2006 to 2020, we have filtered out the effects associated with each year using a year dummy.

Analysis of individual-level data shows that younger age groups are more tolerant of corruption and reject it less than older age groups (see Annex 2). Accordingly, younger people are expected to be less likely to take action against corruption. It is, therefore, worth reflecting on the

proportion of younger age groups in a given region. If this share is high, we expect the corruption risk to be higher. We put into the regressions the share of the 20-40-year-old population in the total population living in the region.

According to research on good governance the low state capacity is associated with higher corruption (Rothstein and Teorell, 2008; Bersch et al, 2017). The size of the public sector can have a contradictory effect on state capacity and, hence, on the level of corruption (Bersch et al., 2017); on the one hand, the larger the number of public sector employees, the more likely it is that specialized expertise that allows for more substantial control of corruption will emerge within the sector, thus increasing state capacity (LaPorta et al., 1999) and lowering corruption. On the other hand, a relatively higher public sector may increase the incentives for corruption or likelihood of corrupt activities (Glaeser and Saks, 2006). To filter out these possible effects, we have included in the regression estimates the weight of the public sector in a given region, measured by the share of public sector employees in total employment.

We run four estimates using regional-level data. Estimates complemented these run on contract-level data (see Annex 5 for the latter results). For two indicators of corruption risk (CR and CoCR), we examined the relationship between the level of corruption risk and the level of education (A3.1.1 and A3.1.2.) and the relationship between the change in their level (A3.2.1. and A3.2.2.):

$$lnCR_{it} = \alpha + \beta lnEDU_{it} + \sum_{t=2006}^{2020} \gamma_t X_{it} + \sum_{t=2006}^{2020} \delta_t YEAR_t + \varepsilon_{it} \quad \text{(A3.1.1)}$$

$$lnCoCR_{it} = \alpha + \beta lnEDU_{it} + \sum_{t=2006}^{2020} \gamma_t X_{it} + \sum_{t=2006}^{2020} \delta_t YEAR_t + \varepsilon_{i,t} \quad \text{(A3.1.2)}$$

and

$$dCR_{ip} = \alpha + \beta dEDU_{ip} + \sum_{p=1}^{4} \gamma_p X_{ip} + \sum_{p=1}^{n} \delta_p PER_p + \varepsilon_{i,p} \quad \text{(A3.2.1)}$$

$$dCoCR_{ip} = \alpha + \beta dEDU_{ip} + \sum_{p=1}^{4} \gamma_p X_{ip} + \sum_{p=1}^{n} \delta_p PER_p + \varepsilon_{i,p} \quad \text{(A3.2.2)}$$

In equations A3.1.1. and A3.1.2., $lnCR_{it}$ is the logarithm of the share of public tenders without competition, $lnCoCR_{it}$ is the logarithm of the share of public tenders with at least four bidders, $lnEDU_{it}$ is the logarithm of the share of the population with at least tertiary education, X_{it} is the vector of control variables in the region *i*, year *t* and β is the statistical effect of the level of education on corruption risk. We expect that for *lnCR* that β value will be negative: i.e., a higher level of education is associated with a lower level of corruption risk, while for *lnCoCR*, on the contrary, β will be positive. The higher the level of education in a region, the more likely it is that the region's public procurement market is subject to more substantial control of corruption risk by the (more educated) population. Additionally, we use an instrumental variable (IV) for education. In educational research, the distance between the individual and the educational institution is commonly employed as an instrumental variable (Pokropek, 2016). As we work with regional-level data, such specific data are unavailable. Hence, we sought an instrument that could influence an individual's access to education. Ultimately, we selected the density of the rail network as it reflects the density of passenger rail transport.

For the A3.2.1, and A3.2.2 equations, we have examined the period 2009-2020 by year and then by dividing the period 2009-2020 into four three-year periods (2009-2011, 2012-2014, 2015-2017, and 2018-2020). The idea was to distinguish between short-term (year-to-year) and medium-term effects between educational attainment and corruption risk and filter out possible short-term cyclical effects. Accordingly, we explored the analysis where n=11 - year-to-year impacts - and where n=4 (three-year period impacts). The dCR_{ip} is the change in the proportion of tenders without competition, $dCoCR_{ip}$ is the change in the proportion of tenders with at least four bidders between period *p* and period *p-1*. $dEDU_{ip}$ is the change in the proportion of the population with at least tertiary education, X_{ip} is the vector of changes in the control variables in region *i* and period *p* compared to period *p-1*. β represents the statistical effect of the change in the level of education on the change in the level of corruption risk. We expect that for *dCR*

the β will be negative: that is, an increase in education is associated with a decrease in corruption risk. In contrast, we expect β to be positive for *dCoCR*: the more the level of education increases in a region from one period to the next, the more the control of corruption risks in the region's public procurement market increases, because an increase in the proportion of more educated people also increases the proportion of those who are less tolerant of corruption.

4 Results

4.1. Regional distribution of key indicators

The regional level data underline considerable differences in corruption risk or control of corruption risk between the regions surveyed. There are regions with the lowest possible corruption risk between 2006 and 2020 (CR=0) and regions with an extremely high risk (CR=0.78). There is also a complete lack of control of corruption risk (CoCR=0) and an extremely high value (CoCR=0.9). (See Table A4.1.) Noteworthy regional differences can be seen mainly between the Western and Central-Eastern parts of Europe in terms of both corruption risk measures that seem to be persistent throughout the examined period.

However, it is also apparent that the *CoCR* was likely to decrease, and the *CR* tended to increase between 2006 and 2020 in most of the regions covered by the study (see Fig. 4.1.1a-d). The regions with the lowest corruption risk include Austrian, French, German, Spanish regions.

Spectacular differences occur in the ratio of the graduated population among the investigated regions. Regions with more educated populations are mainly found in France and Germany, while in some regions in Spain and Romania, graduates are relatively scarce (see Fig. 4.1.1.e-f). We present the detailed descriptive analyses of the regional data in Annex 4.

Fig. 4.1.1a-f here

4.2. Estimations

First, we ran estimates on the correlations between level of education and level of corruption risk. The results indicate that the higher levels of educational attainment of the population are associated with lower levels of corruption risk and higher levels of control of corruption risk (See Table 4.2.1 and Table 4.2.2.). This correlation holds even when we include the economic development and regional public procurement market characteristics (number of contracts per year, average size of contracts, and share of tenders with EU funding) in the estimation. Our results show that the one percent increase in the education level (share of college and university graduates in the regional population) reduces the level of corruption risk by 0.41-0.48 percent and increases the level of control of corruption risk by 0.31-0.37 percent.

Table 4.2.1 here

Based on regional data, these estimation results support the correlation between educational attainment and the level of corruption found in other research. Our results underline the validity of this relationship. Based on population surveys (WVS data), the educated are less tolerant of corruption and less accepting of its justification than the uneducated. (see Annex A2). In regions with a higher proportion of tertiary-educated individuals, the population tends to be less tolerant of corruption. These regions exhibit lower levels of corruption risk and more substantial control of corruption risk in the public procurement market. The results, including the instrumental variable (length of railway track per square kilometer), support the aforementioned effects: regions with a more educated population show lower corruption risk and more effective control of corruption risk (Table 4.2.1 and Table 4.2.2.).

Table 4.2.2 here

The yearly data analysis supports both the persistence of these effects and the downward trend in the strength of these effects for both corruption risk and corruption risk control (See Fig. 4.2.1a-b.).

Fig. 4.2.1a-b here

The novelty of our research lies in the fact that we investigated not only the relationship between levels of education and corruption indicators but also the relationship between changes in these levels.

This line of analysis focuses on how changes in educational attainment levels may induce institutional changes: to what extent they may contribute to a region's contracting entities conducting public procurement with lower levels of corruption risk and to more robust control of corruption risk. Of course, these effects can come from three sides: (1) from the contracting authorities (their higher human capital results in lower levels of corruption); or (2) from the competing companies for public tenders (the companies with higher educated staff could enter more the public procurement market and increase the volume of tenders and thus indirectly reduce the corruption risk in tenders than companies with lower educated staff); or (3) from the population of the region as voters. A more educated population pays more attention to corrupt phenomena and is more forceful in pushing for a curb on corruption.

The results suggest that there is a correlation between increasing educational attainment is associated with decreasing corruption. The increase in the level of education in a region is associated with a decrease in the level of corruption risk and an increase in the control of corruption risk. This effect is statistically significant but weak (See Table 4.2.3 and 4.2.4.).

Table 4.2.3 here

Table 4.2.4 here

5 Conclusions

In this paper, we analyzed how the level of education and its change impacts the level of corruption using regional data from 16 EU countries. One of the novelties of our study is that we investigate the factors influencing corruption risk on a regional level, and we use objective data instead of subjective indicators to characterize potential fraudulent activities. The other novelty is that apart from analyzing the links between the level of education and corruption risk, we investigate the links between the change in corruption risk and the change in education level. Corrupt transactions result from the actors' decisions; however, their direct observation is not feasible. Existing researches cited in the literature review support that higher educational attainment might prevent corrupt transactions.

Additionally, our study has provided empirical evidence at the national level, indicating that countries with better-educated residents are likely to face lower corruption risks in public procurement. This conclusion holds even when considering factors such as GDP, the share of EU-funded public contracts, and total net contract value. Moreover, our findings at the individual level indicate that individuals with higher levels of education are more likely to reject corruption compared to those with lower education. This not only enhances the likelihood of non-corrupt behavior but also makes it more challenging for corrupt politicians to succeed in elections, as the highly educated population are not likely to vote for candidates with a history of corruption.

As we aggregate the public procurement corruption risk data to the level of regions, its investigation may lead to a better understanding of the relationship between education and corruption. Our key findings reveal a negative correlation between corruption risk and educational attainment in the NUTS2 regions of 16 European Union member states. In addition, the results point out that the increasing level of education is associated with a decreasing level

of corruption risk. These are our contributions to corruption research, as this aggregation level is closer to those actors who may participate in or avoid corrupt transactions.

The locals' higher educational attainment limits the risk of public procurement corruption. Educated people may complain about the misbehavior of officials or politicians and thereby encourage them to avoid corruption, which can explain the link between education and the quality of government (Botero et al., 2013). More educated people are less tolerant of corruption. Moreover, education impacts democracy, the rule of law, and political liberty positively independent from wealth (GDP) as higher stages of moral judgment may be fostered by cognitive ability, which may also lead to the increased competence and willingness to seek information necessary for political decisions (Rindermann, 2008).

The impact of education can take place through several mechanisms. The first is when the more educated population can better understand corruption's adverse social and economic effects. Thus, as described above, they are less tolerant of corrupt politicians or officeholders during elections. The other mechanism is that the higher education level in a region can affect both sides of the public procurement market. A higher share of high education in the local labor market also suggests that contracting authorities can better select educated experts and, thus, reduce the risk of corruption in public procurement procedures. Furthermore, the same factor can also positively affect the supply side of the market: local firms are more able to employ highly educated workers, who are more able to submit bids to public tenders -- thus increasing the number of bids.

Nevertheless, overall, all three possible mechanisms point in the same direction: an increase in education is associated with a decrease in corruption risk.

References

Abdou, A., Basdevant, O., Dávid-Barrett, E. and Fazekas, M. (2022) *Assessing Vulnerabilities to Corruption in Public Procurement and Their Price Impact, IMF Working Papers*. WP/22/94. Washington D.C. Available at: https://doi.org/10.5089/9798400207884.001

Alshira'h, A.F. *et al.* (2020) 'A socio-economic model of sales tax compliance', *Economies*, 8(4), pp. 1–15. Available at: https://doi.org/10.3390/economies8040088

Babic, V. and Zarić, S. (2022) 'Tax Evasion and Education Level: Evidence from the European Union Countries', *Knowledge – International Journal*, 51(1), 17–21. https://ikm.mk/ojs/index.php/kij/article/view/5035

Barrington, R. (2014) *Measuring Corruption – A Discussion of Methodology*. Available at: http://www.transparency.org.uk/measuring-corruption-discussion/ (Accessed: 28 May 2017).

Bersch, K., Prac, S. and Taylor, M. (2017) 'State Capacity, Bureaucratic Politicization, and Corruption in the Brazilian State', *Governance. An International Journal of Policy, Administration, and Institutions*, 30(1), pp. 105–124. Available at: https://doi.org/10.1111/gove.12196

Botero, J., Ponce, A. and Shleifer, A. (2013) 'Education, Complaints, and Accountability', *The Journal of Law & Economics*, 56(4), pp. 959–996. Available at: https://doi.org/https://doi.org/10.1086/674133

Broms, R., Dahlström, C. and Fazekas, M. (2017) *Procurement and Competition in Swedish Municipalities, Quality of Government Working Papers Series*. Gothenburg. Available at: https://core.ac.uk/download/pdf/85145196.pdf

Bruner, J. S. and Korchin, S. J. (1946) 'The Boss and the Vote: Case Study in City Politics', *The Public Opinion Quarterly*, 10(1), pp. 1-23. Available at: https://www.jstor.org/stable/2745202

Burguet, R., Ganuza, J.-J. and Garcia Montalvo, J. (2016) *The Microeconomics of Corruption. A Review of Thirty Years of Research, Barcelona GSE Working Paper Series*. 908. Barcelona. Avialable at: https://www.iae.csic.es/investigatorsMaterial/a16209105703sp27669.pdf

Cooray, A., Dzhumashev, R. and Schneider, F. (2016) 'How Does Corruption Affect Public Debt? An Empirical Analysis', *World Development*, 90. Available at: https://doi.org/10.1016/j.worlddev.2016.08.020

Dávid-Barrett, E. and Fazekas, M. (2016) *Corrupt contracting: Controlling Partisan Favouritism in Public Procurement*. 49. Available at: https://www.againstcorruption.eu/wp-content/uploads/2016/07/WP-49-Corrupt-Contracting.pdf

Dimant, E. and Tosato, G. (2018) 'Causes and Effects of Corruption: What Has Past Decade'S Empirical Research Taught Us? a Survey'. *Journal of Economic Surveys*, *32*(2), 335–356. Available at: https://doi.org/10.1111/joes.12198

Dridi, M. (2014) 'Corruption and Education: Empirical Evidence', *International Journal of Economics and Financial Issues*, 4(3), pp. 476–493. Available at: https://www.econjournals.com/index.php/ijefi/article/view/781/pdf

European Commission (2014) *Single Market Scoreboard*. Available at: https://single-market-scoreboard.ec.europa.eu/_en

European Commission (2020) *Study to Develop an Upgraded Single Market Scoreboard as a Governance Tool for the Single Market*. Available at: https://ec.europa.eu/internal_market/scoreboard/_docs/2020/survey/final-report_en.pdf

European Commission, European Education and Culture Executive Agency, Baïdak, N., De Coster, I. and Sicurella, A. (2021) The structure of the European education systems 2021/22 – Schematic diagrams, Publications Office of the European Union, 2021. Brussels. Available at: https://data.europa.eu/doi/10.2797/972889

Fazekas, M., Chvalkovska, J., Skuhrovec, J., Tóth, I.J. and King L.P. (2013) '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*. Berlin: Barbara Budrich Publishers, pp. 68–89. Available at: https://doi.org/10.2139/ssrn.2363545

Fazekas, M. and Tóth, I.J. (2016a) 'Corruption in EU Funds? Europe-wide evidence of the corruption effect of EU-funded public contracting', in J. Bachtler et al. (eds) *EU Cohesion Policy: Reassessing Performance and Direction*. London: Routledge, pp. 186–205. Available at: https://doi.org/10.4324/9781315401867

Fazekas, M. and Tóth, I.J. (2016b) 'From Corruption to State Capture: A New Analytical Framework with Empirical Applications from Hungary', *Political Research Quarterly*, 69(2), pp. 320–334. Available at: https://doi.org/10.1177/1065912916639137

Fazekas, M., Tóth, I.J. and King, L.P. (2014) *Anatomy of grand corruption: A composite corruption risk index based on objective data*, *Discussion Papers*. MT-DP – 2014/3. Budapest. Available at: https://doi.org/10.2139/ssrn.2331980

Fazekas, M., Tóth, I.J. and King, L.P. (2016) 'An Objective Corruption Risk Index Using Public Procurement Data', *European Journal on Criminal Policy and Research*, 22(3), pp. 369–397. Available at: https://doi.org/10.1007/s10610-016-9308-z

Gagliarducci, S. and Coviello, D. (2010) *Building Political Collusion : Evidence from Procurement Auctions*. 4939. Bonn. Available at: https://docs.iza.org/dp4939.pdf

Glaeser, E.L. and Saks, R.E. (2006) 'Corruption in America', *Journal of Public Economics*, 90(6-7), pp. 1053– 1072. Available at: https://doi.org/10.1016/j.jpubeco.2005.08.007

Golden, M.A. and Picci, L. (2005) 'Proposal for a New Measure of Corruption, Illustrated with Italian Data', *Economics and Politics*, 17(1), pp. 37–75. Available at: https://doi.org/10.1111/j.1468-0343.2005.00146.x

Haerpfer, C. *et al.* (eds) (2022) *World Values Survey: Round Seven - Country-Pooled Datafile Version 5.0*. Madrid & Vienna: JD Systems Institute & WVSA Secretariat. Available at: https://doi.org/doi:10.14281/18241.20

Hajdu, M. (2022) *Corruption risks of Hungarian municipalities: Quantitative and Qualitative Analyses*. Corvinus University of Budapest. Available at: https://phd.lib.uni-corvinus.hu/1283/1/hajdu_miklos_den.pdf

Heywood, P.M. and Rose, J. (2013) 'Close but no Cigar: The measurement of corruption', *Journal of Public Policy*, 34(3), pp. 507–529. Available at: https://doi.org/10.1017/S0143814X14000099

Johnson, S., Kaufmann, D. and Zoido-Lobaton, P. (1998) 'Regulatory Discretion and the Unofficial Economy', *American Economic Review*, 88(2), pp. 387–393. Available at: https://www.jstor.org/stable/116953

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2007) *Measuring corruption: Myths and realities, Africa Region Findings & Good Practice Infobriefs*, 273. Washington, DC. Available at: http://hdl.handle.net/10986/9576

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2011) 'The Worldwide Governance Indicators: Methodology and Analytical Issues', *Hague Journal on the Rule of Law*, 3(2), pp. 220–246. Available at: https://doi.org/10.1017/S1876404511200046

Kenny, Ch. and Musatova, M. (2010) *'Red Flags of Corruption' in World Bank Projects. An Analysis of Infrastructure Contracts, Policy Research Working Paper*, 5243. Washington D.C. Available at: https://documents1.worldbank.org/curated/en/790591468321562564/pdf/WPS5243.pdf

Kurtz, M.J. and Schrank, A. (2007) 'Growth and Governance: Models, Measures, and Mechanisms', *Journal of Politics*, 69(2), pp. 538–554. Available at: https://doi.org/10.1111/j.1468-2508.2007.00549.x

Lambsdorff, J.G. (2007) *The Institutional Economics of Corruption and Reform: Theory, Evidence, and Policy*, *Cambridge University Press*. Available at: https://doi.org/10.1017/CBO9781107415324.004

LaPorta, R., Lopes-de-Silanes, F., Shleifer, A., Vishny, R. (1999) 'The quality of government'. *Journal of Law, Economics, and Organization*, 15(1), pp. 222–279. Available at: https://doi.org/10.1093/jleo/15.1.222

Martínez Rosón, M. del M. (2016). 'Prefer the Corrupt One: A Profi le of Citizens who Choose Dishonest but Competent Politicians', *Revista Española de Investigaciones Sociológicas*, (153), pp. 77-92. Available at: http://dx.doi.org/10.5477/cis/reis.153.77

Mauro, P. (1995) 'Corruption and Growth', *The Quarterly Journal of Economics*, 110(3), pp. 681–712. Available at: https://doi.org/10.2307/2946696

McGee, R.W. (2012) 'Education Level and the Ethics of Tax Evasion', in McGee, R. (eds) *The Ethics of Tax Evasion.* Springer, New York, NY, pp. 451-457. Available at: https://doi.org/10.1007/978-1-4614-1287-8_28

Mungiu-Pippidi, A. (2015) *The Quest for Good Governance: How Societies Develop Control of Corruption*. Cambridge University Press. Available at: https://doi.org/10.1017/CBO9781316286937

Mungiu-Pippidi, A. and Dadašov, R. (2016) 'Measuring Control of Corruption by a New Index of Public Integrity', *European Journal on Criminal Policy and Research*, 22(3), pp. 415–438. Available at: https://doi.org/10.1007/s10610-016-9324-z

Orviska, M. and Hudson, J. (2003) 'Tax evasion, civic duty and the law abiding citizen', *European Journal of Political Economy*, 19(1), pp. 83–102. Available at: https://doi.org/10.1016/S0176-2680(02)00131-3

Palguta, J. and Pertold, F. (2017) 'Manipulation of procurement contracts: Evidence from the introduction of discretionary thresholds', *American Economic Journal: Economic Policy*, 9(2), pp. 293–315. Available at: https://doi.org/10.1257/pol.20150511

Rindermann, H. (2008) 'Relevance of education and intelligence for the political development of nations: Democracy, rule of law and political liberty', *Intelligence*, 36(4), pp. 306–322. Available at: https://doi.org/10.1016/j.intell.2007.09.003

Rohwer, A. (2009) *Measuring Corruption: A Comparison between the Transparency International's Corruption Perceptions Index and the World Bank's Worldwide Governance*

Indicators, *CESifo DICE Report*, 7(3), pp. 42-52. Available at: https://www.ifo.de/en/publications/2009/article-journal/measuring-corruption-comparison-between-transparency

Rose, R. and Peiffer, C. (2012) *Paying Bribes to Get Public Services: A Comparative Guide to Concepts and Measures*, *CSPP Studies in Public Policy*. 494. Glasgow. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2180793

Rothstein, B., and Teorell, J. (2008) 'What Is Quality of Government? A Theory of Impartial Government Institutions', *Governance. An International Journal of Policy, Administration, and Institutions*, 21(2), pp. 165–190. Available at: https://doi.org/10.1111/j.1468-0491.2008.00391.x

Rustiarini, N.W. *et al.* (2019) 'Why people commit public procurement fraud? The fraud diamond view', *Journal of Public Procurement*, 19(4), pp. 345–362. Available at: https://doi.org/10.1108/JOPP-02-2019-0012

Shleifer, A. (1997) 'Government in transition', *European Economic Review*, 41(3–5), pp. 385–410. Available at: https://doi.org/10.1016/S0014-2921(97)00011-1

Sik, E. (2002) 'The Bad, The Worst and the Worst', in S. Kotkin and A. Sajó (eds) *Political Corruption in Transition. A Skeptic's Handbook*, New York: Central European University Press, pp. 91–114.

Smith, C.D. (2022) 'Public Procurement Fraud', in *White-Collar Crime and the Public Sector: An Interdisciplinary Approach to Public Procurement Fraud*. Cham: Springer International Publishing, pp. 23–31. Available at: https://doi.org/10.1007/978-3-031-12567-6_4

Suryadarma, D. (2008) *Corruption, Public Spending, and Education Outcomes: Evidence from Indonesia*. Available at: http://dx.doi.org/10.2139/ssrn.1118323

Szűcs, F. (2017) *Discretion and Corruption in Public Procurement*. Available at: https://events.ceu.edu/sites/default/files/szucsjmp.pdf

The World Bank (2017) *Worldwide Governance Indicators*. Available at: http://info.worldbank.org/governance/wgi/#home

Thompson, T. and Shah, A. (2005) *Transparency International's Corruption Perceptions Index: Whose Perceptions Are They Anyway*. Washington, D.C.

Tóth, I.J. and Hajdu, M. (2017) *Intensity of Competition, Corruption Risks and Price Distortion in the Hungarian Public Procurement – 2009-2016*. Budapest. Available at: https://www.crcb.eu/wp-content/uploads/2017/12/eu_hpp_2016_report_170616_.pdf

Tóth, I.J. and Hajdu, M. (2018) *Magyar kórházak közbeszerzései a közgazdasági elemzés boncasztalán*. Budapest. Available at: http://www.crcb.eu/wp-content/uploads/2018/01/hospital_2017_paper_180122_.pdf

Tóth, I.J. and Hajdu, M. (2022) 'Cronyism in the Orbán Regime: An Empirical Analysis of Public Tenders, 2005–2021', in M. Csanádi et al. (eds) *Dynamics of an Authoritarian System: Hungary, 2010-2021*. Budapest: Central European University Press, pp. 230–274. Available at: https://doi.org/https://doi.org/10.7829/j.ctv2cw0s7k.9

Transparency International (2017) *Perceptions Index 2016*. Available at: https://www.transparency.org/news/feature/corruption_perceptions_index_2016

Treisman, D. (2000) 'The causes of corruption: a cross-national study', *Journal of Public Economics*, 76(3), pp. 399–457. Available at: https://doi.org/10.1016/S0047-2727(99)00092-4

Wachs, J. *et al.* (2019) 'Social capital predicts corruption risk in towns', *Royal Society Open Science*, 6(4). Available at: https://doi.org/10.1098/rsos.182103

Wolfe, D.T. and Hermanson, D.R. (2004) 'The Fraud Diamond: Considering the Four Elements of Fraud', *The CPA Journal*, 74(12), pp. 38–42. Available at: https://digitalcommons.kennesaw.edu/cgi/viewcontent.cgi?article=2546&context=facpubs

Winters, M. S. and Weitz-Shapiro. R. (2013) 'Lacking Information or Condoning Corruption: When Do Voters Support Corrupt Politicians?', *Comparative Politics*, 45(4), pp. 418-436. Available at: http://www.jstor.org/stable/43664074

Annex

A1. Country-Level Analysis

A1.1. Data

In the present section, we demonstrate the correlation between our corruption risk indicators (CR and CoCR) and the perception indicators, the Corruption Perception Index (CPI) of the Transparency International and Control of Corruption[2] (CoC) of the World Bank and do a country level analysis about the relationship between public procurement corruption risk (CR), control of corruption risk (CoCR) and the educational attainment in order to underpin the validity of the approach used in our study.

During this country analysis, we used several data sources from The World Bank, Transparency International, Eurostat, and 'TED - Tenders Electronic Daily' of the European Union. We list the variables and their definitions in Table A1.1.1., and their sources and used original data files in Table A1.1.2.

Table A1.1.1 here

Table A1.1.2 here

A1.2. Descriptive statistics

The Table A1.2.1., and Table A1.2.2. contain descriptive statistics for the variables analyzed. Fig. A1.2.1. shows their histograms, and Table A1.2.3. shows their pairwise correlations.

Table A1.2.1 here

Table A1.2.2 here

[2] We calculated with its inverse since its original version ranges from the highly corrupt cases to the very clean ones.

Fig. A1.2.1a-f here

Fig. A1.2.2a-f here

Table A1.2.3 here

The results of pairwise correlations (Table A1.2.3.) show that the corruption perception index (CPI_INV) and corruption risk (CR) are negatively correlated and the corruption control indicators - the CoCR and CoC - are positively correlated with the level of education (the coefficients are -0.29, -0.59, 0.23, and 0.58, respectively). The CPI_INV is negatively correlated with World Bank's CoC characterizing how the fraudulent activities are limited in a country and is in positive relationship with the gauge of the Transparency International indicating the perceived corruption. The CoCR is in positive relationship with the CoC and in negative relationship with the inverse of Corruption Perceptions Index (CPI_INV), as it might be expected. It is also worth noting that there is a nearly deterministic correlation between the World Bank's and the Transparency International's measures (CoC and CPI_INV) – the coefficient is -0.97 – accounting for the similar absolute values of the coefficients in the table below. This is partly natural, as CoC is one of sub-indicators of CPI. There is also a close relationship between the two objective corruption indicators (CR and CoCR): the coefficient is -0.87. The scatterplots of the main variables are shown in Fig. 1.2.3.

In Table A1.2.3 the correlations are indicated between the indicators based on the public procurement performance aggregated to the level of countries and the Control of Corruption and Corruption Perceptions Index variables.

Fig. A1.2.3 here

The lnEU, lnNCV, and lnGDP are closely related. The share of contracts supported by the EU is significantly lower in developed countries, and the average size of contracts is more significant than in less developed countries. This fact should be taken into account in the model specifications.

Table A1.2.4 here

A1.3. Empirical Strategy and Results

We also run regression models explaining the Corruption Risk (CR) and the Control of Corruption Risk (CoCR) indicators run on the country-level dataset according to the following specification:

$$lnY_i = \alpha + \beta lnEDU_i + \lambda X_i + \varepsilon_i, \qquad \text{(A1.1)}$$

where 'i' identifies the countries and *X* the vector of control variables.

The model also supports the finding that higher educational attainment is negatively correlated with public procurement corruption risk (CR), even if the GDP is taken into consideration in the analyses as a control variable, furthermore, the educational attainment seems to be a more important predictor than the GDP according to results below (see Table A1.3).

Table A1.3 here

A2. Education and Refusal of Bribery: The World Value Survey (WVS) Data

To analyze the correlation between the level of education and the rejection of corruption in individual level, we use data from waves 3rd, 5th , 6th , and 7th of the World Value Survey (WVS) (Haerpfer *et al.*, 2022) the required variables are available in these data sets. The WVS data was downloaded from: https://www.worldvaluessurvey.org/WVSContents.jsp.

In the WVS, there is no question on the rejection of corruption in general, but there is one question on the rejection of one type of corruption, bribery. The rejection of bribery is a good proxy concerning the rejection of corruption in general.

The description of the variables in each wave is contained in the WVS documentation file (F00003844-

WVS_Time_Series_List_of_Variables_and_equivalences_1981_2022_v3_1.xlsx) downloaded from https://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp.

The following variables are used in the analysis (in brackets are the code and name of the variable, which is given in the WVS documentation file):

- REFBRIBE: refusal to accept a bribe (F117, Justifiable: Someone accepting a bribe),
- EDU: educational attainment, at least tertiary level (X025, Highest educational level attained),
- SEX (X001, Sex),
- AGE (X003, Age),
- SSIZE size of the settlement where the respondent lives (X049, Settlement size),
- SIC: estimated income status (X047_WVS, Scale of incomes),
- YEAR: year of survey (S020, Year survey).

The WVS data were analyzed in two ways: firstly, for all countries included in waves 3rd and 5-7th (N=345,636) and secondly, only for the 16 countries for which data were available in the WVS and which were included in our regional-level analysis (N = 18,372).

The respondent's age was recoded into eight categories (*AGECAT*) as follows: 18-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, and 81 years and over.

This was the original question (F117) in the questionnaire:

"Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this card. (Read out and code one answer for each statement):

	Never justifiable									*Always justifiable*
Someone accepting a bribe in the course of their duties	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

The variable F117 has been recoded as follows:

$$REFBRIBE = 11 - F117.$$

So, the *REFBRBE* can take a value between 1 and 10, and a higher value indicates a stronger rejection of the bribe.

The majority (71%) of respondents strongly oppose accepting bribes as justifiable (see Fig. A2.1a.).

Fig. A2.1a-d here

We have created a variable highlighting the dichotomy of total rejection and incomplete rejection:

$$REFBRIBED = 1 \; if \; REFIBRIBE = 10 \; and$$

$$REFBRIBED = 0 \; if \; REFIBRIBE < 10.$$

When looking at the differences between total refusal (REFBRIBE = 10) and incomplete refusal (REFBRIBE < 10), we see that those with higher education are slightly more likely to completely refuse to accept bribes than those with lower education (72.4% vs. 70.3%). Women are less tolerant than men in this respect (72.1% vs. 69.5%). See Table A2.1.

Table A2.1 here

We examined whether differences in educational attainment persist when considering sex, age, subjective income status, the respondent municipality size, and the survey year. We estimated the following equation for every *i* respondent:

$$REFBRIBE_i = \beta_1 EDU_i + \beta_2 SEX_i + \beta_3 SIC_i + \beta_4 AGECAT_i + \beta_5 SSIZE_i + \\ + \beta_6 YEAR_i + \varepsilon_i \quad \text{..(A2.1.)}$$

and

$$REFBRIBED_i = \beta_1 EDU_i + \beta_2 SEX_i + \beta_3 SIC_i + \beta_4 AGECAT_i + \beta_5 SSIZE_i + \\ + \beta_6 YEAR_i + \varepsilon_i \quad \text{(A2.2.).}$$

The results suggest that the impact of educational attainment on the rejection of bribery persists even when the impacts of sex, age categories, the self-evaluation of personal income level, and settlement size are considered (see Table A2.2a-c). Respondents with a higher level of education are less likely to find bribery acceptable than those with a lower level of education. Presumably, those people less accepting of corruption are more likely to take action against it.

Table A2.2a here
Table A2.2b here
Table A2.2c here

A3. Descriptive Statistics: Contract-level Data

A total of 6,189,532 contracts from 16 countries were included as a first step in the analysis. We considered only European regions in the analysis, excluding French overseas departments. Among these, framework contracts are qualitatively different from other public contracts in several aspects (average contract value, contract length) and therefore have a different level of competition and corruption risk than other contracts. Accordingly, in the second step, framework contracts (1,131,730 contracts) were excluded and were not considered in the analysis. The distribution of the 5,057,802 contracts by year and country on which the analysis is based is shown in Table A3.1a-b.

Table A3.1a here
Table A3.1b here

Histograms of the main variables analyzed are shown in Figures A3.1a-d and A3.2a-g.

Fig. A3.1a-d here
Fig. A3.2a-g here

A4. Descriptive statistics: NUTS2-Level Data

Table A4.1 here
Table A4.2 here

There are relatively significant differences in corruption risk indicators and level of education amongst European regions. The lowest corruption risk was found in the Austrian, German, Spanish, and Italian regions for the period 2006-2020 (see Table A4.3.), while the highest values were found in the Polish region of Opolskie (PL52) in 2014, and the Spanish region of Cantabria (ES13) in 2006. The lowest CR value is 0.0, the highest is 0.79, the CCR ranges from 0.0 to 0.9, and the tertiary education rate ranges from 6.8% to 58.6%. According to Eurostat data, the lowest tertiary education level was in the Czech region of Severozápad (CZ04) in 2008 and the highest in the Lithuanian region of Sostinės region (LT01) in 2020.

Table A4.3 here

The value of the variables analyzed shows relatively high stability between 2009 and 2020, except for the share of EU-funded contracts, which decreased significantly in 2016-2018 compared to 2009. In addition, there is an increasing trend in ln of corruption risk, which reached 120% of the 2009 value by 2020 (see Fig. A4.1.).

Fig. A4.1 here

The pairwise correlations (see Fig A4.2., Table A4.4. and A4.5.) show there is an apparent strong negative relationship (r=-0.82) between the two indicators of the level of corruption risk (lnCR and lnCoCR). A moderately strong relationship (r=0.48) is observed between the share of tertiary education and GDP per capita. Furthermore, regions with a higher share of tertiary education in the population have a more extensive public procurement market, with more public procurement per year (r=0.37). A weak negative correlation (r=-0.10) is observed between the share of tertiary education and the level of corruption risk. A weak positive correlation (r=0.09) is observed between the control of corruption risk.

Fig. A4.2 here

Table A4.4 here

Table A4.5 here

A5. Nuts2 Level Analysis for the Period 2006-2020

In this section we present the estimations conducted on the level of the public procurement contracts in order to check the robustness of the results based on the data aggregated to the level of NUTS2 regions. Two binary variables are used as outcome variables: CR and CoCR which can take values 0 and 1 as described earlier. The percentages of contracts characterized where the Corruption Risk (CR) =1 and Control of Corruption Risk (CoCR) =1 and the number of contracts considered are indicated in Table A5.1.

Table A5.1 here

We repeated the regression analyses on the NUTS2 level dataset without considering the impacts of periods. We calculated for every variable the average value for the period 2006-2020 and run the following estimations:

$$lnCR_i = \alpha + \beta lnEDU_i + \sum_{j=1}^{k} \gamma_j X_{ij} + \varepsilon_i \quad \text{(A5.1)}$$

$$lnCoCR_i = \alpha + \beta lnEDU_i + \sum_{j=1}^{k} \gamma_j X_{ij} + \varepsilon_i \quad \text{(A5.2)}$$

The results of these models indicate that higher educational attainment may result in lower corruption risk and higher control of corruption risk, even if some of the key features of the contracts are taken into account as control variables (see Table A5.2).

Table A5.2 here

Tables and figures

Fig. 4.1.1a-f.: Maps demonstrating the regional distribution of the CR, CoCR and education variables

Fig. 4.1.1a

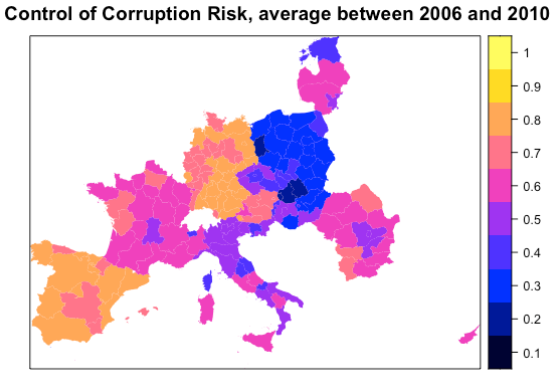


Fig. 4.1.1b

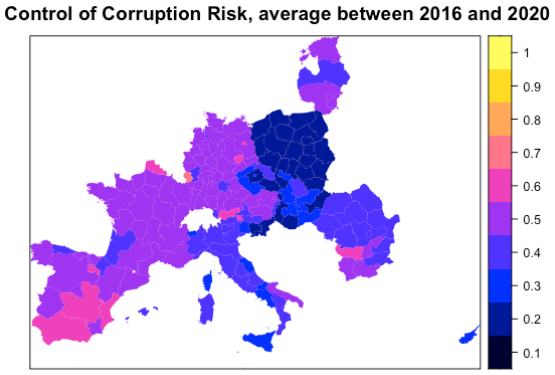


Fig. 4.1.1c

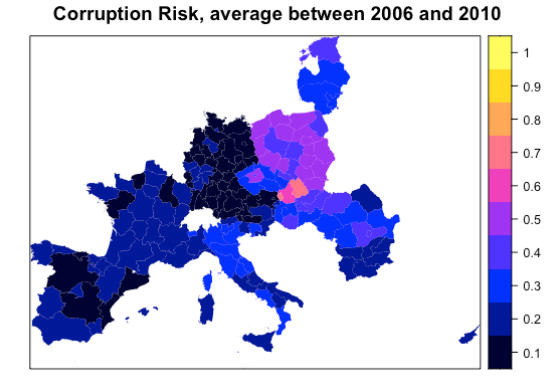


Fig. 4.1.1d

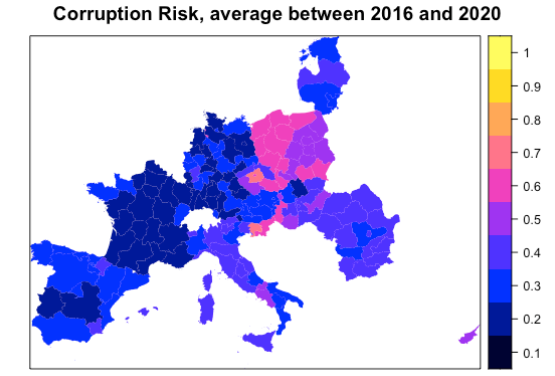


Fig. 4.1.1e

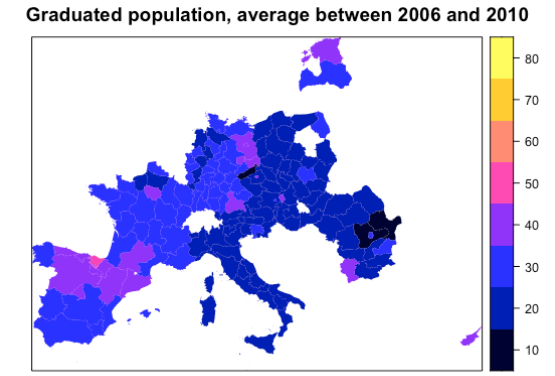


Fig. 4.1.1f

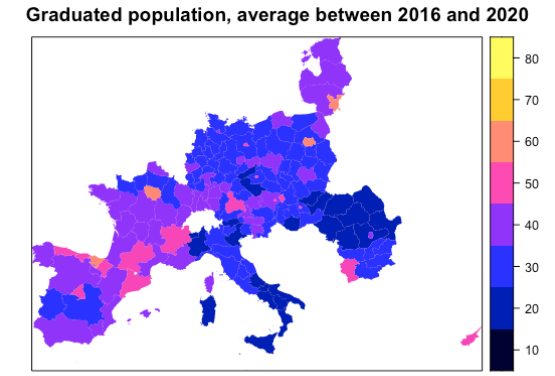


Table 4.2.1.: Effect of Educational Attainment (*lnEDU*) on Corruption Risk (*lnCR*) in the Analyzed European Regions, 2006-2020

	Dependent variable: *lnCR*							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lnEDU	-0.622***	-0.230***	-0.364***	-0.412***	-0.484***	-0.369***	-0.502***	-0.377***
	(0.038)	(0.042)	(0.040)	(0.039)	(0.042)	(0.042)	(0.040)	(0.155)
lnGDP	-	-0.540***	-0.154***	0.082***	0.027***	0.0132	0.039	0.107*
		(0.034)	(0.000)	(0.042)	(0.043)	(0.041)	(0.039)	(0.062)
lnNCV	-	-	-0.190***	-0.213***	-0.168***	-0.143***	-0.128***	-0.116***
			(0.012)	(0.012)	(0.015)	(0.015)	(0.014)	(0.016)
lnEU	-	-	-	0.108***	0.110***	0.092***	0.050***	0.069***
				(0.009)	(0.009)	(0.009)	(0.009)	(0.011)
lnCASEN	-	-	-	-	0.058***	0.057***	0.030***	0.016
					(0.012)	(0.012)	(0.011)	(0.018)
lnSH_STATE						-0.549***	-0.230***	-0.258***
						(0.050)	(0.050)	(0.088)
lnY20_39							1.477***	1.091***
							(0.101)	(0.149)
Year Dummies	Y	Y	Y	Y	Y	Y	Y	Y
Constant	-0.482***	4.307***	3.064***	1.490***	1.351***	-0.023	-4.410***	-3.430***
	(0.122)	(0.309)	(0.296)	(0.305)	(0.310)	(0.319)	(0.424)	(0.543)
F	55.06***	55.41***	77.27***	81.19***	79.68***	92.4***7	103.14***	
1st stage F								120.704***
Tests of Endogeneity								
Dubrin (score) chi2(1)								0.2434
Wu-Hausman F (1,1399)								0.2400
N	2433	1975	1975	1899	1899	1899	1868	1419

*Note: *: p>0.1 **: p>0.05 ***: p>0.01; robust errors are in the brackets*

Table 4.2.2.: Effect of Educational Attainment (*lnEDU*) on Control of Corruption Risk (*lnCoCR*) in the Analyzed European Regions, 2006-2020

	Dependent variable: *lnCoCR*							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lnEDU	0.374***	0.211***	0.314***	0.330***	0.369***	0.358***	0.472***	0.674***
	(0.023)	(0.032)	(0.030)	(0.030)	(0.032)	(0.033)	(0.032)	(0.133)
lnGDP	-	-0.327***	-0.041	-0.178***	-0.151***	-0.146***	-0.154***	-0.320***
		(0.026)	(0.029)	(0.032)	(0.033)	(0.033)	(0.030)	(0.053)
lnNCV	-	-	0.232***	0.238***	0.216***	0.214***	0.191***	0.179***
			(0.009)	(0.009)	(0.012)	0.012	(0.011)	(0.014)
lnEU	-	-	-	-0.071***	-0.072***	-0.068***	-0.035***	-0.054***
				(0.007)	(0.007)	(0.007)	(0.007)	0.009
lnCASEN	-	-	-	-	-0.030***	-0.031***	-0.014	-0.009
					(0.009)	0.009	(0.009)	(0.016)
lnSH_STATE						0.084**	-0.218***	-0.434***
						(0.040)	(0.040)	(0.076)
lnY20_39							-1.288***	-1.518***
							(0.080)	(0.128)
Year Dummies	Y	Y	Y	Y	Y	Y	Y	Y
Constant	-1.552***	-4.577***	3.920***	-2.909***	-2.842***	-2.694***	1.106***	1.880***
	(0.076)	(0.235)	(0.220)	(0.234)	(0.237)	(0.255)	(0.334)	(0.466)
F	56.43***	47.64***	116.68***	108.299***	104.17***	102.87***	113.92***	
1st stage F								120.704***
Tests of Endogeneity								
Dubrin (score) chi2(1)								2.773
Wu-Hausman F (1,1399)								2.739
N	2440	1977	1977	1899	1899	1899	1868	1419

*Note: *: p>0.1 **: p>0.05 ***: p>0.01; robust errors are in the brackets*

Fig. 4.2.1a-b: The Impact of Educational Attainment (*lnEDU*) on Corruption Risk (ln*CR*) and Control of Corruption Risk (*lnCoCR*) in the Analyzed European Regions 2009-2020 by Years Based on Model Specification 5th in Table 4.2.2.

Fig. 4.2.1a.: Corruption Risk (ln*CR*)

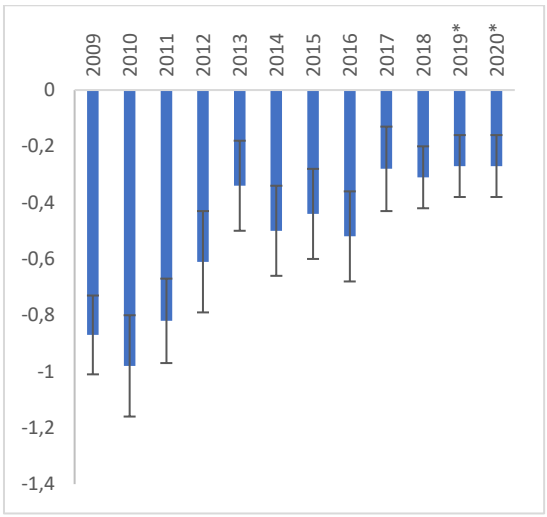
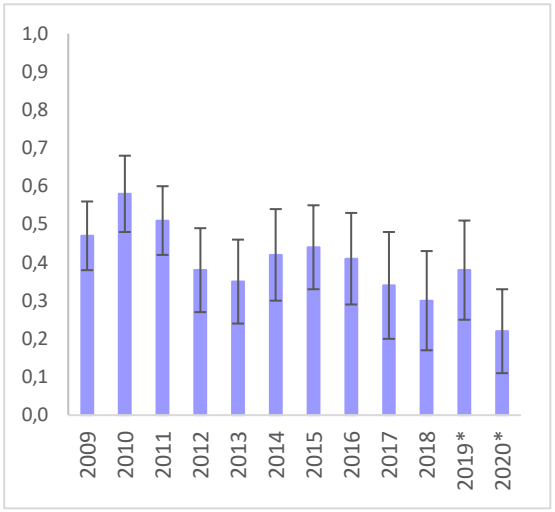


Fig. 4.2.1b.: Control of Corruption Risk (ln*CoCR*)



*Note: *: 2019 and 2020 without the Romanian regions*

Table 4.2.3.: Effect of Change in Educational Attainment (d*EDU*) on Change of Corruption Risk (*dCR*) in the Analyzed European Regions, 2009-2020

	Dependent variable: dCR							
	Model 1		Model 2		Model 3		Model 4	
	y-on-y (1)	3y-on-3y (2)	y-on-y (3)	3y-on-3y (4)	y-on-y (5)	3y-on-3y (6)	y-on-y (7)	3y-on-3y (8)
dEDU	-0.002**	-0.002*	-0.002**	-0.002*	-0.002**	-0.002*	-0.002**	-0.003*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
dCASEN	-	-	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
dlnNCV	-	-	-	-	-0.001	-0.005	-0.003	-0.006
					(0.002)	(0.004)	(0.002)	(0.006)
dEU	-	-	-	-	-	-	0.116***	0.030
							(0.019)	(0.031)
Period Dummies	Y	Y	Y	Y	Y	Y	Y	Y
dGDP	Y	Y	Y	Y	Y	Y	Y	Y
Constant	0.004	0.024***	0.003	0.023***	0.003	0.022***	0.001	0.022***
	(0.004)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.004)	(0.006)
F	4.08	8.25	4.58	8.29	4.26	7.17	6.19	6.15
N	1826	501	1826	501	1826	501	1826	501

*Note: *: p>0.1 **: p>0.05 ***: p>0.01; robust errors are in the brackets*

Table 4.2.4.: Effect of Change in Educational Attainment (d*EDU*) on Change of Control of Corruption Risk (*dCoCR*) in the Analyzed European Regions, 2009-2020

	Dependent variable: dCoCR							
	Model 1		Model 2		Model 3		Model 4	
	y-on-y (1)	3y-on-3y (2)	y-on-y (3)	3y-on-3y (4)	y-on-y (5)	3y-on-3y (6)	y-on-y (7)	3y-on-3y (8)
dEDU	0.003**	0.006***	0.003**	0.006***	0.003**	0.006***	0.003**	0.006***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
dCASEN	-	-	0.000	0.000**	0.000	0.000***	0.000	0.000***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
dlnNCV	-	-	-	-	0.022***	0.035***	0.023***	0.035***
					(0.003)	(0.006)	(0.003)	(0.006)
dEU	-	-	-	-	-	-	-0.063**	0.009
							(0.025)	(0.040)
Period Dummies	Y	Y	Y	Y	Y	Y	Y	Y
dGDP	Y	Y	Y	Y	Y	Y	Y	Y
Constant	-0.002	-0.054***	-0.002	-0.051***	0.000	-0.047***	0.000	-0.047***
	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.000)	(0.007)
F	7.36	5.38	6.81	4.92	9.66	10.44	9.51	8.90
N	1826	501	1826	501	1826	501	1826	501

*Note: *: p>0.1 **: p>0.05 ***: p>0.01*

Table A1.1.1.: List and Definition of Variables

#	Variable name	Definition
	GDP	GDP per capita, PPP (current international USD)
	CoC	Control of Corruption: Estimate. „Control of Corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5." [https://datacatalog.worldbank.org/search/dataset/0038026/Worldwide-Governance-Indicators]
	CPI	Corruption Perceptions Index „The 2022 Corruption Perceptions Index (CPI) shows that most countries are failing to stop corruption. The CPI ranks 180 countries and territories around the world by their perceived levels of public sector corruption, scoring on a scale of 0 (highly corrupt) to 100 (very clean)." [https://www.transparency.org/en/cpi/2022]
	CPI_INV	CPI_INV = 10 – (CPI/10)
	EDU	Tertiary education (levels 5-8), Population by educational attainment level, sex and NUTS 2 regions (%)
	CR	Corruption Risk; $CR_{it} = \sum_{j=1}^{n} \frac{cr_{itj}}{n_{it}}$ where the *cr* is a dummy variable with the value 0 if the contract was awarded with more than one bid; *cr* has the value 1 if there was only one bid. The $\sum_{j=1}^{n} cr_{itj}$ is the number of contract awarded with one bidder and n_{it} is the number of contracts in *i* country and *t* year.
	CoCR	Control of Corruption Risk $CoCR_{it} = \sum_{j=1}^{n_{it}} \frac{cocr_{itj}}{n_{it}}$ where the *cocr* is a dummy variable with the value 0 if the contract was awarded with less than four bidders; *cocr* has the value 1 if there were at least four bids. The $\sum_{j=1}^{n_{it}} cocr_{itj}$ is the number of contract awarded with more than three bidders and n_{it} is the number of contracts in *i* country and *t* year.
	lnNCV	Mean Value of Logarithm of Net Contract Value $lnNCV_{it} = \sum_{j=1}^{n} \frac{lnncv_{itj}}{n_{it}}$ where $lnncv_{jtj}$ is the logarithm of net contract value of *j* contract in *i* country and *t* year; and n_{it} the number of contracts in *i* country and *t* year.
	lnEU	Logarithm of Share of Contract Funded by EU in All Contract $EU_{it} = \sum_{j=1}^{n_{it}} \frac{eu_{itj}}{n_{it}}$ where the *eu* is a dummy variable with the value 0 if the contract was funded by domestic sources; *eu* has the value 1 if the contract was funded by EU subsidies. The $\sum_{j=1}^{n_{it}} eu_{itj}$ is the number of contract funded by EU, and n_{it} is the number of contracts in *i* country and *t* year. $lnEU_{it}$ = the logarithm of EU_{it} *in i* country and *t* year.

Table A1.1.2.: Data Sources, Links and Data Files of Indicators Analyzed, 2000-2022

Variable name	Data source	Link	Files
GDP	The World Bank	https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD	API_NY.GDP.PCAP.PP.CD_DS2_en_csv_v2_4770425.csv
GDP [NUTS2 level]	Eurostat	https://ec.europa.eu/eurostat/databrowser/view/TGS00005/default/table	tgs00005_linear.csv
CoC	The World Bank	https://datacatalog.worldbank.org/search/dataset/0038026/Worldwide-Governance-Indicators	wgidataset-fixed.dta
CPI	Transparency International	https://www.transparency.org/en/cpi/	CPI-2000_200603_083012.csv CPI-2001_200603_082938.csv CPI-2002_200602_115328.csv CPI-2003_200602_113929.csv CPI-2004_200602_110140.csv CPI-2005_200602_104136.csv CPI-2006-new_200602_095933.csv CPI-2007-new_200602_092501.csv CPI-Archive-2008-2.csv CPI-2009-new_200601_120052.csv CPI-2010-new_200601_105629.csv CPI-2011-new_200601_104308.csv CPI2012_Results.xls CPI2013_DataBundle_2022-01-20-182851_xyum.zip CPI2014_DataBundle-2.zip CPI_2015_FullDataSet_2022-01-18-145020_enyn_2022-01-20-180010_mabu.xlsx CPI2016_Results.xlsx CPI2017_Full_DataSet-1801.xlsx CPI2018_Full-Results_1801.xlsx CPI2019-1.xlsx CPI-2021-Full-Data-Set.zip CPI2022_GlobalResultsTrends.xlsx
EDU	Eurostat	https://ec.europa.eu/eurostat/databrowser/view/edat_lfse_04/default/table?lang=EN	edat_lfse_04_linear.csv.gz
CR, CoCR, lnNCV, lnEU	TED - Tenders Electronic Daily	https://data.europa.eu/data/datasets/ted-csv?locale=en	https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2006.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2007.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2008.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2009.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2010.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2011.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2012.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2013.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2014.zip

			https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2015.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2016.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2017.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2018.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2019.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2020.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2021.zip https://data.europa.eu/api/hub/store/data/ted-contract-award-notices-2022.zip

Table A1.2.1.: Descriptive Statistics of the Corruption Indicators by Years, 2000-2022

	Corruption Risk (CR)			Corruption Perceptions Index (inv.) (CPI_INV)			Control of Corruption Risk (CoCR)			Control of Corruption (CoC)		
	Mean	Median	Standard deviation	Median	Standard deviation	Standard deviation	Mean	Median	Standard deviation	Mean	Median	Standard deviation
2000				3.58	3.60	2.25				1.11	1.11	0.89
2001				3.59	3.40	2.15						
2002				3.57	3.70	2.20				1.07	1.16	0.91
2003				3.63	3.40	2.30				1.11	1.16	0.85
2004				3.58	3.45	2.24				1.09	1.07	0.84
2005				3.51	3.45	2.22				1.10	1.03	0.82
2006	0.13	0.07	0.12	3.46	3.35	2.14	0.63	0.66	0.16	1.12	1.06	0.85
2007	0.17	0.12	0.13	3.43	3.45	1.94	0.56	0.57	0.16	1.11	1.05	0.88
2008	0.17	0.12	0.14	3.48	3.45	1.83	0.52	0.55	0.19	1.09	1.12	0.86
2009	0.18	0.13	0.14	3.58	3.50	1.91	0.55	0.58	0.17	1.06	1.05	0.86
2010	0.18	0.11	0.14	3.62	3.65	1.95	0.55	0.58	0.17	1.06	1.03	0.84
2011	0.18	0.15	0.12	3.65	3.75	2.04	0.51	0.51	0.17	1.06	1.06	0.85
2012	0.20	0.17	0.12	3.49	3.50	1.62	0.51	0.50	0.16	1.07	1.13	0.88
2013	0.21	0.18	0.13	3.55	3.75	1.60	0.48	0.49	0.14	1.07	1.08	0.87
2014	0.21	0.17	0.12	3.46	3.70	1.57	0.48	0.51	0.16	1.05	0.99	0.84
2015	0.22	0.21	0.12	3.34	3.65	1.59	0.45	0.49	0.14	1.07	0.96	0.85
2016	0.23	0.17	0.12	3.45	3.80	1.59	0.44	0.48	0.14	1.06	0.86	0.84
2017	0.26	0.22	0.13	3.46	3.80	1.52	0.42	0.45	0.14	1.03	0.83	0.82
2018	0.26	0.24	0.13	3.46	3.80	1.51	0.40	0.43	0.14	1.04	0.85	0.84
2019	0.27	0.25	0.14	3.49	3.80	1.53	0.40	0.39	0.16	1.02	0.85	0.83
2020	0.29	0.31	0.13	3.52	3.85	1.52	0.38	0.37	0.14	1.05	0.80	0.82
2021				3.50	3.85	1.51				1.05	0.81	0.82
2022				3.53	3.80	1.45						

Table A1.2.2.: Descriptive Statistics of the Main Variables

	mean	median	Std. deviation	N
CR	0.21	0.18	0.13	482
CPI_INV	3.52	3.60	1.83	725
CoCR	0.48	0.49	0.17	482
CoC	1.07	1.04	0.84	672
GDP	33186.03	30863.02	18589.04	770
EDU	27.93	28.2	9.69	693

Fig. A1.2.1a-f.: Histograms of the Corruption Indicators (Perception and Objective Indicators)

Fig. A1.2.1a

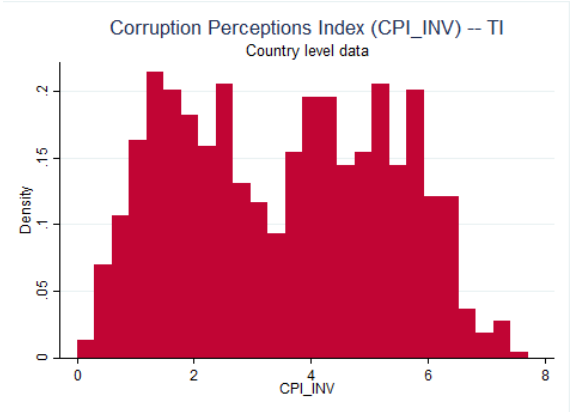


Fig. A1.2.1b

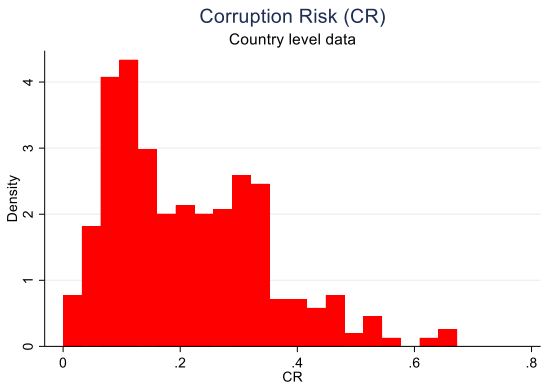


Fig. A1.2.1c

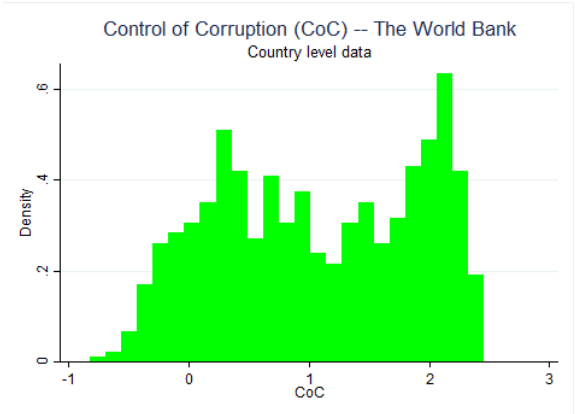


Fig. A1.2.1d

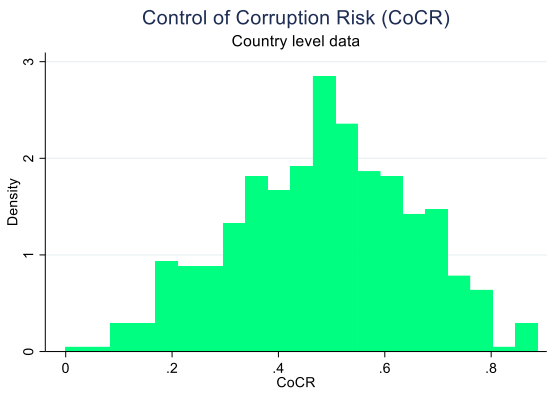


Fig. A1.2.2a-f.: Histogram of the GDP per in PPS, the Share of Tertiary Education in the Population Aged 25-64, the Logarithm of Mean Contract Value and the Logarithm of Share of EU Funded Contracts in All Contracts

Fig. A1.2.2a

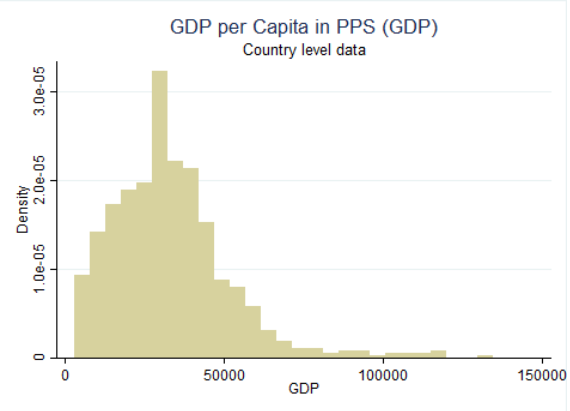


Fig. A1.2.2b

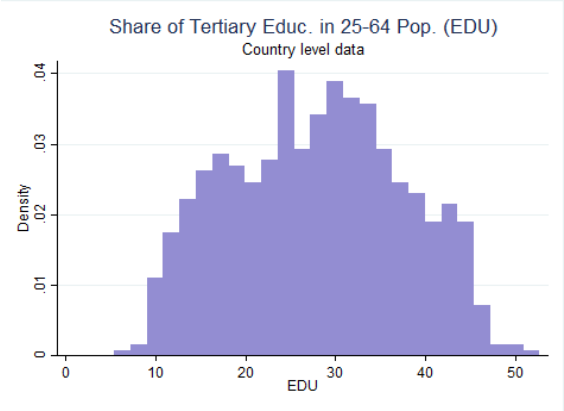


Fig. A1.2.2c

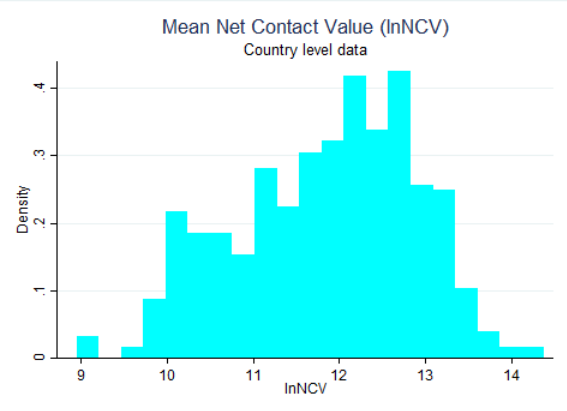


Fig. A1.2.2d

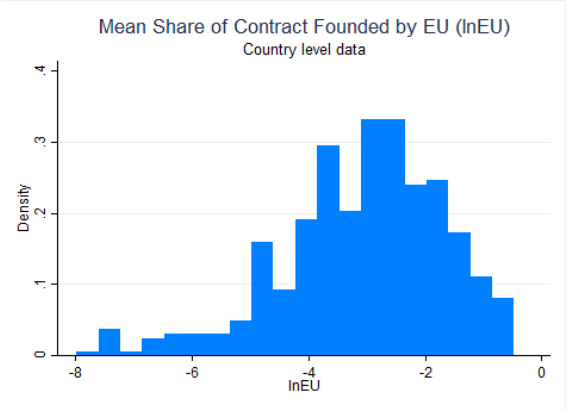


Table A1.2.3.: Correlation Matrix of the Main Variables

	CR	CoCR	CPI_INV	CoC	EDU
CR	1.000				
CoCR	-0.868***	1.000			
CPI_INV	0.621***	-0.553***	1.000		
CoC	-0.639***	0.567***	-0.974***	1.000	
EDU1	-0.290***	0.233***	-0.588***	0.577***	1.000
GDP2	-0.381***	0.226***	-0.6317***	0.651***	0.659***

*: p<0.1 **: p<0.05 ***: p<0.01

Fig. A1.2.3: Scatterplots of the Main Variables

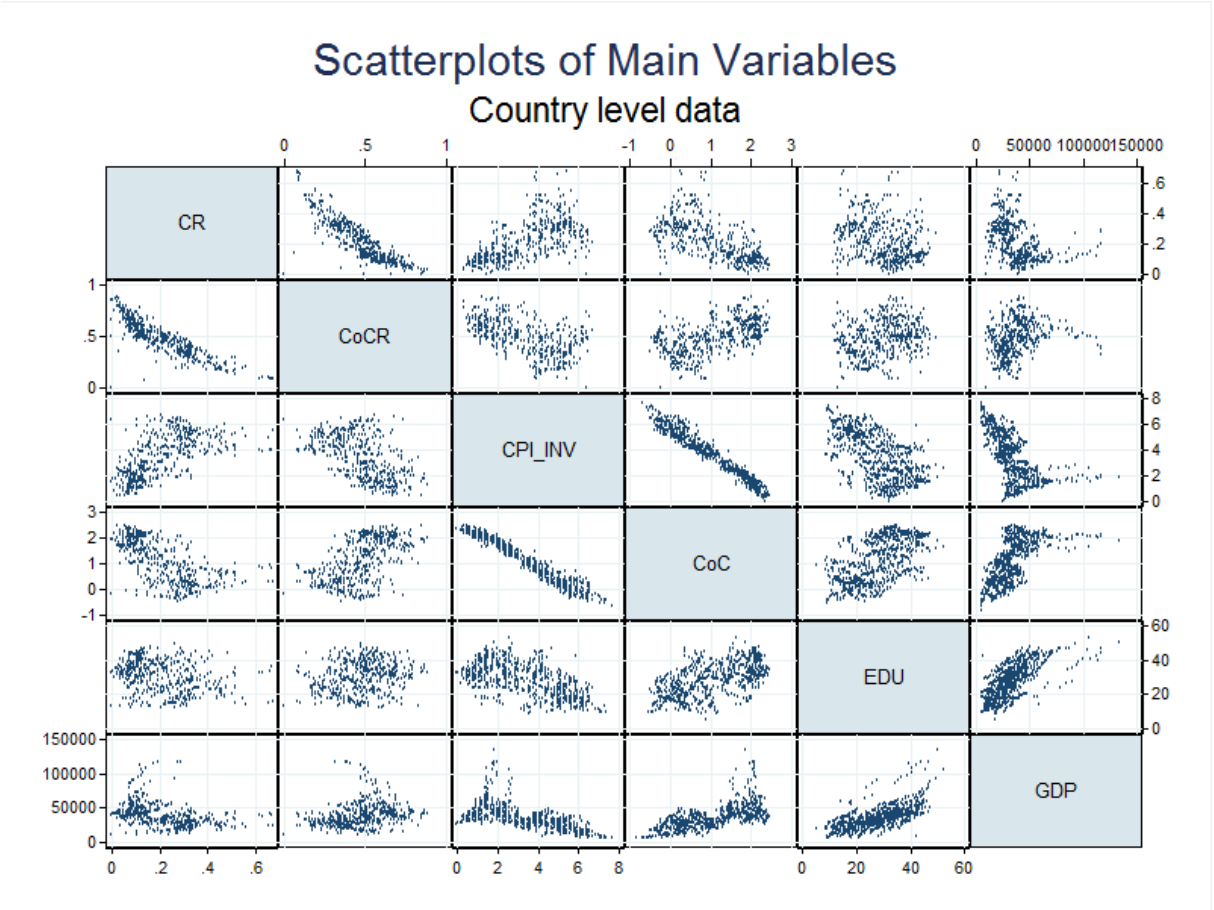


Table A1.2.4.: Correlation Matrix of lnEU, lnNCV and lnGDP – Country Level Data

	lnEU	lnNCV
lnEU	1.000	
lnNCV	-0.412***	1.000
lnGDP	-0.414***	0.592***

Table A1.3: Beta-coefficients of the Educational Attainment and the GDP According to the OLS-models with Robust Standard Errors on the Country Level Dataset.

	lnCR (1)	lnCR (2)	lnCPI_INV (3)	lnCoCR (4)	lnCoCR (5)	lnCoC (6)
lnEDU	-0.486***	-0.469***	-0.514***	0.251***	0.268***	0.812***
	(0.073)	(0.091)	(0.048)	(0.051)	(0.054)	(0.071)
lnNCV	-0.262***	-	-	0.125***	-	-
	(0.028)			(0.020)		
lnEU	0.081***	-	-	-0.035***	-	-
	(0,015)			(0.011)		
lnGDP	-0.072	-0.790***	-0.810***	-0.063	0.236***	1.225***
	(0.077)	(0.068)	(0.036)	(0.054)	(0.041)	(0.060)
Year Dummies	Y	Y	Y	Y	Y	Y
Constant	3.343***	7.296***	7.249***	-2.212***	-3.752***	-14.538***
	(0.059)	(0.130)	(0.277)	(0.388)	(0.356)	(0.548)
F value	41.58	29.48	55.03	18.66	15.41	43.22
N	432	466	684	432	467	594

**: p<0.1 **: p<0.05 ***: p<0.01*

Fig. A2.1a-d: Histogram of Variables Analyzed – The WVS dataset

Fig. A2.1a

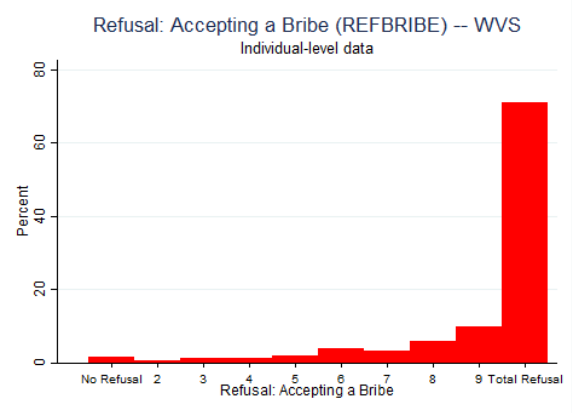


Fig. A2.1b

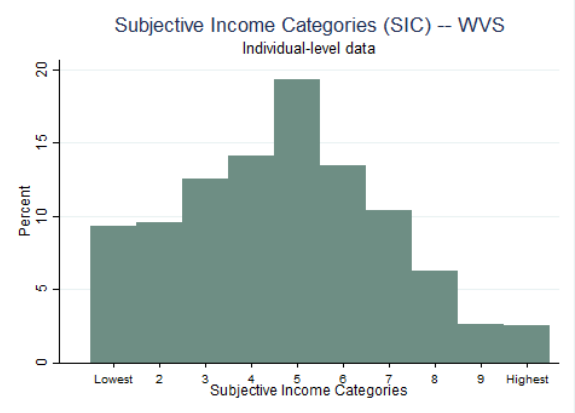


Fig. A2.1c

Fig. A2.1d

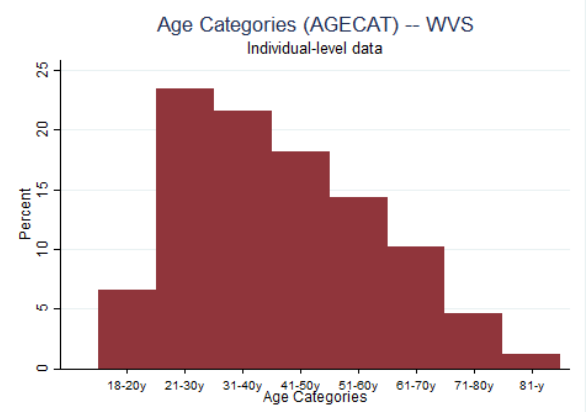


Table A2.1.: Rejection of Bribery by Education, Sex and Age Categories – The WVS Dataset.

		Incomplete Refusal (*REFBRIBED*=0)	Complete Refusal (*REFBRIBED*=1)	Total	N
Education	Less than tertiary	29.6	70.4	100.0	246,721
	Tertiary or highest	27.6	72.4	100.0	78,110
Sex	Female	27.9	72.1	100.0	174,195
	Male	30.5	69.5	100.0	159,619
Age cat.	18-20	37.0	63.0	100.0	21,890
	21-30	34.5	65.5	100.0	77,586
	31-40	30.0	70.0	100.0	71.354
	41-50	28.1	71.9	100.0	60.082
	51-60	25.3	74.7	100.0	47,653
	61-70	22.1	77.9	100.0	33,709
	71-80	19.1	80.9	100.0	15,196
	81-	17.9	82.1	100.0	3,788

Table A2.2a: Impact of Education on Rejection of Bribery *(REFBRIBE*) – All Surveyed Countries in 3-7 Wave of WVS Dataset.

Variable Name:	REFBRIBE				
	(1)	(2)	(3)	(4)	(5)
EDU	1.087***	1.090***	1.156***	1.194***	1.207***
	(0.010)	(0.010)	(0.011)	(0.012)	(0.013)
SEX	-	1.132***	1.118***	1.127***	1.123***
		(0.009)	(0.009)	(0.009)	(0.010)
SIC	-	-	0.960***	0.968***	0.968***
			(0.002)	(0.002)	(0.002)
AGECAT	-	-	-	1.155***	1.159***
				(0.003)	(0.003)
SSIZE	-	-	-	-	1.006***
					(0.002)
YEAR Dummies	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.011	0.011	0.012	0.016	0.017
N	324,831	324,640	304,035	301,907	241,572

**: p<0.1 **: p<0.05 ***: p<0.01*

Note: ordered logic estimations, odds ratios are in cells and standard errors are in brackets

Table A2.2b: Impact of Education on Rejection of Bribery (*REFBRIBE*) – Only 16 European Countries in WVS Dataset.

Variable Name	REFBRIBE				
	(1)	(2)	(3)	(4)	(5)
EDU	1.070*	1.063***	1.130***	1.166***	1.193***
	(0.043)	(0.043)	(0.048)	(0.050)	(0.053)
SEX	-	1.277***	1.240***	1.244***	1.261***
		(0.043)	(0.043)	(0.043)	(0.045)
SIC	-	-	0.954***	0.982**	0.982***
			(0.009)	(0.009)	(0.009)
AGECAT	-	-	-	1.183***	1.186***
				(0.012)	(0.013)
SSIZE	-	-	-	-	0.981***
					(0.007)
YEAR Dummies	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.050	0.052	0.054	0.061	0.062
N	17,934	17,932	16,837	16,826	15,759

**: p<0.1 **: p<0.05 ***: p<0.01*

Note: ordered logic estimations, odds ratios are in cells and standard errors are in brackets

Table A2.2c: Impact of Education on Rejection of Bribery (*REFBRIBED*) – WVS Dataset.

Variable Name	REFBRIBED			
	Probit		Logit	
	All countries	Only surveyed countries	All countries	Only surveyed countries
	(1)	(2)	(3)	(4)
EDU	0.107***	0.082***	0.177***	0.141***
	(0.007)	(0.027)	(0.011)	(0.046)
SEX	0.067***	0.145***	0.111***	0.246***
	(0.005)	(0.022)	(0.009)	(0.037)
SIC	-0.019***	-0.013**	-0.033***	-0.022**
	(0.001)	(0.006)	(0.002)	(0.010)
AGECAT	0.084***	0.096***	0.144***	0.164***
	(0.002)	(0.007)	(0.003)	(0.011)
SSIZE	0.001***	-0.012***	0.002	-0.020***
	(0.001)	(0.004)	(0.002)	(0.007)
YEAR Dummies	Yes	Yes	Yes	Yes
Pseudo R^2	0.028	0.103	0.028	0.103
N	241,572	15,769	241,572	15,769

**: p<0.1 **: p<0.05 ***: p<0.01*

Note: probit and logit estimations, robust standard errors are in brackets

Table A3.1a: Number of Analyzed Contracts by Year and Countries, 2006-2020

ccode	2006	2007	2008	2009	2010	2011	2012	2013
AT	1,702	2,387	2,961	2,828	2,712	2,828	2,841	2,572
BG	0	1,761	3,818	3,934	6,153	6,241	8,340	10,456
CY	623	661	825	868	1,093	1,049	902	994
CZ	2,704	3,069	3,984	5,076	5,069	4,986	6,631	6,663
DE	12,476	15,883	16,070	19,000	20,803	22,363	23,667	24,773
EE	641	707	821	641	1,320	1,128	1,626	1,248
ES	13,295	16,122	16,745	18,510	22,423	18,499	14,757	15,501
FR	77,945	98,918	98,812	90,576	86,871	89,593	93,309	86,410
HU	3,864	4,020	5,202	6,128	6,131	5,923	4,491	6,529
IT	10,361	14,095	15,232	17,344	17,257	18,468	18,653	17,496
LT	3,661	3,847	3,988	3,423	5,846	8,371	13,331	6,387
LV	1,881	2,719	2,853	2,711	8,289	7,758	6,131	5,567
PL	52,214	54,787	67,467	75,440	86,321	102,221	108,897	116,181
RO	0	6,609	13,203	8,036	7,569	7,267	5,563	5,011
SI	2,143	2,964	2,941	3,254	3,411	3,727	3,625	2,916
SK	790	1,090	1,598	1,251	1,301	1,438	2,176	1,743
Total	184,300	229,639	256,520	259,020	282,569	301,860	314,940	310,447

Note: without framework agreements

Table A3.1b: Number of Analyzed Contracts by Year and Countries, 2006-2020

ccode	2014	2015	2016	2017	2018	2019	2020	Total
AT	2,522	2,412	3,060	3,157	3,432	3,887	4,625	43,926
BG	11,505	10,595	11,999	18,558	18,294	18,996	18,091	148,741
CY	918	562	832	531	672	803	498	11,831
CZ	7,063	7,533	8,397	16,178	23,642	25,449	24,483	150,927
DE	25,179	27,720	33,799	41,207	45,723	56,869	57,736	443,268
EE	1,238	1,434	1,280	1,378	1,840	2,645	2,015	19,962
ES	17,909	17,308	19,605	22,539	28,044	35,259	32,756	309,272
FR	79,476	78,960	70,372	69,419	65,408	63,447	47,909	1,197,425
HU	5,280	5,526	5,786	7,583	10,411	9,473	10,670	97,017
IT	18,068	18,353	18,668	16,301	19,979	24,402	20,146	264,823
LT	14,922	7,262	8,359	9,343	13,842	11,484	12,372	126,438
LV	5,191	5,369	4,082	4,668	6,231	7,030	7,463	77,943
PL	118,949	113,364	100,096	126,815	137,636	143,644	147,435	1,551,467
RO	5,031	5,551	4,390	5,546	18,208	145,041	149,782	386,807
SI	2,246	2,505	5,614	33,820	39,045	45,659	48,467	202,337
SK	1,806	1,845	1,504	1,953	2,347	2,273	2,503	25,618
Total	317,303	306,299	297,843	378,996	434,754	596,361	586,951	5,057,802

Note: without framework agreements

Fig. A3.1a-f.: Histograms of Variables Analyzed by NUTS2-level

Fig. A3.1a

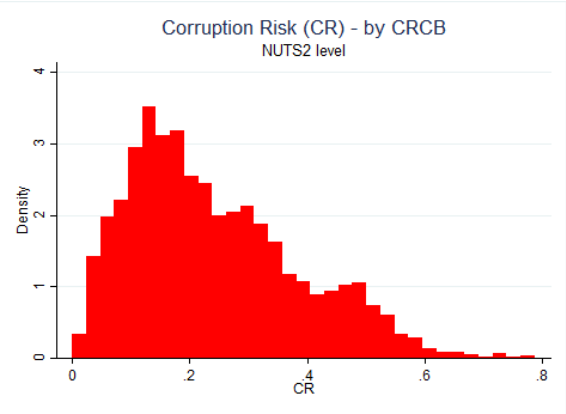


Fig. A3.1b

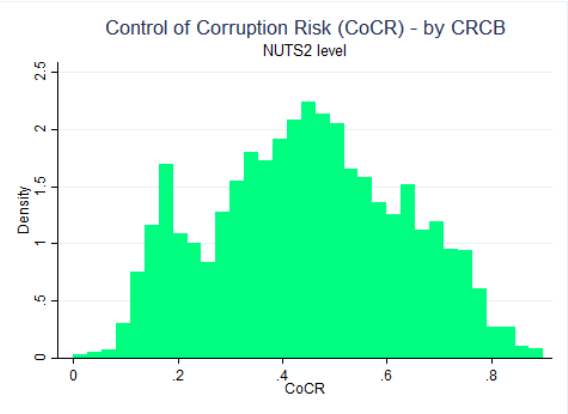


Fig. A3.1c

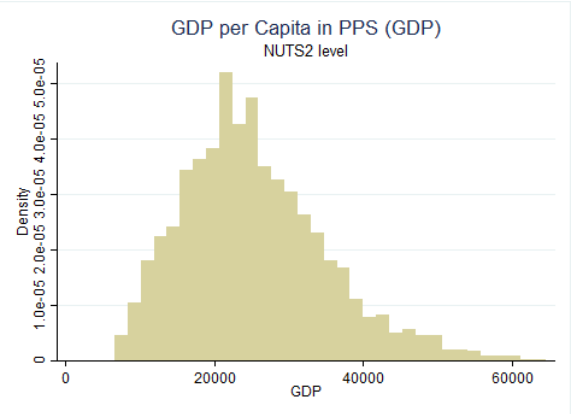


Fig. A1.1d

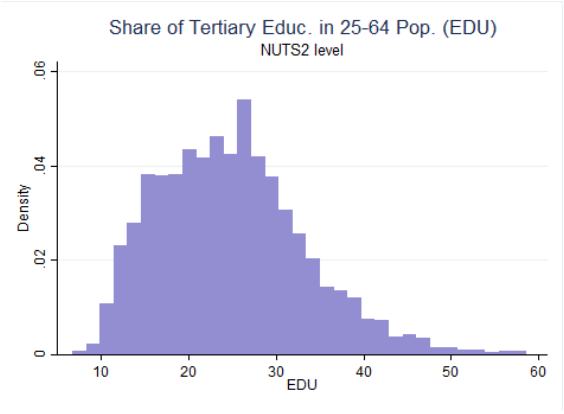


Fig. A3.2a-g: Histograms of Variables Analyzed by NUTS2-level

Fig. A3.2a

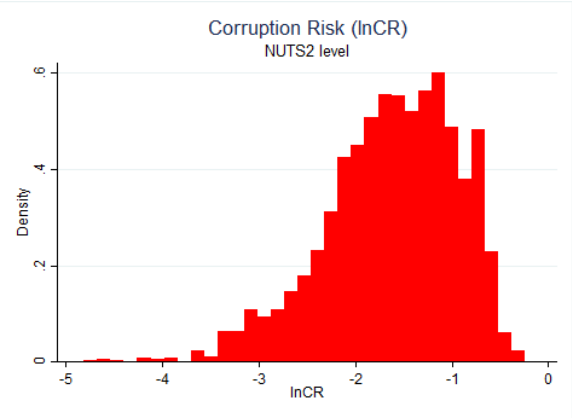


Fig. A3.2b

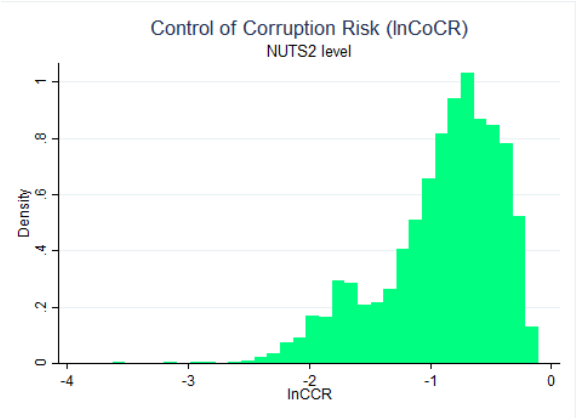


Fig. A3.2c

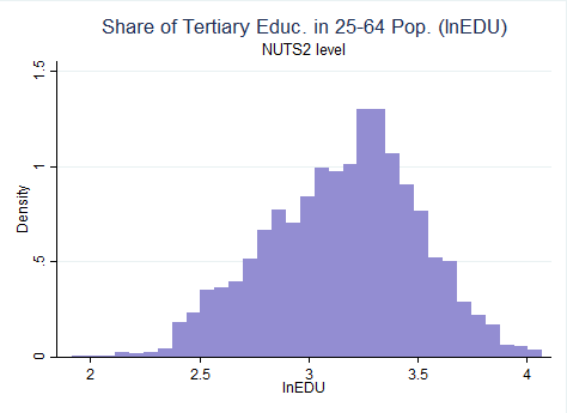


Fig. A3.2d

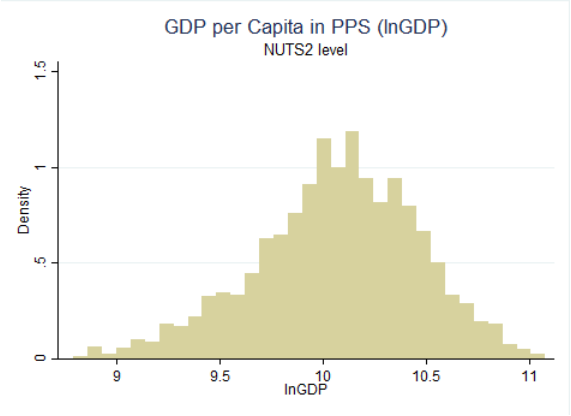


Fig. A3.2e

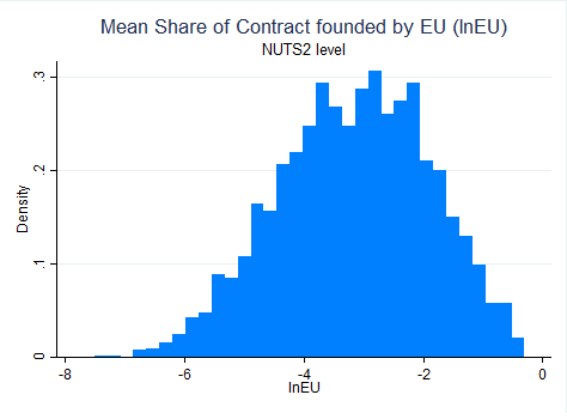


Fig. A3.2f

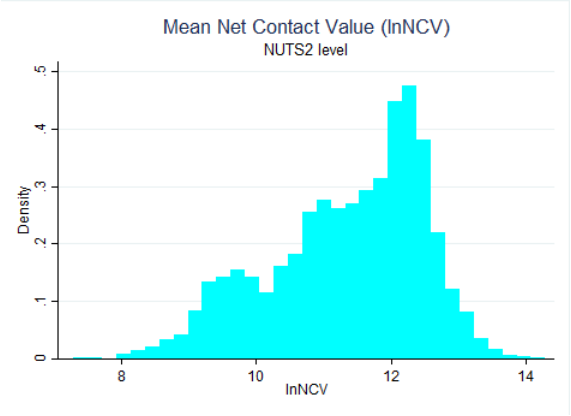


Fig. A3.2g

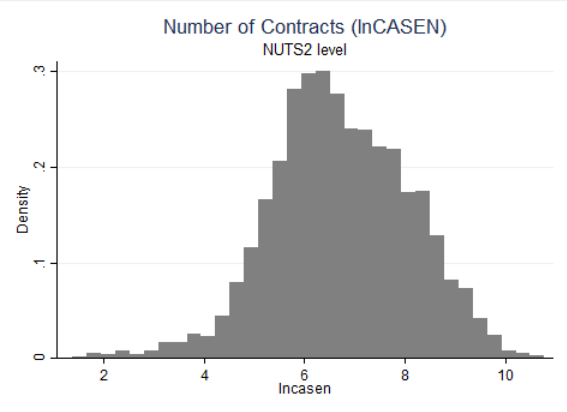


Table A4.1.: Descriptive Statistics of Variable Analyzed 2006-2020*- NUTS2 Level Dataset

Variable Name	Min.	Max.	Median	Mean	Standard Deviation	N
Corruption Risk (CR)[1]	0.000	0.786	0.212	0.241	0.143	2,491
Control of Corruption Risk (CoCR)[1]	0.000	0.900	0.449	0.447	0.183	2,491
Mean Contract Value (lnNCV)[1]	7.282	14.300	11.577	11.351	1.151	2,491
Mean Rate of EU Funded Contracts (EU)[1]	0.000	0.741	0.041	0.081	0.107	2,491
Regional GDP, PPS per Inhabitant (GDP)[2]	6,600	64,600	24,200	25,387.970	9,656.939	2,004
Share of Tertiary Educ. in Pop. 25-64, % (EDU)[3]	6.8	58.6	24.4	24.9	8.432	2,441

*Note: *: without the French overseas departments*
Sources:

1: own calculations from Tenders Electronic Daily data
2: Eurostat data (https://ec.europa.eu/eurostat/databrowser/view/TGS00005/default/table)
3: Eurostat data (https://ec.europa.eu/eurostat/web/products-datasets/-/edat_lfse_04)

Table A4.2.: Descriptive Statistics of Main Variable in 2006, 2009 and 2020 – NUTS2 Level Dataset.

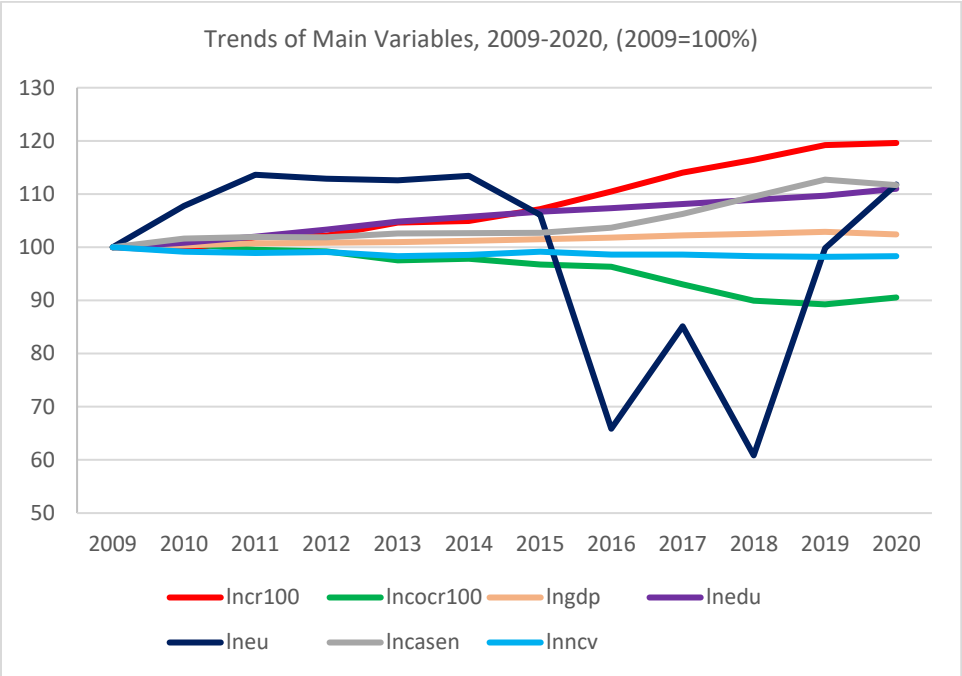
Variable Name, Year	Mean	Median	Min.	Max.	Standard Deviation	N
lnCR, 2006	-2.31	-2.32	-4.65	-0.60	0.87	149
lnCR, 2009	-1.76	-1.76	-4.14	-0.30	0.73	166
lnCR, 2020	-1.20	-1.19	-2.03	-0.33	0.39	167
lnCoCR, 2006	-0.51	-0.38	-1.79	-0.11	0.35	153
lnCoCR, 2009	-0.81	-0.64	-3.14	-0.14	0.53	167
lnCoCR, 2020	-1.17	-1.05	-2.42	-0.49	0.42	167
lnNCV, 2006	11.50	11.78	8.40	13.74	1.19	153
lnNCV, 2009	11.47	11.68	9.04	13.86	1.11	167
lnNCV, 2020	11.27	11.63	8.05	13.82	1.22	167
lnEU, 2006	-3.30	-3.31	-6.17	-1.20	1.22	137
lnEU, 2009	-3.10	-2.95	-6.14	-0.57	1.20	160
lnEU, 2020	-2.92	-2.90	-5.43	-0.62	1.04	160
lnGDP, 2009	9.92	9.97	8.79	10.82	0.41	167
lnGDP, 2020	10.16	10.15	9.28	11.01	0.35	167
lnEDU, 2006	2.95	2.95	2.08	3.73	0.34	145
lnEDU, 2009	3.03	3.05	2.13	3.77	0.34	159
lnEDU, 2020	3.36	3.42	2.47	4.07	0.31	167

Table A4.3.: Regions with Lowest or Highest Values of the Variables Analyzed*.

Variable Name	Min. Value (NUTS2 Region and Year)	Max. Value (NUTS2 Region and Year)
CR	Burgenland (AT11), 2006 Vorarlberg (AT34), 2006 Weser-Ems (DE94), 2006 Trier (DEB2), 2006, 2009 Ciudad Autónoma de Ceuta (ES63), 2011 Ciudad Autónoma de Melilla (ES64), 2007 Provincia Autonoma di Bolzano/ Bozen (ITH1), 2012	Opolskie (PL52), 2014
CoCR	Provincia Autonoma di Bolzano/ Bozen (ITH1), 2012	Cantabria (ES13), 2006
GDP	Северозападен [Severozapaden] (BG31), 2009	Praha (CZ01), 2019
EDU	Severozápad (CZ04), 2008	Sostinės regionas (LT01), 2020

*Note: *: without the French overseas departments*

Fig. A4.1.: Trends of Main Variables Analyzed * (2009= 100%).



*Note: lncr100 = lnCR*100;lncocr100= lnCOCR*100*

Fig. A4.2.: Scatterplot of lnCR, lnCoCR, lnEDU and lnGDP in 2009 and 2020.

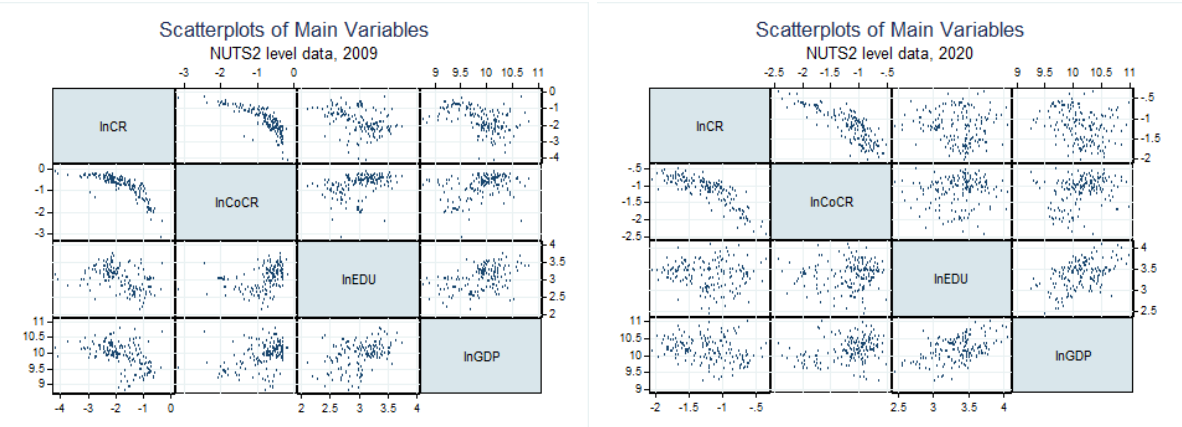


Table A4.4.: Pairwise Correlations amongst the Variables Analyzed, 2006-2020.

Variable name	lnCR	lnCoCR	lnNCV	lnEU	lnGDP	lnCASEN	lnEDU
lnCR	1.00						
lnCoCR	-0.82***	1.00					
lnNCV	-0.40***	0.45***	1.00				
lnEU	0.24***	-0.23***	-0.11***	1.00			
lnGDP	-0.29***	0.26***	0.54***	-0.40***	1.00		
lnCASEN	0.27***	-0.26***	-0.56***	-0.09***	0.01	1.00	
lnEDU	-0.10***	0.09***	0.05***	-0.09***	0.48***	0.37***	1.00

*Note: *: p < 0.1 **: p < 0.05 ***: p < 0.01*

Table A4.5.: Pairwise Correlations amongst the Main Variables Analyzed in 2009 and 2020.

	2009			2020		
	lnCR	lnCoCR	lnGDP	lnCR	lnCoCR	lnGDP
lnCR	1.00			1.00		
lnCoCR	-0.79***	1.00		-0.82***	1.00	
lnGDP	-0.47***	0.39***	1.00	-0.26***	0.27***	1.00
lnEDU	-0.39***	0.37***	0.42***	-0.10	0.08	0.48***

*Note: *: p < 0.1 **: p < 0.05 ***: p < 0.01*

Table A5.1: Percentage of Contracts Marked by the Corruption Risk (CR) and the Control of Corruption Risk (CoCR) Indicators, 2006-2020, Percent.

	CR=1	CoCR=1	N
2006	15.1	63.8	192,435
2007	23.5	47.0	244,832
2008	26.0	41.8	281,554
2009	25.6	43.8	310,018
2010	25.5	44.2	352,112
2011	26.9	41.5	387,515
2012	27.9	39.4	402,884
2013	27.2	39.6	408,635
2014	27.1	40.6	422,445
2015	27.5	39.4	423,813
2016	27.5	40.2	409,459
2017	32.1	35.1	553,445
2018	34.5	31.7	635,920
2019	34.3	30.9	869,683
2020	35.9	32.3	871,382

Table A5.2: Odds Ratios Related to the Educational Attainment and the GDP According to the Logit Models on the Contract Level Dataset

	Dependent Variable	
	Corruption Risk (CR)	Control of Corruption Risk (CoCR)
Educational Attainment	0.99***	1.01***
GDP	1.00***	0.99***
Log of Contract Value	Yes	Yes
EU-fund	Yes	Yes
Sector Dummies	Yes	Yes
Year Dummies	Yes	Yes
Country Dummies	Yes	Yes
Pseudo R-Square	0.09	0.11
N	2,939,446	2,940,255

*Note: ***: p<0.01 **: p<0.05 *: p<0.1; robust errors are in the brackets*