

# Modelling Corruption Perceptions: Evidence from Eastern Europe and Central Asian Countries

Giorgio d'Agostino<sup>1</sup> · Luca Pieroni<sup>2</sup>

Accepted: 21 March 2018 / Published online: 15 May 2018  
© Springer Science+Business Media B.V., part of Springer Nature 2018

**Abstract** This work proposes a multidimensional framework that is based on a latent class model to identify various types of corruption and to outline their importance. A dataset of Eastern European and Central Asian countries is used to identify four groups of corrupt activities, which go beyond the usual classification of corruption into administrative and political corruption. Our estimates are validated by means of a direct administrative corruption index that is derived from the same dataset and also by a comparison with the corruption perception rankings that are published by Transparency International. The potential of the proposed approach is illustrated with an application to the relationship between firms' competitiveness and the latent classes of corruption that we have identified.

**Keywords** Corruption · Eastern Europe and Central Asian economies · Latent class models · Multidimensional item response theory · Firm competitiveness

**JEL Classification** C51 · C52 · D22 · D73

## 1 Introduction

Progress in fighting corruption on all fronts requires a measurement of corruption itself to diagnose problems and monitor results. This recognition had renewed interest in many international organisations and also among aid donors, aid recipients, investors, and civil society. Many recent estimations have emphasised the considerable detrimental effects of corruption on growth rate.<sup>1</sup>

---

<sup>1</sup> Several studies have shown that a one standard-deviation increase in corruption lowers investment rates by three percentage points and also lowers average annual growth by about one percentage point (Mauro

✉ Giorgio d'Agostino  
giorgio.dagostino@uniroma3.it

<sup>1</sup> Department of Economics, Roma Tre University, Via Silvio D'Amico 77, 00145 Rome, Italy

<sup>2</sup> Department of Political Science, University of Perugia, Via Alessandro Pascoli, 06123 Perugia, Italy

The question of exactly how corruption can be estimated has been widely debated (Kaufmann et al. 2007). In several cases, corruption is measured by drawing informal views of relevant stakeholders and the composite subjective index is an aggregate measure of these perceptions,<sup>2</sup> which involve an irreducible element of uncertainty<sup>3</sup> and a lack of transparency, especially in combining information from different sources. However, because corruption usually leaves no paper trail, perceptions of corruption based on an individual's actual experiences are viewed as the best—and sometimes the only—information available.

A related concern is the excessively narrow current definition of corruption, which is characteristically described as the misuse of public power for private benefit,<sup>4</sup> which leaves many conceptual ambiguities open to question (Svensson 2005). For example, although the term “private benefit” may describe the receipt of money or some kind of asset of value, it may involve increased political power or status, in addition to advantages deriving from receiving promises for future favours or benefits for closely connected persons (nepotism or favouritism). In addition, this definition does not consider corruption as a multidimensional phenomenon. For example, there is a certain degree of consensus in recognising at least two relevant dimensions of corruption—defined as *political and administrative corruption*—which should be taken into account when anti-corruption policies are implemented (Bardhan 1997, 2006; Warren 2004).

Based on this strand of the literature, this paper intends to examine some of the methodological issues concerning the identification and measurement of multidimensional corruption indices from multiple indicators. We exploit data from the Business Environment and Enterprise Performance Survey 2002 (BEEPS) to evaluate the multiple dimensions of corruption across the countries of Eastern Europe and Central Asia. The multidimensional framework (M-IRT) that was proposed by Bartolucci (2007) includes latent class (LC) analysis (Goodman 1974, 1978) along with Item Response Theory (IRT). It allows us to characterise the role that is played by the various types of information contained in responses to the questionnaire (items) and to identify the underlying LCs of corrupt

---

Footnote 1 (continued)

1995; Knack and Keefer 1995), while d'Agostino et al. (2016a, b) find that the negative impact of corruption at least doubles the reduction in the real per capita growth, irrespective of location.

<sup>2</sup> As alternatives to perception corruption measures, new data sources allow us to measure corruption based on actual episodes. For example, Seligson (2006) used victim surveys to obtain quantitative data on the prevalence of bribery, and Reinikka and Svensson (2006) used public expenditure tracking surveys to quantify embezzlement of public funds and enterprise surveys to quantify bribery at the micro level. Olken (2009), and Ferraz and Finan (2008) relied on external audits to measure fraud in local government offices. Gorodnichenko and Sabirianova-Peter (2007) use gaps between the incomes and expenses of public officials for similar purposes. Lastly, Dimitrova-Grajzl et al. (2012) measured an index of the perception of corruption using responses to eight questions regarding the frequency of making unofficial payments/gifts. In particular, these authors considered aspects such as resorting to the courts for civil matters and requesting official documents from authorities.

<sup>3</sup> The measurement error stems from two main problems: (1) a specific measurement noise in specific corruption measures, and (2) specific measures of corruption are imperfectly related to overall corruption (Kaufmann et al. 2007).

<sup>4</sup> For a discussion of the various definitions of corruption, see Lambsdorff (2007).

activities.<sup>5</sup> This method also includes the well-known multidimensional Rasch model (Rasch 1961), as a particular case.

Although the IRT model was initially used in psychometrics to examine ability/achievement tests (see, Bock 1997), it has since been applied in social and economic research. For example, Cappellari and Jenkins (2006) applied the IRT model to explain several problems in the deprivation scales of widely used sum-score deprivation indices. Similar methods were used by Kuklys (2004) to analyse housing and healthcare functioning and Faye et al. (2009) used a set of food insecurity indicators to derive a food deprivation scale. The M-IRT extension was also applied in identifying the latent health status of elderly patients who are currently receiving healthcare assistance in Italy (Bartolucci et al. 2010, 2012).

The proposed analysis identifies a multidimensional structure of corruption and finds that four LCs exist, with increasing levels of corruption. The contribution of our findings is emphasised by the robust association of our identified LCs of administrative corruption with the indices of (mainly) administrative corruption derived from BEEPS and the CPI of Transparency International (TI). In addition, our illustrative example, which uses our different LCs to estimate the relationship between firms' competitiveness and corruption, attempts to disentangle several contrasting expected signs that an aggregate indicator of corruption generally yields in the cited relationship.

The rest of this paper is organised as follows. Section 2 describes the dataset that is used for our analysis. Section 3 discusses the empirical framework that we use to quantifying the number of LCs and presents our main results. Section 4 shows the robustness of these results by comparing the aggregate indices of firms' components of corruption with those extracted directly from BEEPS and CPI of TI. Section 5 presents the estimates of the relationship between a firm's competitiveness and corruption using the identified LCs of corruption. Finally, some conclusions are drawn in Sect. 6.

## 2 Corruption Data from the BEEPS Firm Survey

Our analysis uses data on corrupt attitudes perceived by firms from BEEPS, which is jointly implemented by the European Bank for Reconstruction and Development (EBRD) and the World Bank. These surveys have been carried out, usually every three years, since 1999. They examine several important problems affecting firms in Eastern Europe and Central Asia, such as questions about firms' sales, investments, innovations, and access to financing, along with business regulations, taxation, and qualitative perception of the business environment. They also contain details of the firms' characteristics, such as sector and size in terms of employees. Although other firms' surveys have been conducted by the World Bank, the extension of these investigations were designed by a single country. In addition, this sample of microdata is of particular interest because it includes countries that had been undergoing transition towards freer markets and which were potentially associated with corrupt practices.

---

<sup>5</sup> The statistical literature provides alternative techniques for inferring a measurement scale of corruption from a list of pre-ordered (or pre-classified) items of the questionnaire. One basic approach is the sum-score technique, which simply consists of a weighed or unweighed sum of the indicators. Factor analysis techniques are also widely used to check whether or not a set of items fits a one-dimensional measurement scale. For example, Knack (2007) provides analyses of corruption using BEEPS data, descriptively confirming the distinction between political and administrative corruption.

One section of BEEPS is addressed to government policies and practices, and it includes items regarding perceived corruption by firms. One of the most interesting features of BEEPS is the sample design that is used to collect the data, which was of particular interest for our aims. Based on the managers' perceptions and related to the line of business in which they operate, these subjective measures of corrupt practices cover the interviewees' direct experience and they more clearly solicit answers on corruption. Obviously, the subject matter of BEEPS is sensitive. Consequently, it was natural to ask why firms would be willing to answer questions about their involvement in corruption. A number of features were incorporated into our survey to improve the quality of the data, according to three key components (Hellman et al. 2003): (1) before beginning an interview, the nature and purpose of the research project was explained; (2) data were recorded without any mention of the name of the respondent or the company, and they were collected by international organisations and not by national governmental authorities; and, (3) on sensitive topics, our questions were carefully worded to remove any implication of blame from the respondents, who were usually indirectly asked to describe the typical situation of a similar firm.<sup>6</sup>

In this paper, we used data from the survey carried out in 2002 because, in contrast to the other years, the data for that year do not contain many missing values in corruption responses. This provides some reassurance that most firms in the countries where BEEPS was carried out were in fact willing to talk about corruption, even if under-reporting remains an issue. If firms in some countries were simply unwilling to talk about corruption, then it seems logical that they would give favourable responses for corruption questions and, therefore, they would induce a downward bias.<sup>7</sup>

Our aim is to characterise corruption from different underlying latent causes. We exploit the potential of this questionnaire, which distinguishes *a priori* between items that can proxy *administrative* and *political* corruption. Administrative corruption, which is aimed at influencing the proper implementation of laws and regulations, is often referred to as bribery by individuals, groups or firms in the private sector. In contrast, political corruption refers to how firms influence the content and implementation of specific laws and regulations for the benefit of narrow private interests rather than the broad public interest (Scott 1972; Bardhan 1997, 2006; Warren 2004).

The questionnaire items are listed in Table 1. Looking at these proxied definitions of administrative corruption, we selected 10 items, covering a wide range of features related to corrupt practices (Anderson and Gray 2006). Three items concern corrupt payments in the field of public services and permits, which are closely connected with the fixed costs of doing business. A second set of items is linked with bribes to weaken the activities of public inspections within a firm, or inspections of buildings, health and safety regulations, and environmental safety. The remaining sub-set of items describes informal payments that are required in dealings with the public administration, imports and customs, courts and tax collection. It should be noted that Items 9 and 10 are included in the administrative corruption section because they reflect the use of bribery to speed procedures necessary for making overdue payments, and the influence—rather than manipulation—of the contents of new legislations and rules, respectively.

---

<sup>6</sup> See also Svensson (2003) for a discussion of the technique of administering questionnaires containing sensitive topics.

<sup>7</sup> While non-response may also be a 0, thereby inducing an upward bias on corruption measure, this is less reliable when we investigate corruption, which is mainly a crime.

**Table 1** Description of the dataset

	How often does your firm make payments/gifts for the following purposes (proxies of <i>administrative corruption</i> )
Item01	(1) To be connected to and maintain public services (e.g., electricity and telephone lines)
Item02	(2) To obtain business licenses and permits
Item03	(3) To obtain government contracts
Item04	(4) To deal with occupational health and safety inspections
Item05	(5) To deal with fire and building inspections
Item06	(6) To deal with environmental inspections
Item07	(7) To deal with taxes and tax collection
Item08	(8) To deal with customs/imports
Item09	(9) To deal with courts
Item10	(10) To influence the contents of new legislation rules, decrees, etc.
	To what extent have the following practices had a direct impact on your business (proxies of <i>political corruption</i> )
Item11	(11) Private payments/gifts or other benefits to parliamentarians to affect their votes
Item12	(12) Private payments/gifts or other benefits to government officials to affect the content of government decrees
Item13	(13) Private payments/gifts or other benefits to judges to affect decisions in criminal court cases
Item14	(14) Private payments/gifts or other benefits to judges to affect decisions in commercial cases
Item15	(15) Private payments/gifts or other benefits to central bank officials to affect central bank policies and decisions
Item16	(16) Illegal contributions to political parties and/or election campaigns to affect the decisions of elected officials
	Data from Business Environment and Enterprise Performance Surveys (BEEPS), jointly implemented with data from European Bank for Reconstruction and (EBRD) and World Bank

As regards political corruption—which in our case refers specifically to whether payments by other firms affect respondents' firms—a set of six items was used.<sup>8</sup> In particular, managers responded to questions regarding private payments or gifts that are made with the aim of affecting the votes of parliamentarians and government officials on specific laws, the contents of governmental decrees, and the decisions of elected officials. In addition, private payments or gifts were considered to potentially influence the decisions of judges working for the criminal and civil courts, together with benefits to bank officials in influencing central bank policies and decisions.

All of the items reported in the Table 1 range in an interval 0–4, where 0 represents the lowest degree of impact of the corruption item on the firms' business. For the purpose of the empirical analysis, these items are transformed in dummy variables recording as 0 when the original variable has values of 0 “no impact”, 1 “minor” and 2 “moderate”, and as 1 when the original variable has values of 3 “major” and 4 “decisive”.

Our dataset contains details of 6667 firms in 27 countries. Table 2 lists by country the descriptive statistics of firms and the itemised responses to the questions on administrative and political corruption, reporting the fraction responding “YES” in each set. Although the proportion of firms interviewed across countries does not differ greatly, the number of interviewed firms (minimum 170) was sufficient to allow us to extract consistent information on corruption issues within countries. In the appropriate columns of Table 2, the binary responses of items in terms of perception of corruption indicate not only the great variety of administrative and political corruption across countries but also the fact that the items referring to the two types of corruption do not necessarily go hand in hand. This preliminary evidence emphasises the idea of corruption as a complex phenomenon, which requires a suitable statistical technique to identify and measure the main characteristics of corruption.

### 3 Framework of Analysis

In this section, we describe a LC model that explains the multidimensionality of corruption. According to the constrained version of this model, we view the tendency to be involved in corruption as a latent characteristic affecting the conditional probability of responding in a certain way to each item. According to this characteristic, the respondents were grouped into a certain number of LCs. Statistical criteria were mainly used to choose the number of such classes.

#### 3.1 The Statistical Model

We first describe the M-IRT model proposed by Bartolucci (2007). Let  $y_j$  denote the response variable for the  $j$ -th item of the questionnaire, which is binary ( $y_j = 0, 1$ ). We denote by  $n$  the number of firms in the sample, and we assume that they respond to  $J$  items measuring  $D$  different latent characteristics (tendency to be involved in corruption) and we also assume that every item measures only one of these characteristics. This model allows us to examine the correlation between the latent dimensions of corruption. Here,  $n = 4610$

<sup>8</sup> See also Svensson (2005) for a discussion of the definition of political corruption.

**Table 2** Statistical distribution of firms and questionnaire items of perceived corruption, by country

Country	FIRMS	Items																		
		Administrative corruption								Political corruption										
		Number	(%)	1	2	3	4	5	6	7	8	9	10	Mean	11	12	13	14	15	16
Albania	170	2.55	0.71	0.51	0.49	0.44	0.42	0.40	0.79	0.74	0.52	0.33	0.53	0.32	0.38	0.43	0.48	0.20	0.43	0.37
Armenia	171	2.56	0.10	0.19	0.09	0.09	0.15	0.13	0.27	0.13	0.10	0.07	0.13	0.07	0.08	0.11	0.13	0.08	0.05	0.09
Azerbaijan	170	2.55	0.36	0.55	0.18	0.22	0.31	0.23	0.53	0.31	0.13	0.13	0.29	0.17	0.22	0.20	0.21	0.22	0.17	0.20
Belarus	250	3.75	0.19	0.48	0.29	0.40	0.48	0.30	0.33	0.33	0.24	0.09	0.31	0.05	0.08	0.12	0.22	0.14	0.07	0.11
Bosnia and Herze	182	2.73	0.52	0.62	0.58	0.64	0.56	0.45	0.61	0.59	0.59	0.49	0.56	0.40	0.40	0.49	0.50	0.42	0.48	0.45
Bulgaria	250	3.75	0.33	0.68	0.39	0.48	0.54	0.42	0.51	0.51	0.52	0.32	0.47	0.40	0.46	0.33	0.50	0.17	0.46	0.39
Croatia	187	2.80	0.31	0.51	0.40	0.48	0.39	0.37	0.42	0.43	0.40	0.33	0.40	0.19	0.23	0.22	0.28	0.17	0.28	0.23
Czech Rep.	268	4.02	0.22	0.48	0.44	0.47	0.38	0.37	0.33	0.30	0.21	0.16	0.34	0.18	0.17	0.22	0.23	0.13	0.18	0.19
Estonia	170	2.55	0.16	0.48	0.56	0.38	0.40	0.34	0.20	0.33	0.17	0.31	0.33	0.12	0.15	0.08	0.10	0.09	0.16	0.12
FYROM	170	2.55	0.43	0.62	0.48	0.61	0.51	0.41	0.43	0.39	0.43	0.29	0.46	0.44	0.40	0.41	0.48	0.39	0.46	0.43
Georgia	174	2.61	0.39	0.61	0.43	0.45	0.50	0.41	0.76	0.47	0.43	0.31	0.47	0.31	0.35	0.29	0.30	0.31	0.28	0.31
Hungary	250	3.75	0.24	0.24	0.45	0.37	0.34	0.32	0.27	0.33	0.25	0.24	0.30	0.11	0.14	0.10	0.12	0.10	0.17	0.12
Kazakhstan	250	3.75	0.26	0.46	0.30	0.24	0.42	0.29	0.55	0.29	0.24	0.11	0.32	0.12	0.10	0.11	0.16	0.09	0.06	0.11
Kyrgyz Rep.	173	2.59	0.39	0.58	0.38	0.35	0.53	0.48	0.76	0.38	0.41	0.26	0.45	0.27	0.25	0.26	0.35	0.17	0.16	0.24
Latvia	176	2.64	0.13	0.37	0.34	0.37	0.47	0.31	0.33	0.30	0.26	0.19	0.31	0.28	0.32	0.26	0.29	0.17	0.28	0.27
Lithuania	200	3.00	0.07	0.26	0.25	0.36	0.35	0.28	0.26	0.21	0.15	0.12	0.23	0.19	0.18	0.19	0.21	0.10	0.21	0.18
Moldova	174	2.61	0.38	0.67	0.23	0.34	0.56	0.40	0.68	0.51	0.17	0.14	0.41	0.23	0.28	0.26	0.49	0.23	0.21	0.28
Poland	500	7.50	0.16	0.29	0.33	0.38	0.35	0.31	0.27	0.25	0.19	0.20	0.27	0.13	0.15	0.12	0.17	0.09	0.20	0.14
Romania	255	3.82	0.25	0.59	0.36	0.58	0.44	0.44	0.38	0.39	0.37	0.22	0.40	0.22	0.22	0.20	0.26	0.16	0.25	0.22
Russia	506	7.59	0.36	0.52	0.30	0.36	0.60	0.33	0.54	0.25	0.19	0.17	0.36	0.09	0.13	0.12	0.18	0.07	0.09	0.11
Serbia and Monte	170	2.55	0.54	0.49	0.51	0.52	0.51	0.41	0.42	0.54	0.47	0.32	0.47	0.33	0.36	0.37	0.36	0.29	0.40	0.35
Slovak Rep.	188	2.82	0.17	0.57	0.55	0.39	0.40	0.38	0.39	0.47	0.38	0.21	0.39	0.20	0.23	0.24	0.32	0.13	0.27	0.23
Slovenia	176	2.64	0.19	0.28	0.24	0.24	0.21	0.19	0.16	0.16	0.13	0.16	0.20	0.22	0.24	0.16	0.20	0.15	0.20	0.20

Table 2 (continued)

Country	FIRMS	Items																		
		Administrative corruption								Political corruption										
		Number	(%)	1	2	3	4	5	6	7	8	9	10	Mean	11	12	13	14	15	16
Tajikistan	514	7.71	0.65	0.75	0.53	0.40	0.60	0.49	0.82	0.47	0.55	0.33	0.56	0.21	0.38	0.47	0.55	0.42	0.16	0.37
Turkey	463	6.94	0.43	0.54	0.32	0.30	0.32	0.34	0.44	0.40	0.28	0.23	0.36	0.34	0.35	0.28	0.33	0.30	0.35	0.33
Ukraine	260	3.90	0.32	0.53	0.38	0.54	0.59	0.40	0.55	0.33	0.30	0.15	0.41	0.13	0.17	0.19	0.27	0.11	0.20	0.18
Uzbekistan	250	3.75	0.22	0.41	0.27	0.22	0.37	0.26	0.50	0.24	0.23	0.12	0.28	0.18	0.26	0.26	0.30	0.28	0.17	0.24
Total	6667	100	0.31	0.48	0.36	0.39	0.44	0.35	0.46	0.35	0.29	0.21	0.36	0.20	0.24	0.22	0.28	0.18	0.23	0.22

Each item indicates share of successful response (modality 1) of corruption. Questionnaire items are listed in Table 1. Each item is constructed in the same way and are considered as 0 when the original variable has values of “no impact”, “minor” and “moderate” and as 1 when the original variable has values of “major” and decisive. “FYROM”: Former Yugoslav Republic of Macedonia



and  $J = 16$ , and, according to the items defined in the dataset, we expect at least two different dimensions, so that  $D \geq 2$ .

The model that we adopt is based on the following parameterisation of the conditional probabilities of success  $\lambda_{j|c}$  as a logit function:

$$\text{logit}(\lambda_{j|c}) = \gamma_j \left( \sum_d \delta_{jd} \theta_{cd} - \beta_j \right), \quad j = 1, \dots, J, \tag{1}$$

where with reference to item  $j$ ,  $\lambda_{j|c} = (p(y_j = 1) | \text{firm is in class } c)$  and  $\delta_{jd}$  is a dummy variable equal to 1 if  $j \in \mathcal{J}_d$  and to 0 otherwise, where  $\mathcal{J}_d$  is the sub-set of items measuring dimension  $d$ ;  $\theta_{cd}$  is a measure of the latent trait (dimension  $d$ ) for subjects in LC  $c$ , and  $\beta_j$  indicates the IRT *difficulty* parameter<sup>9</sup> as the overall tendency to respond 0 to item  $j$  (for a more detailed discussion of this point, see Bartolucci 2007). Parameters  $\gamma_j$  may be set equal to a fixed value according to one-parameter logistic parameterisation (1PL) or they may be left unconstrained and, thus, estimated as in two-parameter logistic parameterisation (2PL). In the second case, the model allows for the different sensitivity of the item measuring the latent trait. The relative importance of the difference between a firm’s trait level and the item threshold is, therefore, determined by the magnitude of the *discriminatory* power of item ( $\gamma_j$ ).

Under the assumption of *local independence*,<sup>10</sup> the distribution of  $\mathbf{Y}$  for subjects in the  $c$ -th LC is:

$$p(\mathbf{y}|c) = p(\mathbf{Y} = \mathbf{y} | \Theta = \theta_c) = \prod_j \lambda_{j|c}^{y_j} (1 - \lambda_{j|c})^{1-y_j}, \tag{2}$$

where  $p(\mathbf{Y} = \mathbf{y} | \Theta = \theta_c)$  is the conditional probability that a subject with latent vector  $\theta$  provides to response configuration  $\mathbf{y}$ . Through a finite mixture, we can express the distribution of  $\mathbf{Y}$  as:

$$p(\mathbf{y}) = \sum_c p(\mathbf{y}|c) \pi_c, \tag{3}$$

where  $\pi_c = p(\Theta = \theta_c)$  are the weights corresponding to each LC.

The log-likelihood function, which is used to estimate the above multidimensional LC Rash model, is thus:

$$\ell(\boldsymbol{\phi}) = \sum_{\mathbf{y}} n(\mathbf{y}) \log(p(\mathbf{y})), \tag{4}$$

where  $\boldsymbol{\phi}$  is the vector containing all identifiable parameters of the model and  $p(\mathbf{y})$  is computed as a function of  $\boldsymbol{\phi}$ . In particular, to make the parameters identifiable, we use constraint  $\beta_j = 0$ , when the parameterisation is of 1PL type, and  $\beta_j = 0, \gamma_j = 1$  when it is of 2PL type, where  $J = j_1, \dots, j_D$ , and  $j_d$  denotes a specific element of  $\mathcal{J}_d$ .

<sup>9</sup> In a typical application of the IRT model, the response variable is referred to test item behaviours and is 1 when we have a correct response. Because the probability of this event decreases with  $\beta_j$ , this is called difficulty parameter.

<sup>10</sup> Following the *local independence* assumption, the response variables are conditionally independent, given the latent variables.

To estimate the model parameters, we maximise log-likelihood  $\ell(\boldsymbol{\phi})$  by the EM algorithm (Dempster et al. 1977); for further details, see also Bartolucci (2007) and Bartolucci et al. (2012). Let us assume that we know frequencies  $m(\mathbf{y}, c)$  of the contingency table in which these subjects are cross-classified according to response configuration ( $\mathbf{y}$ ) and the LC ( $c$ ). The EM algorithm is based on the so-called *complete* log-likelihood:

$$\ell^*(\boldsymbol{\phi}) = \sum_{\mathbf{y}} \sum_c m(\mathbf{y}, c) \log p(c, \mathbf{y}). \tag{5}$$

At the E-step of this algorithm, we compute the expected value of  $\bar{m}(\mathbf{y}, c)$  for each  $\mathbf{y}$  and  $c$ , given  $n(\mathbf{y})$ , and at the M-step we maximise *complete* log-likelihood, in which every frequency  $m(\mathbf{y}, c)$  is replaced by the corresponding expected value  $\hat{m}(\mathbf{y}, c)$ . The EM alternates these two steps until convergence.

The estimated parameters from the EM algorithm are then used to compare (two) multidimensional models. The hypothesis test is of type:

$$H_0 : \mathbf{g}(\boldsymbol{\phi}) = \mathbf{0} \tag{6}$$

where  $\mathbf{g}(\boldsymbol{\phi})$  is a vector-valued function and  $\mathbf{0}$  denotes the column vector of zeros of suitable dimensions. To select the most suitable model, we use the likelihood ratio (LR) test statistic

$$LR = 2 \sum_{\mathbf{y}} n(\mathbf{y}) \log \left[ \frac{\hat{p}_D(\mathbf{y})}{\hat{p}_{D-1}(\mathbf{y})} \right] \tag{7}$$

where  $\hat{p}_D(\mathbf{y})$  refers to the assumed model with  $D$  different dimensions and  $\hat{p}_{D-1}(\mathbf{y})$  to the constrained model with  $D - 1$  dimensions. When the response probability is modelled by an M-IRT model, the resulting  $LR$  statistic has a  $\chi^2$  distribution with a number of degrees of freedom equal to the difference between the number of parameters of the full multidimensional and restricted models, in which we merge two distinct dimensions.

This method is used to cluster items in homogeneous groups, corresponding to various types of corruption. A crucial point in identifying various types of corruption is the strength of the correlation between two distinct dimensions. We compute this correlation as:

$$\hat{\rho}_{d_1, d_2} = \sum_{c=1}^C \hat{\theta}_{c, d_1} \hat{\theta}_{c, d_2} \pi_c \tag{8}$$

where  $\theta_{c, d_1}$  and  $\theta_{c, d_2}$  are the standardised estimates of the support points referring to two specific dimensions,  $d_1$  and  $d_2$ , for subjects in LC  $c$  (e.g.,  $\hat{\pi}_c$  is the estimated weight of this class with  $c = 1, \dots, C$ ). After identifying the LCs in according to  $\hat{\theta}_{cd}$  and  $\hat{\rho}_{d_1, d_2}$ , as a final step, we compute the firm's trait scores by using the expected value of frequencies  $\hat{m}(\mathbf{y}, c)$  estimated at the E-step, defined as:

$$\hat{m}_{\mathbf{y}|c} = n(\mathbf{y}) \frac{p(\mathbf{y}|c)\pi_c}{\sum_h p(\mathbf{y}|h)\pi_h}. \tag{9}$$

### 3.2 Results

Before presenting the main results of the analysis, we have to choose the number of classes, which is crucial for model identification in the multidimensional LC Rash

**Table 3** Output of hierarchical cluster analysis

Seq.	Dim.	Clusters	Deviance	LR test	<i>p</i> value
1	15	1.3.4.5.6.7.9.10.11.12.13.14.15.16.(2.8)	0.194	0.194	0.908
2	14	1.3.4.5.6.7.11.12.13.14.15.16.(2.8).(9.10)	0.500	0.306	0.858
3	13	1.3.4.7.11.12.13.14.15.16.(2.8).(9.10).(5.6)	1.109	0.609	0.737
4	12	1.3.4.7.11.12.14.16.(2.8).(9.10).(5.6).(13.15)	1.789	0.680	0.712
5	11	3.4.11.12.14.16.(2.8).(9.10).(5.6).(13.15).(1.7)	2.690	0.901	0.637
6	10	3.4.11.12.14.(2.8).(9.10).(5.6).(1.7).(13.15.16)	3.939	1.249	0.535
7	9	3.4.11.12.14.(9.10).(5.6).(13.15.16).(1.2.7.8)	6.185	2.246	0.325
8	8	3.4.14.(9.10).(5.6).(13.15.16).(1.2.7.8).(11.12)	9.666	3.481	0.175
9	7	3.4.(9.10).(5.6).(1.2.7.8).(11.12).(13.14.15.16)	13.296	3.630	0.163
10	6	3.(9.10).(4.5.6).(1.2.7.8).(11.12).(13.14.15.16)	18.388	4.592	0.101
11	5	(9.10).(4.5.6).(1.2.3.7.8).(11.12).(13.14.15.16)	23.668	5.580	0.061
12	4	(9.10).(4.5.6).(1.2.3.7.8).(11.12.13.14.15.16)	42.936	19.268	0.000
13	3	(4.5.6).(1.2.3.7.8.9.10).(11.12.13.14.15.16)	84.831	41.895	0.000
14	2	(1.2.3.4.5.6.7.8.9.10).(11.12.13.14.15.16)	207.149	122.318	0.000
15	1	(1.2.3.4.5.6.7.8.9.10.11.12.13.14.15.16)	3059.953	2852.805	0.000

Column (1) lists the sequence of steps (seq), Column 2 lists the number of dimensions for each step (dim), and Column 3 lists the combinations of items in each dimension for each sequential step (clusters). Deviance for each step (Deviance), LR test (LR test), and *p* values appear in last three columns. For classifications and definitions of items, see Table 1

model. Although this selection is mainly based on information criteria [i.e., BIC and consistent Akaike information criterion (e.g., *CAIC*)], we complement it with indices measuring goodness-of-fit. A detailed description of the proposed model selection strategy is reported in Table 9 of “Appendix A”. This analysis is based on information criteria and it strongly suggests that the optimal number of LCs is 4, this result is also confirmed by the goodness-of-fit measures (which are given at the end of the table).

We will now summarise the main results of the multidimensional approach to corruption. As described in the previous section, the procedure consists of running a sequence of nested models beginning with an LC model, in which the number of dimensions is equal to the number of selected items (unconstrained model). An analytical overview of the results from hierarchical clustering analysis is given in Table 3. Column 3 lists the combinations of items for each of the sequential steps; Columns 4 and 5 list the deviance from the initial LC model and from the previous model, respectively. Finally, the *p* value of LR test statistics appears in Column 6.

Following the LR test proposed in Eq. (6), Table 3 lists multidimensionality, keeping five significant dimensions, as also confirmed by hierarchical clustering analysis (see, Appendix B). Differences within types of corruption appear in the items of administrative corruption, which characterise Dimensions 1, 2, and 3; the items of political corruption are given in Dimensions 4 and 5.

Looking at the item definitions (Table 1), we can progressively identify the dimensions of administrative corruption as follows:

- Bureaucrats’ need to influence legally established regulations and the timing of applications of laws;

**Table 4** Estimates of conditional probabilities and discriminatory power of items of selected model

Dim.	Item	Class1	Class2	Class3	Class4	$\gamma$
3	Item1	0.044	0.194	0.433	0.777	1.000
3	Item2	0.102	0.417	0.723	0.934	1.118
3	Item3	0.055	0.248	0.528	0.848	1.057
2	Item4	0.031	0.144	0.644	0.881	1.000
2	Item5	0.046	0.205	0.735	0.919	1.001
2	Item6	0.012	0.080	0.596	0.887	1.192
3	Item7	0.075	0.371	0.705	0.937	1.208
3	Item8	0.047	0.226	0.507	0.842	1.086
1	Item9	0.011	0.206	0.371	0.877	1.000
1	Item10	0.011	0.139	0.246	0.730	0.850
4	Item11	0.005	0.657	0.035	0.731	1.000
4	Item12	0.008	0.738	0.055	0.799	0.979
5	Item13	0.002	0.649	0.038	0.822	1.000
5	Item14	0.008	0.800	0.113	0.900	0.895
5	Item15	0.002	0.499	0.035	0.686	0.861
5	Item16	0.008	0.613	0.081	0.759	0.752

Column (1) lists dimension of each item (dimension) and column 2 item code (item). Conditional probabilities and discriminatory power ( $\gamma$ ) of items of selected model. For classifications and definitions of items, see Table 1

- Unofficial payments for inspections in the fields of occupational health and safety, fire and buildings, and environmental works;
- Application of government contracts, business licences and contracts, public services, and general taxation problems.

The items which cluster political corruption are:

- Corruption by firms to influence the contents of specific laws and regulations;
- Corruption to influence decisions by the judges of criminal courts and commercial cases, and decisions or policies of the officials of the central bank.

Table 4 lists the estimated conditional probabilities of the LC model for each selected item, together with estimates of parameters  $\gamma_j$ . It shows that there is a dissimilar order between the first 10 items (administrative corruption) and the last 6 (political corruption). For a clearer interpretation, we refer to the components of *breadth* and *depth* of corruption, as defined by Mocan (2008). The former describes the extent to which corruption is widespread in a country, and the latter describes the extent to which each LC component affects corruption. Our results indicate that administrative corruption items follow an increasing order across the four LCs, and the last six political corruption items show higher conditional probabilities for LC 1, which achieves a (modest) influence in Items 2 and 7 (see Table 4), which are associated with administrative corruption. This implies that the estimated level of corruption in LC 1, mainly corresponding to tax evasion or improving access to public services, does not seem to be an obstacle for the business environment. However, as argued previously, we cannot exclude the possibility that firms under-report the true perception of

**Table 5** Estimates of  $\theta$  and correlations among dimensions

	Dim1	Dim2	Dim3	Dim4	Dim5	Weight
Impact of dimensions on classes						
Class1	-3.072	-3.027	-4.498	-6.372	-5.326	0.439
Class2	-1.427	-1.352	-1.350	0.615	0.651	0.108
Class3	-0.269	1.020	-0.526	-3.233	-3.311	0.313
Class4	1.247	2.425	1.961	1.527	1.000	0.140
	Dim1	Dim2	Dim3	Dim4	Dim5	
Correlation among dimensions						
Dim1	1.000					
Dim2	0.989	1.000				
Dim3	0.990	0.959	1.000			
Dim4	0.747	0.645	0.834	1.000		
Dim5	0.676	0.563	0.773	0.994	1.000	

The first part of table lists standardised estimate of latent trait level ( $\theta_{c,d}$ ) for each class  $c$  and dimension  $d$ , with estimated weight in each class ( $\pi_c$ ). The second part lists correlations among dimensions

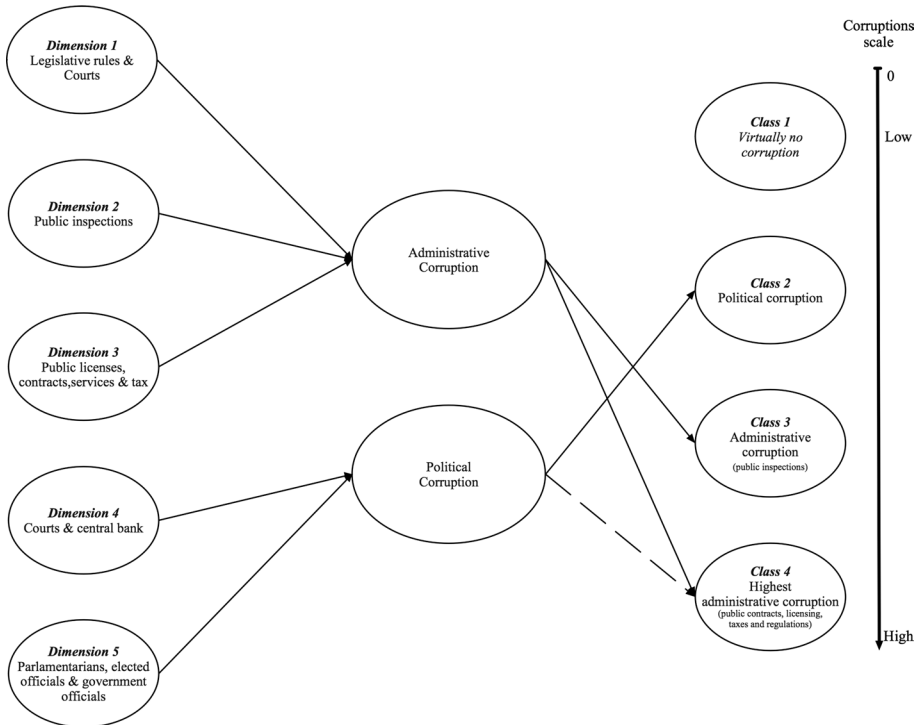
corruption, or at least do not provide answers to questions on this sensitive topic. Accordingly, we exercise caution by classifying the estimates by the conditional probability of this LC and refer to *virtually no corruption*.

This discussion suggests a crucial point of the present study is the interpretation of the LCs. Useful suggestions for this point come from the estimate of dimensions on different classes and from the correlation measures between identified dimensions (Table 5).

This analysis has three main results. First, estimated parameter  $\theta_{c,d}$  in Eq. (1) formally identifies the four LCs according to their different dimensions. As shown in the upper part of Table 5, within the four LCs, all five dimensions are increasingly ordered, the lowest values appearing in class 1 and the highest in class 4. This also means that we can order corruption levels progressively.

Second, as a brief index of how widespread corruption is in firms, we observe that most firms (43.9%) belong to Class 1, followed by Class 3 (31.3%), and Class 2 is the smallest with only 10.8%. The remaining firms are in Class 4 (14%).

Third, looking at the values of the dimensions in each class, we identify Class 3 as the class that has problems in administrative corruption, and Class 2 as characteristically political corruption. In more detail, and according to the results shown in Table 4, because Dimension 2 prevails in LC 3, we can identify it as corresponding to administrative corruption because it is more closely linked with public inspections, apart from taxes, licences and permits. In addition, although the items concerning political corruption may partly affect the results of Class 4, the prevalence in that class of the effects of administrative corruption Dimensions 1, 2, and 3 shows that the highest level of corruption is more closely associated with general administrative corruption. Figure 1 shows the identification process derived from the empirical framework.



**Fig. 1** Identification of corruption indices. *Notes* Analytic description of LCs and dimensions, according to results of Table 5. The continuous vertical arrow (right) measures scale of corruption, from 0 (absence of corruption) to high

## 4 Comparative Analysis

In this section, the expected value of posterior probabilities, computed at the E-step, allows us to build a binary variable in which each firm is attributed 1 in the LC with the highest level of posterior probability, and is 0 otherwise. This new variable identifies the perceived prevalence of corrupt practices for a firm classified in one of the LCs.

We then assess the robustness of the results by: (1) comparing our estimated corruption variables with a corruption ranking index extracted from the BEEPS, and (2) comparing the country's perception corruption ranking with CPI. In the first case, we compare our estimates with the perceived index of administrative corruption, which was directly obtained from the BEEPS dataset in which firms responded to the following question: "Is it common for firms in your line of business to have to make irregular additional payments or gifts to get things done as regards customs, taxes, licenses, regulation services etc.?". The answer includes six modalities on an increasing scale of perceived corruption (*Never, Seldom, Sometimes, Frequently, Usually, Always*). However, for empirical comparison with our four LCs, we rescale it into four modalities by combining responses *Seldom* and *Sometimes* and *Frequently* and *Usually*. The new corruption index thus mentions *Absence, Low, Medium, and High*, which tests an increase in the quantitative perception of corruption and is in line with the estimated LCs.

Table 6 lists the results of the LR test statistic in the 27 countries, which assumes an independence hypothesis between each estimated LC and the corresponding binary index at firm-level of BEEPS administrative corruption. LCs 1 and 4 appear to be closely associated with the directly perceived administrative corruption index. Here we discuss the novelty of the proposed corruption measure by characterising other identified LCs. In Class 3, the data reject the independence hypothesis at the 5% level between our firm score variable and that obtained with *medium corruption*, as extracted directly from BEEPS. We only note a non-significant association for a few countries. This result strengthens our findings and it indicates that although the index was built with a direct item of the perceived administrative corruption in BEEPS, its specificity generates the clear-cut effect of public inspections and tax evasion in some countries with respect to general administrative corruption (e.g., Bulgaria, Croatia, Hungary, Latvia, Uzbekistan and Yugoslavia). Finally, Column 2 lists the results of a “false experiment” specification, in which *low administrative corruption*, as extracted from BEEPS, was expected to be independent of the political corruption variable estimated as LC 2. We find confirmation that these two components of corruption are not related and that, independently of corruption level, they affect country corruption differently. Appendix C also graphically illustrates the correlation results between the estimated LCs and the perception of corruption index extracted from BEEPS. These indices are aggregated at country level by counting the “1” in each firm of the LC and in the corruption modalities of the variables that were extracted from BEEPS. In line with the results of the LR test statistic, there is a significant correlation between the identified administrative LC and the BEEPS modalities, whereas the political corruption class is not correlated with the BEEPS measure (the pairwise correlation coefficient is only 0.038).

Second, to validate our estimates, we exploit the perceived corruption scores at country level—the maximum value of the index represents the worst performance—to obtain the corresponding ranking. In particular, to obtain the country index from the LCs, we use the average of positive responses for each country and we weight these averages by the sample size of each country. Then, from these country averages, we create four different corruption ranking variables in a range of 0–100, where 0 records the lowest value of the variable at country level. We then compare these rankings with those of the CPI, which is the most prominent index measuring corruption at country level. It should be noted that CPI represents the best government performance with the maximum value of the index (CPI perception index range: 0–10). This comparison is particularly important because it suggests the potential external validity of our estimating procedure in building a country-ranking index.

Our hypothesis considers that, if our score estimates for each LC are correct, then the specific LC related to “virtually no corruption” should partially reproduce the CPI country classification because the latter measures the best government performance with the maximum value of the index. The analysis in Table 7 shows that in some cases (i.e. Bosnia & Herzegovina, Czech Republic, Georgia, Lithuania, Moldova, Romania, Russia, Slovenia and Ukraine) the rankings of the first latent class are in line with those of CPI. In other cases (i.e. Azerbaijan, Bulgaria, Kazakhstan, Slovak Republic and Uzbekistan) we do not find a link between the two indices. The explanation given here, which generally reflects criticism reported in the literature, is that CPI scores are not weighted means of policy and that certain firm-level characteristics influence corruption which we instead extract from the multidimensional framework adopted.

Some examples serve to clarify the main features of these LC differences. In Kazakhstan, the high estimated score in public inspections leads to an increase in the mean level of the perception of corruption, as measured by CPI. In Azerbaijan and

**Table 6** Tests for independence between identified LCs and direct BEEPS measures of corruption

	Virtually no corruption (LC1)		Political corruption (LC2)		Administrative corruption (inspections) (LC3)		Highest administrative corruption level (public contracts, licensing taxes and regulations) (LC4)	
	VS	BEEPS index of corruption (absence)	VS	BEEPS index of corruption (low)	VS	BEEPS index of corruption (medium)	VS	BEEPS index of corruption (high)
Albania	213.708*	0.811	57.638*	101.677*	(0.000)	(0.368)	(0.001)	(0.001)
Armenia	172.723*	0.064	49.305*	58.879*	(0.000)	(0.801)	(0.015)	(0.015)
Azerbaijan	370.101*	0.055	113.989*	49.467*	(0.000)	(0.815)	(0.026)	(0.026)
Belarus	321.844*	15.058	126.818*	159.203*	(0.000)	(0.220)	(0.000)	(0.000)
Bosnia & Herzegovina	688.497*	12.560	429.223*	211.593*	(0.000)	(0.262)	(0.000)	(0.000)
Bulgaria	412.088*	22.689	0.156	226.317*	(0.000)	(0.132)	(0.000)	(0.000)
Croatia	202.493*	82.793*	0.554	119.937*	(0.000)	(0.004)	(0.000)	(0.000)
Czech Republic	664.122*	10.352	392.161*	137.012*	(0.000)	(0.309)	(0.000)	(0.000)
Estonia	599.359*	35.750	373.577*	116.024*	(0.000)	(0.059)	(0.000)	(0.000)
FYROM	600.870*	0.756	365.622*	80.978*	(0.000)	(0.385)	(0.000)	(0.000)



Table 6 (continued)

	Virtually no corruption (LC1)		Political corruption (LC2)		Administrative corruption (inspections) (LC3)		Highest administrative corruption level (public contracts, licensing taxes and regulations) (LC4)	
	VS	BEEPS index of corruption (absence)	VS	BEEPS index of corruption (low)	VS	BEEPS index of corruption (medium)	VS	BEEPS index of corruption (high)
Georgia	496.622* (0.000)	0.004 (0.949)	265.603* (0.000)	122.376* (0.000)				
Hungary	258.069* (0.000)	50.323* (0.025)	12.937 (0.255)	152.746* (0.000)				
Kazakhstan	255.949* (0.000)	29.099 (0.088)	45.286* (0.033)	71.213* (0.008)				
Kyrgyzstan	268.411* (0.000)	0.169 (0.681)	92.779* (0.002)	91.552* (0.002)				
Latvia	428.327* (0.000)	54.366* (0.020)	16.125 (0.204)	75.011* (0.006)				
Lithuania	323.478* (0.000)	0.445 (0.505)	143.077* (0.000)	65.670* (0.010)				
Moldova	170.909* (0.000)	0.546 (0.460)	52.514* (0.022)	59.833* (0.014)				
Poland	417.191* (0.000)	27.214 (0.099)	161.722* (0.000)	120.334* (0.001)				
Romania	385.822* (0.000)	20.116 (0.156)	182.290* (0.000)	55.487* (0.018)				
Russia	399.378* (0.000)	23.117 (0.128)	42.141* (0.040)	114.458* (0.001)				

Table 6 (continued)

	Virtually no corruption (LC1)		Political corruption (LC2)		Administrative corruption (inspections) (LC3)		Highest administrative corruption level (public contracts, licensing taxes and regulations) (LC4)	
	VS	BEEPS index of corruption (absence)	VS	BEEPS index of corruption (low)	VS	BEEPS index of corruption (medium)	VS	BEEPS index of corruption (high)
Slovak Republic	81.053* (0.004)	0.107 (0.744)		113.735* (0.001)		20.003 (0.157)		
Slovenia	448.439* (0.000)	0.612 (0.434)		301.775* (0.000)		42.201* (0.040)		
Tajikistan	389.076* (0.000)	11.664 (0.280)		199.459* (0.000)		54.259* (0.020)		
Turkey	525.004* (0.000)	59.588* (0.015)		123.636* (0.000)		190.284* (0.000)		
Ukraine	941.732* (0.000)	0.866 (0.352)		812.055* (0.000)		52.655* (0.022)		
Uzbekistan	479.520* (0.000)	0.295 (0.587)		10.885 (0.297)		149.778* (0.000)		
Yugoslavia	292.500* (0.000)	0.849 (0.357)		34.337 (0.064)		62.623* (0.012)		

The independence test is the LR statistic test at country level. Asterisks: 5% significance level. All of our tests have as a null hypothesis the independence between our estimated country score indicators of corruption and those of BEEPS

Uzbekistan, the difference between our estimated index of virtually no corruption and CPI is shown by the relatively greater perception of political corruption. Instead, CPI estimates an excellent ranking for Estonia, which is probably associated with political corruption. The positions for the Slovak Republic and Bulgaria are more distant on the corruption index. CPI indicates that these two countries are in positions 3 and 7, respectively. In 2002, Bulgaria experienced the highest degree of political corruption and it was greatly constrained by the inefficiencies, or failure, of the legal system (Delavallade 2011).

Country rankings for administrative corruption in LCs 3 and 4 are particularly affected by the level of taxation, permits and inspections, and government contracts. While government contracts are linked to regulatory quality and the enforcement of property rights. Restrictive regulation and taxation, which are associated with insufficient enforcement of property rights, have been proven to be key determinants of administrative corruption in Russia, Ukraine, Moldova and Albania (Tanzi 1998).

The determinants of corruption are slightly different from the situation which arose in Bosnia with respect to the former USSR countries, in which administrative corruption is shown to rise with weak enforcement of property rights, revealing the incapacity of the courts to implement the law (Johnson et al. 1999). As an example of the highest levels of administrative corruption in Bosnia, Nowak (2001) attributed the space for corrupt activities to the post-war context, which complicated power structures and fragmented administration. Figure 2 graphically summarises the ranking of corruption components in the Eastern Europe countries.

## 5 An Illustrative Example: The Relationship Between Firm Competitiveness and Corruption

In this section, we will revise the empirical relationship between competitiveness and corruption by exploiting the results of the M-IRT model, which identifies four monotonically increasing LCs. Many papers have been published about the determinants of corruption, with particular emphasis on the role of market competition (Ades and Di Tella 1999; Herzfeld and Weiss 2003; Dreher et al. 2007; Pieroni and d'Agostino 2013). In the traditional view, the competitiveness of firms is presumed to lead to more corruption because they may spend part of their profits on bribes paid to public officials.<sup>11</sup> Ades and Tella (1997) and Clarke and Xu (2002) empirically examined the hypothesis of a positive correlation between firm competitiveness with corruption in East European and Central Asian countries. Their results support the mechanism in question; that is, the most competitive firms have much more money with which to bribe bureaucrats and, thus, increase corruption. These findings are in line with the estimates regarding the same relationship by Svensson (2003) and Reinikka and Svensson (2006) in the case study of Uganda. However, this economic literature also proposed and tested an alternative hypothesis: if comparative disadvantages in competitiveness exist, then firms invest in bribery to stay in the market. This hypothesis was supported by Gauthier and Reinikka (2001) and was empirically confirmed

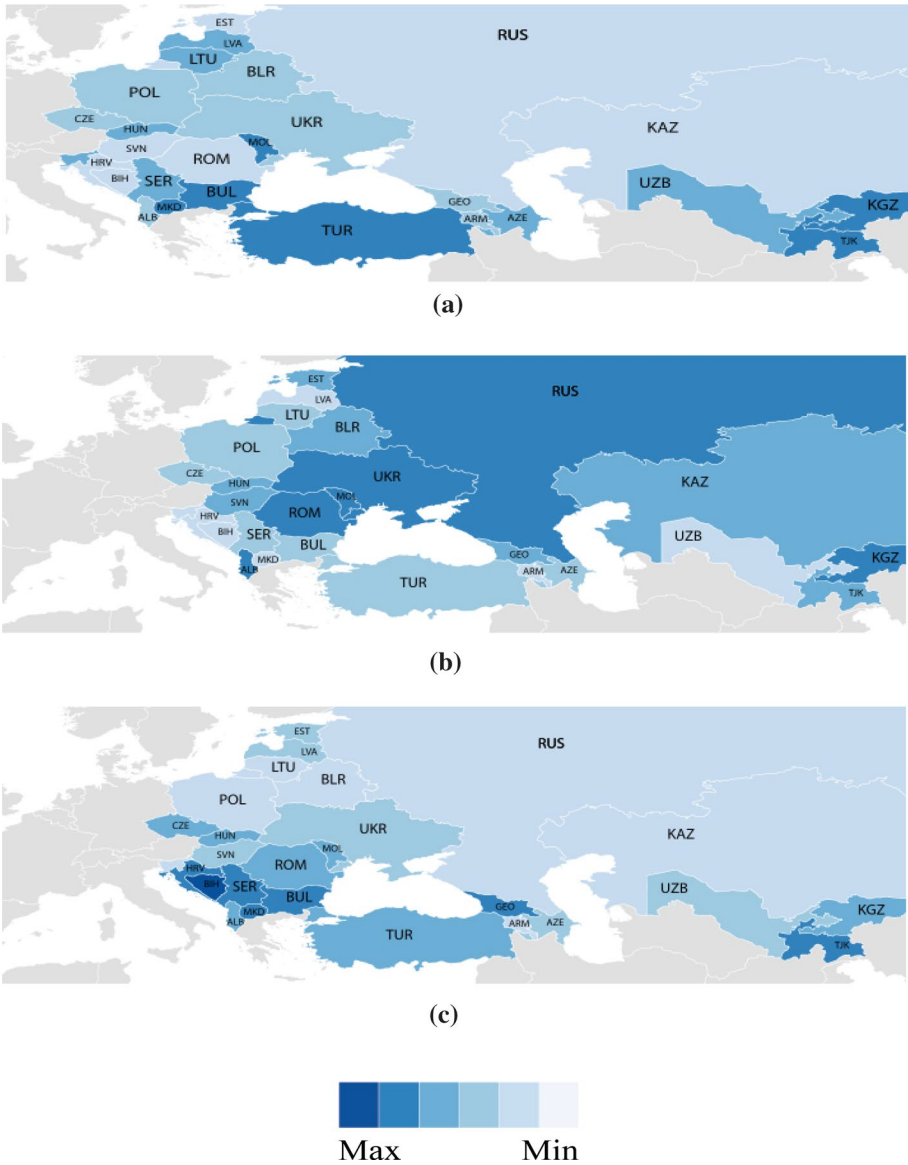
---

<sup>11</sup> In particular, the model developed by Bliss and Tella (1997) found that a firm's corruption increases with its profitability, so that less competitive firms leave the market. In turn, higher profitability increases bribery, consolidating a non-transparent system of behaviour on the part of bureaucrats.

**Table 7** Country corruption scores and rankings

Country	Virtually no corruption (LC1)		Political corruption (LC2)		Administrative corruption (inspections) (LC3)		Highest administrative corruption level (public contracts licenses, taxes and regulations) (LC4)		CPI (2002) index	
	Score	Ranking	Score	Ranking	Score	Ranking	Score	Ranking	Score	Ranking
	Albania	21.40	(22)	10.20	(13)	49.00	(22)	19.40	(18)	4.16
Azerbaijan	47.90	(12)	14.60	(19)	27.10	(8)	10.40	(8)	3.33	(22)
Belarus	49.20	(9)	6.20	(8)	37.80	(16)	6.70	(3)	8.00	(4)
Bosnia & Herzegovina	27.80	(19)	5.600	(5)	21.10	(4)	45.6	(22)	3.66	(20)
Bulgaria	23.00	(21)	23.60	(22)	29.90	(11)	23.60	(20)	6.66	(7)
Croatia	51.80	(6)	5.50	(4)	20.00	(3)	22.70	(19)	6.33	(8)
Czech Republic	48.30	(10)	7.70	(11)	29.50	(10)	14.50	(15)	6.16	(9)
Estonia	48.00	(11)	4.00	(2)	36.00	(15)	12.00	(9)	9.33	(2)
Georgia	32.30	(18)	9.80	(12)	33.10	(14)	24.80	(21)	4.00	(18)
Hungary	54.10	(5)	2.60	(1)	33.00	(13)	10.30	(7)	4.90	(10)
Kazakhstan	49.50	(8)	6.00	(6)	39.10	(17)	5.40	(1)	3.83	(19)
Latvia	51.50	(7)	11.30	(14)	23.70	(5)	13.40	(12)	6.16	(11)
Lithuania	55.30	(4)	12.90	(16)	24.70	(6)	7.10	(4)	8.00	(5)
Moldova	27.00	(20)	18.30	(20)	41.30	(18)	13.50	(13)	35.0	(21)
Poland	57.00	(3)	7.20	(10)	28.70	(9)	8.00	(6)	6.66	(6)
Romania	37.10	(16)	6.20	(7)	41.60	(19)	15.20	(17)	4.33	(15)
Russia	41.40	(14)	4.80	(3)	48.40	(21)	6.10	(2)	4.50	(14)
Slovak Republic	41.80	(13)	11.80	(15)	31.80	(12)	14.50	(16)	8.16	(3)
Slovenia	60.40	(1)	14.20	(18)	18.30	(2)	7.10	(5)	10.00	(1)
Turkey	39.40	(15)	19.70	(21)	26.50	(7)	14.40	(14)	5.333	(12)
Ukraine	36.40	(17)	7.00	(9)	44.30	(20)	12.30	(10)	4.00	(17)
Uzbekistan	58.00	(2)	14.20	(17)	17.30	(1)	12.30	(11)	4.83	(13)

We have excluded Yugoslavia, FYROM, Hungary, Tajikistan and Kyrgyzstan from our comparisons because TI does not report their country measures for 2002. The ranking for each country is given in round brackets. Columns 1 and 5 are ranked in terms of good government (100 = best degree of government institutions). Columns 2 to 4 are ranked in terms of highest corruption level (100 = most corrupt countries)



**Fig. 2** Representation of types of corruption, by country. **a** Political corruption. **b** Administrative corruption (inspections). **c** Highest administrative corruption level (public contracts, licenses, taxes and regulations). *Notes* Corruption by country is constructed by corruption scores of Table 7

for a sample of North African countries by Delavallade (2011), in which a negative relationship was found.

To illustrate the role of firm competitiveness as a determinant of corruption, the literature generally distinguishes two components: administrative and state capture indices (which we call political corruption). The main shortcoming of this approach is that the

empirical tests must be carried out to ascertain whether or not the aggregate index of corruption can be pooled with either administrative or political corruption. It is certainly possible that perceived corruption indicates the existence of an overlap between these indices. Instead, our model allows some firms in these LCs of corruption to be similar to others, although they comprise statistically distinct groups.

Thus, we set up an empirical framework based on logit models, in which the dependent variables are firms' dummy variables as derived from the estimates of the expected value of frequencies (see Sect. 4). Competitiveness is also a discrete variable with three modalities, in which the highest score indicates greater competitiveness and is proxied by the increase in sales over the previous two years (i.e., sales in 1999). We extend the concept of competitiveness by also including firms which, at least, did not suffer reduced sales. This allows us to avoid the general criticism concerning firm competitiveness linked with increased sales (i.e., to maximise firm profits) and we substitute it with the more general aim of firm survival in a competitive market.

The model specification includes the size of the firm (*size*) and the size of the sector in which it operates (*sector*). Following the results of Hellman et al. (2000), we assume that small firms tend to engage more in administrative than in political corruption, probably because the former is less costly. We include the variable *size* dichotomised, which is based on full-time employees of firms calculated as "small" (e.g., 2–49, reference modality), "medium" (e.g., 50–249), or "large" (e.g., > 250), expecting the size of the firm to negatively affect administrative corruption. In addition, this specific literature finds conflicting evidence that corruption differs among sectors. Industrial sectors, in which projects involve large amounts of money or high rent-generating public procurements, may be more open to corruption, particularly of the political kind. Thus, we consider a dummy variable according to whether a firm works in the "manufacturing" or "services" sector, the latter being the reference modality.

The coefficient estimates of the logit model are listed in Table 8 in terms of odds ratios. We emphasise three important results. First, firm competitiveness has a positive effect in cases where there is a perception of virtually no corruption (e.g., LC 1) but it has a negative effect when the weight of corruption increases. Long-term survival strategy is characterised by firms that are stable or which increase their competitiveness, indicating that a representative firm has less need to bribe bureaucrats to obtain, for example, public procurements in its line of business. We estimate that, at most, few corrupt activities (i.e., virtually no corruption) does not constrain economic development, finding that competitive (stable) firms increase the probability of not operating in corrupt countries by about 40%. However, the effects of competitiveness change for components characterised by higher levels of corruption, starting from the political kind. Competitive firms that occupy a stable position in their markets have a significantly lower probability than less competitive firms of being involved in corruption. This may imply that less competitive firms resort to unofficial payments to compensate for their competitive position, distorting the rules of competition for countries mainly affected by high levels of corrupt activities.

Second, by including dummies for the specific effects of size, we find significant coefficients in the LCs identified as the highest levels of corruption. In particular, large enterprises reduce the propensity (e.g., odds ratio = 0.771 and 0.805, for corruption components 3 and 4, respectively) to seek influence in inspections, public contracts and licenses, taxes, and so on. This matches the previous literature on the transition economies, which finds a positive relationship between administrative corruption and small firm size. In contrast, we do not find significant differences between the sectors of the major lines of business in each component of corruption.

**Table 8** Firm competitiveness and corruption: logit and ordered logit regressions

	Virtually no corruption	Political corruption	Administrative corruption: (inspections)	Highest administrative corruption level (public contracts, licensing, taxes and regulations)	Ordered logit
<i>Size (ref. small)</i>					
Medium	1.124 (0.091)	0.776* (0.104)	0.922 (0.078)	1.085 (0.122)	0.957 (0.070)
Large	1.418*** (0.128)	0.926 (0.134)	0.771*** (0.076)	0.805 (0.109)	0.739*** (0.061)
<i>Sector (ref. services)</i>					
Manufacturing	1.073 (0.095)	0.924 (0.128)	1.038 (0.097)	0.846 (0.103)	0.920 (0.073)
<i>Competitiveness (ref. decreasing)</i>					
Stable	1.396*** (0.133)	0.869* (0.119)	0.786** (0.080)	0.775* (0.111)	0.737*** (0.064)
Increasing	1.077 (0.084)	0.933 (0.112)	0.909 (0.074)	1.094 (0.120)	0.965 (0.067)
<i>Country (ref. Georgia)</i>					
Yugoslavia	1.044 (0.289)	1.580 (0.640)	0.766 (0.221)	1.008 (0.307)	0.902 (0.220)
FYROM	1.128 (0.335)	1.609 (0.690)	0.471** (0.161)	1.407 (0.443)	0.976 (0.262)
Albania	0.540** (0.167)	1.100 (0.490)	2.029** (0.559)	0.723 (0.237)	1.258 (0.304)
Croatia	2.207*** (0.596)	0.578 (0.297)	0.542** (0.164)	0.819 (0.256)	0.515*** (0.130)
Turkey	1.347 (0.286)	2.308*** (0.736)	0.723 (0.157)	0.524*** (0.129)	0.576*** (0.107)

Table 8 (continued)

	Virtually no corruption	Political corruption	Administrative corruption: (inspectors)	Highest administrative corruption level (public contracts, licensing, taxes and regulations)	Ordered logit
Bosnia	0.697 (0.214)	0.579 (0.316)	0.579* (0.185)	2.708*** (0.800)	2.412*** (0.647)
Slovenia	3.088*** (0.756)	1.555 (0.572)	0.470*** (0.128)	0.230*** (0.083)	0.274*** (0.061)
Poland	2.755*** (0.591)	0.704 (0.252)	0.818 (0.179)	0.269*** (0.075)	0.343*** (0.066)
Ukraine	1.150 (0.254)	0.714 (0.262)	1.666** (0.363)	0.425*** (0.113)	0.757 (0.146)
Belarus	1.942*** (0.459)	0.636 (0.267)	1.262 (0.301)	0.224*** (0.079)	0.463*** (0.098)
Hungary	2.364*** (0.559)	0.250** (0.135)	1.045 (0.252)	0.345*** (0.108)	0.427*** (0.092)
Czech Republic	1.804** (0.421)	0.804 (0.316)	0.895 (0.215)	0.537** (0.152)	0.546*** (0.115)
Slovak Republic	1.442 (0.389)	1.298 (0.541)	0.977 (0.270)	0.512** (0.173)	0.604** (0.145)
Romania	1.203 (0.293)	0.640 (0.274)	1.474 (0.354)	0.527** (0.154)	0.776 (0.166)
Bulgaria	0.623* (0.162)	2.980*** (1.023)	0.827 (0.208)	0.965 (0.262)	1.031 (0.219)
Moldova	0.766 (0.211)	2.134** (0.798)	1.440 (0.374)	0.458** (0.152)	0.828 (0.187)
Latvia	2.176*** (0.603)	1.231 (0.535)	0.639 (0.193)	0.468** (0.169)	0.424*** (0.108)



Table 8 (continued)

	Virtually no corruption	Political corruption	Administrative corrup- tion: (inspections)	Highest administrative corruption level (public contracts, licensing, taxes and regulations)	Ordered logit
Lithuania	2.538*** (0.616)	1.349 (0.505)	0.685 (0.176)	0.235*** (0.085)	0.333*** (0.073)
Estonia	1.887** (0.518)	0.395 (0.232)	1.174 (0.328)	0.403** (0.149)	0.533** (0.134)
Armenia	5.748*** (1.585)	1.077 (0.446)	0.344*** (0.108)	0.050*** (0.037)	0.147*** (0.038)
Kazakhstan	1.940*** (0.463)	0.604 (0.259)	1.348 (0.323)	0.180*** (0.069)	0.456*** (0.097)
Azerbaijan	1.712* (0.477)	1.636 (0.676)	0.807 (0.239)	0.386** (0.152)	0.484*** (0.122)
Uzbekistan	2.819*** (0.694)	1.586 (0.588)	0.435*** (0.121)	0.418*** (0.131)	0.326*** (0.073)
Russia	1.424 (0.306)	0.490* (0.187)	1.970*** (0.420)	0.195*** (0.058)	0.608*** (0.114)
Tajikistan	0.397*** (0.121)	2.168** (0.810)	0.907 (0.244)	1.623* (0.448)	1.782** (0.413)
Kyrgyzstan	0.619* (0.179)	1.791 (0.698)	1.430 (0.379)	0.760 (0.235)	1.098 (0.255)
Constant	0.391*** (0.080)	0.124*** (0.040)	0.541*** (0.110)	0.379*** (0.087)	0.032
Pseudo R <sup>2</sup>	0.053	0.050	0.036	0.070	
LR $\chi^2$	331.513***	156.335***	206.960***	260.869***	359.596***
Log lk	- 2976	- 1481	- 2780	- 1720	- 5509
N	4590	4590	4590	4590	4590

Estimates reported as odds ratios. In brackets: standard errors of estimated parameters; asterisks: significant  $p$  value, \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . LR test statistic for country fixed effects shown at bottom of table

Third, we provide support for the findings of the country classification on corruption. We estimate that Turkey and Bulgaria have the highest propensity for political corruption and significance in the estimated country-dummy parameters of administrative corruption, which accurately replicates our previous estimates.

The last column of Table 8 lists the estimates of the ordered logit regression parameters. The odds of reporting virtually no corruption versus political corruption, and the two LCs of administrative corruption with monotonic increase of intensity are 25–30% lower in firms with stable competitiveness, net of the characteristics of firms (size and sector) and of the specificities of countries. This confirms the findings of the logistic regressions.

## 6 Conclusions

This paper proposes a LC model to identify different dimensions of corruption, which is a latent phenomenon that is often measured by synthesis of observed indices assumed to be correlated with corruption. In particular, the constrained version of the M-IRT model is used to classify corruption using firms' perceptions in a sample of Eastern Europe and Central Asian countries.

The hypothesis that items in a multidimensional perspective can reproduce the distinction between administrative and political corruption is extended and we find that four LCs exist, with increasing levels of corruption. We identify these LCs and validate the results according to a corruption index that is directly derived from the BEEPS and CPI rankings of TI. An illustrative example of the relationship between firm competitiveness and the identified LCs shows that weakly competitive firms are more tempted to resort to bribery, influencing the rules of competition, and this propensity is estimated to be greater in small firms associated with pervasive administrative corruption.

We believe that the strategy that we have used to identify corruption would be useful to empirical researchers in several respects. First, the greater part of the data on corruption uses direct investigation by questionnaires, such as the M-IRT models, which allow us to disentangle the different dimensions that an aggregate indicator of corruption tends to hide. Second, the M-IRT model reduces the bias associated with the measures of impact of a specific relationship, when corruption has been incorrectly classified. From the M-IRT model that is adopted here, the contrasting evidence of the determinants of corruption may be explained by overlapping their heterogeneous effects. This implies that the implementation of anti-corruption policies may also benefit from a correct identification of which type of corruption affects countries or regions, and which tools should be used to combat it.

## Appendix A: Model Selection

The choice of the number of classes in the multidimensional LC Rash model is crucial for model identification. Although selection is mainly based on information criteria, we complement it with indices measuring goodness-of-fit. The most frequently used index is the Bayesian information criterion (BIC; Schwarz 1978), which is based on the index:

$$BIC = -2\hat{\ell} + m \log(n), \quad (10)$$

**Table 9** Selection of number of LCs

Log-likelihood statistics					
	Log-likelihood (log LL)	BIC (log LL)	AIC (log LL)	AIC3 (log LL)	CAIC (log LL)
Class1	- 46461.6580	93058.2930	92955.3173	92971.3173	93074.2930
Class2	- 40417.9682	81114.3238	80901.9364	80934.9364	81147.3238
Class3	- 38985.0967	78391.9925	78070.1933	78120.1933	78441.9925
Class4	- 38266.8741	77242.3707	76701.7482	76785.7482	77326.3707
Class5	- 38204.7666	77261.5674	76811.5331	76812.5331	77362.5674
Class6	- 38404.3510	77373.9129	76942.7020	77009.7020	77440.9129
Class7	- 38171.2934	77338.0329	76578.5869	76696.5869	77456.0329
Class8	- 38148.8201	77436.4980	76567.6402	76702.6402	77571.4980

Goodness of LC model fit					
	E	$R^2_{entropy}$	Standard $R^2_{variance}$	AWE	
				$I^c$	Index
Class1	0.0000	10.0000	10.0000	- 46461.6586	93241.2687
Class2	0.0374	0.8712	0.8922	- 40829.3882	82314.5512
Class3	0.0532	0.8630	0.8754	- 39674.2046	80342.0076
Class4	0.0613	0.8593	0.8633	- 39214.4170	79760.2558
Class5	0.1283	0.7791	0.7547	- 39787.7568	81244.7587
Class6	0.1432	0.7623	0.7295	- 39945.4201	81897.9089
Class7	0.1845	0.7214	0.6728	- 40429.0881	83203.0681
Class8	0.1974	0.7070	0.6541	- 40630.1816	83943.0787

The upper part of the table lists the log-likelihood statistics and estimates the log-likelihood function. The proposed log-likelihood statistics are: Bayesian Information Criterion (BIC), Akaike information criterion (AIC), Akaike information criterion with 3 as penalising factor (AIC3), and Consistent Akaike information criterion (CAIC). Bottom part of table lists several measures and statistics based on ability of model to fit data, and include: estimated proportion of classification errors (E), two pseudo R-square measures ( $R^2_{entropy}$ ) and ( $R^2_{variance}$ ), and Average Weight of Evidence (AWE)

where  $\hat{\ell}$  is the maximum value of log-likelihood test statistics and  $m$  is the number of free parameters. The Akaike information criterion (AIC) is another widely used criterion, (Akaike 1973), which is based on the index:

$$AIC = -2\hat{\ell} + 2 \log(n). \tag{11}$$

Two extended versions of the AIC index were proposed by Andrews and Currim (2003), and they include different weights to estimate the log-likelihood function. In the first case, AIC is penalised with a factor of 3 instead of 2 (AIC3); various penalisation terms are included in AIC to define a CAIC.

We also used measures based on the capacity of the model to fit the data. These measures were based on  $R^2_{entropy}$  and variance  $R^2_{variance}$  (Magidson and Vermunt 2001), the estimated proportion of classification errors (E),<sup>12</sup> and the Average Weight of Evidence

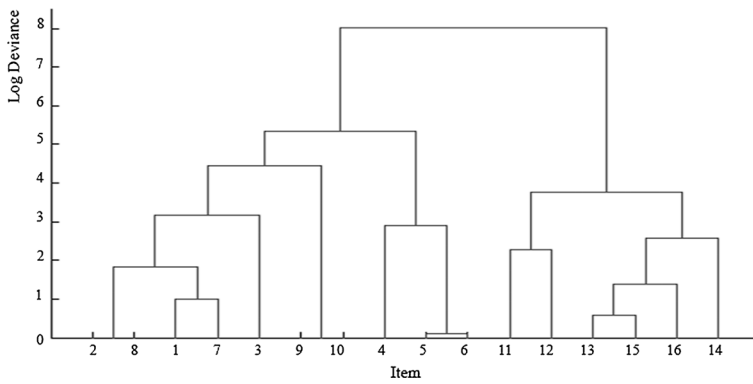
<sup>12</sup> Unlike the other measures, E proves model accuracy when values are close to zero.

(AWE). In particular, the AWE is built by adding a third dimension to the BIC index, which considers the performance of individual classifications within groups, as argued in Banfield and Raftery (1993). The formal specification is:

$$AWE = -2\hat{\ell}^c + 2m \left[ \frac{2}{3} + \log(n) \right] \quad (12)$$

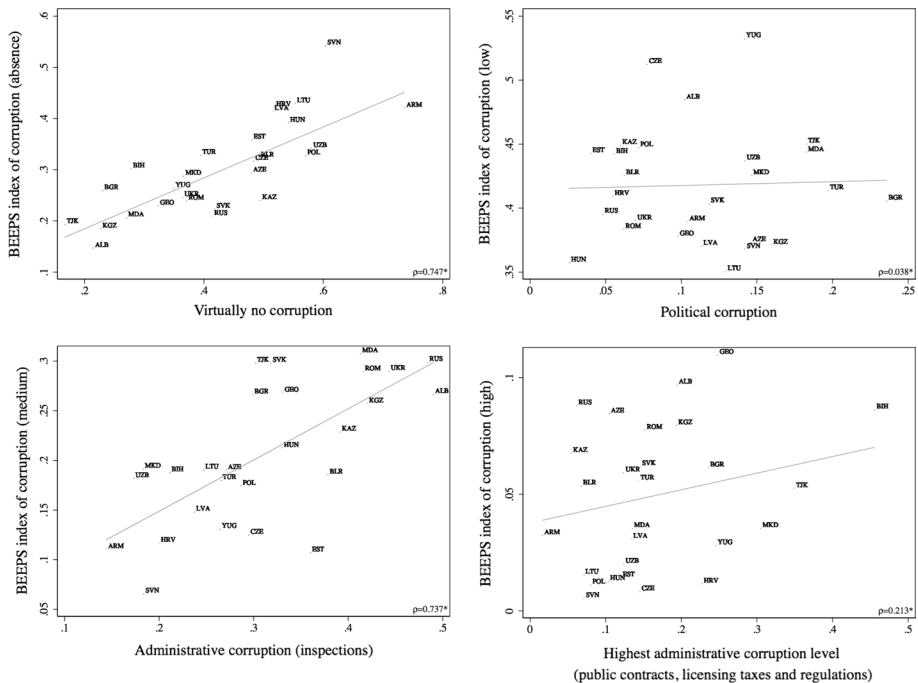
where  $\hat{\ell}^c$  is the log-likelihood classification (Biernacki and Govaert 1998). Table 6 lists the log-likelihood and classification statistics for a number of predetermined LCs, from 2 to 8 (Table 9).

## Appendix B: Dendrogram



*Notes* Classifications and definitions of items are listed in Table 3.

## Appendix C: Country Correlations Between Identified Latent Classes and Direct BEEPS Measures of Administrative Corruption



## References

- Ades, A., & Di Tella, R. (1997). The new economics of corruption: A survey and some new results. *Political Studies*, 45(3), 496–515.
- Ades, A., & Di Tella, R. (1999). Rents, competition, and corruption. *American Economic Review*, 89(4), 982–993.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *2nd international symposium on information theory* (pp. 267–281). Budapest: Akademiai Ki-ado.
- Anderson, J. H., & Gray, C. W. (2006). *Anticorruption in transition 3: Who is succeeding... and Why?*. Washington DC: World Bank.
- Andrews, R. L., & Currim, I. S. (2003). A comparison of segment retention criteria for finite mixture logit models. *Journal of Marketing Research*, 40, 235–243.
- Banfield, J. D., & Raftery, A. E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics*, 49(3), 803–821.
- Bardhan, P. (1997). Corruption and development: A review of issues. *Journal of Economic Literature*, 35(3), 1320–1346.
- Bardhan, P. (2006). The economist's approach to the problem of corruption. *World Development*, 34(2), 341–348.
- Bartolucci, F. (2007). A class of multidimensional IRT models for testing unidimensionality and clustering items. *Psychometrika*, 72(2), 141–157.
- Bartolucci, F., d'Agostino, G., & Montanari, G. E. (2010). *An investigation of the discriminant power and dimensionality of items used for assessing health condition of elderly people*. Technical Report. [arXiv:1008.3268v1](https://arxiv.org/abs/1008.3268v1) [stat.AP].
- Bartolucci, F., Montanari, G. E., & Pandolfi, S. (2012). *Item selection by an extended latent class model: An application to nursing homes evaluation*. SSRN: Technical Report.
- Biernacki, C., & Govaert, G. (1998). *Choosing models in model-based clustering and discriminant analysis*. Institut National de Recherche en Informatique et en Automatique: Technical Report.

- Bliss, C., & Di Tella, R. (1997). Does competition kill corruption? *Journal of Political Economy*, 105(5), 1001–1023.
- Bock, R. (1997). A brief history of item response theory. *Educational Measurement: Issues and Practice*, 16, 21–23.
- Cappellari, L., & Jenkins, S. P. (2006). Summarising multiple deprivation indicators. *ISER Working Paper Series 2006-40*. Institute for Social and Economic Research.
- Clarke, G. R. G., & Xu, L. C. (2002). Ownership, competition, and corruption: Bribe takers versus bribe payers. *Policy Research Working Paper Series 2783*. The World Bank.
- d'Agostino, G., Dunne, J. P., & Pieroni, L. (2016a). Government spending. *Corruption and Economic Growth, World Development*, 84, 190–205.
- d'Agostino, G., Dunne, J. P., & Pieroni, L. (2016b). Corruption and growth in Africa. *European Journal of Political Economy*, 43, 71–88.
- Delavallade, C. (2011). What drives corruption? Evidence from north African firms. *Working Papers 244*. Economic Research Southern Africa.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), 1–38.
- Dimitrova-Grajzl, V., Grajzl, P., & Guse, A. J. (2012). Trust, perceptions of corruption, and demand for regulation: Evidence from post-socialist countries. *Journal of Behavioral and Experimental Economics (formerly The Journal of Socio-Economics)*, 41(3), 292–303.
- Dreher, A., Kotsogiannis, C., & McCorrison, S. (2007). Corruption around the world: Evidence from a structural model. *Journal of Comparative Economics*, 35(3), 443–466.
- Faye, O., de Laat, J., & Zulu, E. (2009). *Poverty dynamics and mobility in Nairobi's informal settlements*. APHRC: Technical Report.
- Ferraz, C., & Finan, F. (2008). Exposing corrupt politicians: The effects of Brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, 123(2), 703–745.
- Gauthier, B., & Reinkka, R. (2001). Shifting tax burdens through exemptions and evasion—An empirical investigation of Uganda. *Policy Research Working Paper Series 2735*. The World Bank.
- Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61(2), 215–231.
- Goodman, L. A. (1978). *Analysing qualitative/categorical data log-linear models and latent-structure analysis*. Boston: Addison-Wesley Pub. Co.
- Gorodnichenko, Y., & Sabirianova-Peter, K. (2007). Public sector pay and corruption: Measuring bribery from micro data. *Journal of Public Economics*, 91(5–6), 963–991.
- Hellman, J. S., Jones G., & Kaufmann, D. (2000). Seize the state, seize the day: State capture, corruption and influence in transition. Technical Report 2444. *World Bank Policy Research Working Paper*.
- Hellman, J. S., Jones, G., & Kaufmann, D. (2003). Seize the state, seize the day: State capture and influence in transition economies. *Journal of Comparative Economics*, 31(4), 751–773.
- Herzfeld, T., & Weiss, C. (2003). Corruption and legal (in)effectiveness: An empirical investigation. *European Journal of Political Economy*, 19(3), 621–632.
- Johnson, S., Kaufmann, D., & Zoido-Lobaton, P. (1999). Corruption, public finances, and the unofficial economy. *Policy Research Working Paper Series 2169*. The World Bank.
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2007). Measuring corruption : Myths and realities. In *Africa region findings & Good practice infobriefs* (no. 273). Washington, DC: World Bank.
- Knack, S., & Keefer, P. (1995). Institutions and economic performance: Cross-country tests using alternative measures. *Economics and Politics*, 7, 207–227.
- Knack, S. (2007). Measuring corruption: A critique of indicators in Eastern Europe and Central Asia. *Journal of Public Policy*, 27(3), 255–291.
- Kuklys, W. (2004). Measuring standard of living in the UK—An application of sen's functioning approach using structural equation models. *Papers on strategic interaction 2004–11*. Max Planck Institute of Economics, Strategic Interaction Group.
- Lambdsdorff, J. G. (2007). *The institutional economics of corruption and reform: Theory, evidence and policy*. Cambridge: Cambridge University Press.
- Mauro, P. (1995). Corruption and growth. *Quarterly Journal of Economics*, 110(3), 681–712.
- Magidson, J., & Vermunt, J. K. (2001). Latent class factor and cluster models, bi-plots, and related graphical displays. *Sociological Methodology*, 31, 223–264.
- Mocan, N. (2008). What determines corruption? International evidence from microdata. *Economic Inquiry*, 46(4), 493–510.
- Nowak, R. (2001). Corruption and transition economies. Technical Report. *Presented to the preparatory seminar for the 9th OCSE economic forum*. Bucharest

- Olken, B. A. (2009). Corruption perceptions vs. corruption reality. *Journal of Public Economics*, 93(7–8), 950–964.
- Pieroni, L., & d'Agostino, G. (2013). Corruption and the effects of economic freedom. *European Journal of Political Economy*, 29(C), 54–72.
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. *Berkeley Symposium on Mathematical Statistics and Probability*, 4, 321–333.
- Reinikka, R., & Svensson, J. (2006). Using micro-surveys to measure and explain corruption. *World Development*, 34(2), 359–370.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464.
- Scott, J. C. (1972). *Comparative political corruption*. Englewood Cliffs, NJ: Prentice-Hall.
- Seligson, M. A. (2006). The measurement and impact of corruption victimization: Survey evidence from Latin America. *World Development*, 34(2), 381–404.
- Svensson, J. (2003). Who must pay bribes and how much? Evidence from a cross section of firms. *The Quarterly Journal of Economics*, 118(1), 207–230.
- Svensson, J. (2005). Eight questions about corruption. *Journal of Economic Perspectives*, 19(3), 19–42.
- Tanzi, V. (1998). Corruption around the world: Causes, consequences, scope, and cures. *IMF Working Papers* 98/63. International Monetary Fund.
- Warren, E. M. (2004). What does corruption mean in a democracy? *American Journal of Political Science*, 48(2), 328–343.