

Kaufmann, Daniel; Mehrez, Gil; Gurgur, Tugrul

## Article

# Voice or public sector management? An empirical investigation of determinants of public sector performance based on a survey of public officials

Journal of Applied Economics

## Provided in Cooperation with:

University of CEMA, Buenos Aires

*Suggested Citation:* Kaufmann, Daniel; Mehrez, Gil; Gurgur, Tugrul (2019) : Voice or public sector management? An empirical investigation of determinants of public sector performance based on a survey of public officials, Journal of Applied Economics, ISSN 1667-6726, Taylor & Francis, Abingdon, Vol. 22, Iss. 1, pp. 321-348,  
<https://doi.org/10.1080/15140326.2019.1627718>

This Version is available at:

<https://hdl.handle.net/10419/314064>

## Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

## Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



## Voice or public sector management? An empirical investigation of determinants of public sector performance based on a survey of public officials

Daniel Kaufmann, Gil Mehrez & Tugrul Gurgur

**To cite this article:** Daniel Kaufmann, Gil Mehrez & Tugrul Gurgur (2019) Voice or public sector management? An empirical investigation of determinants of public sector performance based on a survey of public officials, *Journal of Applied Economics*, 22:1, 321-348, DOI: [10.1080/15140326.2019.1627718](https://doi.org/10.1080/15140326.2019.1627718)

**To link to this article:** <https://doi.org/10.1080/15140326.2019.1627718>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 13 Jun 2019.



[Submit your article to this journal](#)



Article views: 4278



[View related articles](#)



[View Crossmark data](#)



Citing articles: 15 [View citing articles](#)

ARTICLE



# Voice or public sector management? An empirical investigation of determinants of public sector performance based on a survey of public officials

Daniel Kaufmann<sup>a</sup>, Gil Mehrez<sup>b</sup> and Tugrul Gurgur<sup>c</sup>

<sup>a</sup>The Natural Resource Governance Institute, NY, USA; <sup>b</sup>IMF, Washington, DC, USA; <sup>c</sup>Tanyeli Research and Consulting, Ankara, Turkey

## ABSTRACT

Drawing on an in-depth governance micro-survey of public officials in Bolivia, we address empirically the question of the relative importance of the various determinants of governance. We find that commonly made inferences about policy based on simple correlation can be highly misleading due to the high correlation between the various governance determinants, as well as the endogeneity in these variables. We find that undue emphasis may have been given in the previous work to a number of conventional public-sector management variables (such as civil servant wages, internal enforcement of rules, autonomy of agency by fiat, etc.), while undermining the priority of “external” (to public sector management) variables, such as citizen voice and transparency. The latter set of “voice”-related variables has a larger effect on the service delivery performance and corruption than the more traditional public-sector management type of variables.

## ARTICLE HISTORY

Received 27 July 2015  
Accepted 23 April 2019

## KEYWORDS

Governance; transparency;  
public sector management;  
micro-survey; Bolivia

## JEL

C36; D73

## 1. Introduction

Efficient governance has been long recognized as an important element in improving welfare and economic growth. Yet many countries are plagued with incompetent bureaucracy, mismanagement, and corruption. Recently, a number of these countries have undertaken policies to fight corruption and improve their governance. While the empirical literature is growing, theoretical and empirical findings are rather limited to guide these policies. A particular gap is in country-relevant empirical research, where findings and policy recommendations are adapted to a particular country setting. Much of the empirical work is through cross-country regressions, and thus with limited policy applicability. Even when country-specific data are available, inferences are often made based on simple correlation between variables. As a result, reviewing the writings in the literature one encounters a quite diverse and comprehensive list of factors related to good governance, such as public sector management and compensation, voice and accountability, meritocracy, decentralization and the like. The cumulative effect of such studies of different aspects of governance that matter often sheds no light on

their relative influence. This has tended to result in overly long (“Christmas tree-like”) lists of recommendations without providing any priorities.

Additional complications may arise because several policy variables may be endogenous. For example, on the one hand, poor transparency may lead to a higher level of corruption as public officials would be able to hide their illegal activities. On the other hand, corruption may lead to poorer transparency, as corrupt agents, reluctant to be exposed, attempt to weaken information flow. In this case, even partial correlations based on single equation regression analysis (which controls for other factors) may be misleading.

Drawing on an in-depth micro-survey in Bolivia, this paper addresses the question of the relative importance of the various determinants of governance. We, first, develop a simple theoretical model, where corruption, transparency and the level/quality of public services (“performance”) are the three key decision variables in the hands of a public official. Then, using a unique data set from detailed responses of public officials working in various public sector agencies and municipalities in Bolivia, we estimate a system of equations built upon our theoretical work. Based on these results we try to shed light on priorities to improve the quality of governance and public services.

In general, policies to improve governance can be divided into two groups: policies that emphasize the importance of “*citizen power*” (such as voice, transparency, and lack of politicization) and policies that emphasize the importance of “*the structure of institutions*” (such as formal accountability mechanisms, autonomy by fiat, rules, and government pay). The data set used in this paper allows a comparison of these policies after controlling for a large set of relevant factors. We find that the citizen power-related variables have a larger effect on performance and corruption than the structure of institutions. Thus, it may be more effective to emphasize reforms focused on voice and transparency rather than building formal norms, rules, and regulations to strengthen the quality of governance, to reduce corruption, and to improve public services.

In the next section, we discuss the literature. We present a simple model in section 3. Section 4 explains the data and econometric estimations. Section 5 concludes.

## 2. Theoretical considerations

There is a growing literature on corruption that involves both theoretical and empirical studies. Treisman (2007) and Dimant and Tosato (2017) provide a comprehensive review of this literature. Many factors have been discussed and empirically investigated as the determinants of performance and corruption in the public sector, including government pay, meritocracy, decentralization, press freedom, citizen voice, transparency, etc. We briefly discuss this literature and the main findings below.

Economic analyses of the determinants of corruption follow Becker’s seminal model of crime and punishment (Becker, 1968) and the principal-agent theory (Becker & Stigler, 1974). Becker (1968) argued that a person chooses illegal activity if the expected payoff from illegal activity is greater than the expected payoff from legal activities. Policies to fight corruption, therefore, should focus on increasing rewards (wage) on the one hand and strengthening enforcement (punishment and monitoring) on the other hand. Becker and Stigler (1974) extended this analysis to a principal-agent setting,

where because of incomplete information the higher-level manager or the voter (the “principal”) cannot observe the actions of the public official (the “agent”). In this model, information asymmetries enable the corrupt public official to hide his/her corrupt activities. Thus, more transparency may reduce corruption by eliminating information problems.

Implications of the aforementioned models have been debated extensively in the literature. Many scholars have argued that raising government pay may reduce corruption (see, for instance, Mookherjee & Png, 1995). However, the required pay raise may be prohibitively expensive, and it may be cost-effective for governments to pay “capitulation wages” (i.e., wages that attract only the dishonest) rather than to raise wages to the levels required to deter corruption (Besley & McLaren, 1993). Moreover, pay hikes may not be effective as a policy tool if control mechanisms are too weak or too strong (Di Tella & Schargrodsky, 2003) or there is a substantial difference between the actual wage and the “fair” wage as perceived by the agent (Van Rijckeghem & Weder, 2001).

Enforcement is an important element of Becker’s model. In the absence of credible enforcement mechanisms, public officials pay no heed to rules and guidelines that regulate their actions. Lack of accountability fosters excessive discretion, which creates opportunities for corruption. Moreover, when jobs and promotions in the public service are based on political connections, public officials would have less incentive to stay “clean” because their career prospects are linked to “whom they serve” rather than to their performance (Evans & Rauch, 2000). Thus, a bureaucracy that offers long-term careers with chances of advancement based on merit would result in less corruption. Politicization of institutions, on the other hand, causes economic distortion and corruption through inflated budgets and high government pay (Menes, 1999).

Public opinion is a crucial part of enforcement. Putnam (1993), LaPorta, Silanes, Shleifer, and Vishny (1999), Brunetti and Weder (2003), and Islam (2006) highlighted press freedom, civic engagement, and presence of well-informed electorate in a democratic setting as a significant deterrent of corruption. In this context, the role of increased access to information and transparency in public administration may play a significant role in curbing corruption by raising the probability of getting caught. Transparency also reduces information asymmetry, thereby reducing the size and availability of rent-seeking opportunities (Kolstad & Wiig, 2009; Yamamura & Kondoh, 2013). Gray-Molina, de Rada, and Yáñez (1999) found that over-pricing and informal payments to municipal health service providers in Bolivia decline significantly in places where citizens participate in health board meetings. Furthermore, they showed that formal control and supervision mechanisms have no significant effect on corruption. Reinikka and Svensson (2005), using a micro survey in Uganda, showed access to information may limit corruption in public schools. Del Monte and Papagni (2007) and Gurgur (2016) found that citizen voice (captured by participation to voluntary organizations) reduces corruption in Italy and the Philippines, respectively. However, transparency is not a panacea. Unless citizens’ capacity to act upon available information is not strengthened (through, for example, media circulation, free and fair elections, and an independent and effective judiciary) the negative link between transparency and corruption may disappear (Escaleras, Lin, & Register, 2010; Lindstedt & Naurin, 2010).

Decentralization has also been analyzed in the context of good governance. Decentralization may improve the quality of government by bringing public officials

closer to the people. People can compare the performances of local governments and if they are dissatisfied, they can vote the incumbent out of power (Porto & Porto, 2000). Alternatively, they can “vote by their feet” or by “their pocket” (Tiebout, 1956). Other scholars, however, argued that corruption is more prevalent at the local level due to high patronage politics and clientelism emerged from closer interaction between local governments and citizens (Tanzi, 1993). Moreover, inefficiencies due to vertical and horizontal coordination problems, lack of qualified employees, and effective monitoring mechanisms may increase corruption at the local level (Prud’homme, 1994). Empirical research on the premises of decentralization is inconclusive. Fisman and Gatti (2002), Ferraz and Finan (2011), and Gurgur and Shah (2015) reported that decentralization is associated with lower corruption. Treisman (2007), Fan, Lin, and Treisman (2009), and Goel and Nelson (2011) found that corruption is higher in federally structured states, in countries that have more tiers of government or more financial decentralization, especially if there exists ethnic or economic heterogeneity within society.

This paper offers several contributions to the governance literature. First, we develop a model where the performance of public agencies, corruption, and transparency are the three key decision variables in the hands of a public official. The public official decides on their levels to maximize her welfare subject to several constraints. First, a higher level of corruption leads to the loss of bargaining power with the private agent and reputational costs. Second, a lower level of transparency is subject to internal rule-based constraints. Third, a lower level of performance is subject to lower earnings and promotion prospects. This joint determination of the three key decision variables by the public official extends the literature by highlighting the importance of incentive structure along these dimensions in determining the level/quality of public services. In particular, the link between corruption and performance may be either a positive or a negative. This contrasts much of the literature which: either (i) did not model such link with public sector performance at all, or (ii) modeled it as a simple and pre-determined direction where corruption is either “grease” helping businesses by getting around red tape (Lui, 1985), or “sand” exacerbating the red-tape (Kaufmann & Wei, 2000), or (iii) modeled “public service” rather superficially (such as red tape or licenses), rather than actual public services. Hence, our model is a more integrative model of governance and public services.

Second, in contrast to other models where the levels of corruption, transparency, and performance are determined taking the levels of the other two variables as given, our model considers all these variables as endogenous to be determined simultaneously. Measures to improve public services and governance (as well as empirical studies), therefore, should utilize this complex simultaneous system to devise efficient policies.

Third, our model formally incorporates the institutional norms and emphasizes the positive (negative) externalities of having honest (corrupt) public officials. Deviations from the institutional norms are subject to penalty. This highlights the importance of culture within institutions. An individual who works in an institution with high performance or well-established guidelines is more likely to provide better performance than an individual who works in an institution that lacks both. This is “the egg and the chicken” dilemma faced by policymakers: If one could improve the governance or performance within an institution then each public official would improve her

individual performance, but until all or a significant portion of public officials do so the institutional average would not rise.

### 3. A simple model of public sector performance, corruption, and transparency

Consider an economy that is composed of an institution (“principal”), a representative public official (“agent”) and  $n$  heterogeneous households. The utility of household  $i$  depends on two elements: her preference (in terms of quality or quantity) over public services,  $X_i$ , and the service she actually receives from the public official,  $Z_i$ . For example,  $X_i$  could be thought of as the preferred level for water or sanitation services, public roads, building permits and so on, and  $Z_i$  as the level provided by the public official. Specifically, let the utility function of household  $i$  be:

$$U_i = -(X_i - Z_i)^2. \quad (1)$$

The household attains maximum utility when the level of service she receives is equal to her preference and declines the larger the difference between the service she receives and her preference. Note that the optimal level of service from the household’s point of view is not infinity, but rather a finite amount. For example, building too many roads, telephone poles, or collecting garbage 10 times a day may reduce the welfare.

We assume that the institution’s utility (which we consider as the social welfare) is different from the sum of the utilities of households due to externalities or limited resources. For example, construction permits may have negative externalities such as traffic congestion or pollution, or the agency has enough resources to collect garbage only once a month. Specifically, let the institution’s utility be:

$$SW = -\sum_{i=1}^n (\beta X_i - Z_i)^2, \quad (2)$$

where  $\beta$  captures the difference between the social welfare and the individual welfare. If  $\beta < 1$ , then the public service involves negative externalities (or resources are limited). If  $\beta > 1$ , then the public service involves positive externalities. To simplify the exposition, we present the solution for the case where  $\beta < 1$ , which we believe is more realistic. However, one could solve the model for  $\beta > 1$  without loss of generality as well.

The law requires the institution to follow the same policy for everybody, i.e., it cannot issue a policy that links  $Z_i$  to  $X_i$ . For example, the institution cannot collect garbage only from some houses or requires permits/regulations only from some households. This level,  $Z^*$ , maximizes the social welfare and can be calculated from equation (2) as:

$$Z^* = \beta \bar{X}, \quad (3)$$

where  $\bar{X} = \sum_i X_i / n$  is the level of service preferred by households on the average. Note that since  $\beta < 1$ , equation (3) implies that the public service is provided at a level that is



less than the average of public preferences due to negative externalities or lack of enough resources.

The public official in charge of supplying public services, however, has a different objective. By differentiating the level of service given to some households, the public official seeks to maximize her expected welfare. In particular, she provides  $Z^p$  (the standard level of service) to any household that does not bribe her and provides  $Z_i$  (the customized level of service) to household  $i$  who pays a bribe,  $b_i$ . We define  $b_i$  as a function of the difference between the utility of the standard service,  $(X_i - Z^p)^2$  and the utility of the customized service,  $(X_i - Z_i)^2$ :

$$b_i = \frac{1}{2} (X_i - Z^p)^2 - \frac{1}{2} (X_i - Z_i)^2 \quad (4)$$

We assume that the value of the bribe to the public official, or alternatively the amount of bribe that the public official receives to deviate from  $Z^p$  to  $Z_i$  declines with the level of transparency (i.e., information available to the public). Intuitively, a higher level of transparency reduces the bargaining power of the public official and hence she can capture a smaller share from the households. Alternatively, a higher level of transparency increases the probability of being caught and hence reduces the value of the bribe. To capture this effect, we assume that the value of bribe to the public official,  $g(T)$ , is discounted according to the level of transparency,  $T$ , where  $g'(T) < 0$ . Alternatively,  $g(T)$  can be seen as the rent paid to the official and  $(1 - g(T))$  is the rent captured by the household. One could consider also other forms of transparency, such as transparency about the deviation of the standard level from the norm or transparency about the variation of service levels among households. We leave these forms of transparency for future research.

We assume that the level of transparency is controlled by the public official. In the absence of any restrictions set by the institution, the public official can maximize the bribe she receives by setting the standard service and transparency at their lowest levels. The institution, however, provides incentives to prevent such behavior. Accordingly, the public official's expected utility depends on three elements. First, it depends on the difference between the socially optimal level of service,  $Z^*$ , and the standard level,  $Z^p$ . Every organization provides some rules and policies to guide the decision-making process in the administration. Officials who deviate from the institutional norm have lower (higher) probability of promotion (demotion). Second, it depends on the difference between the customized level of service,  $Z_i$ , that household  $i$  gets and the standard level,  $Z^p$ . The more the public official deviates from the standard practices and does favor some households the larger the risk of being caught and punished. Third, it depends on the difference between transparency level,  $T$ , set by the public official and the transparency standards,  $T^*$ , set by the institution. The public official can reduce the transparency level by hiding records, withholding information, not reporting transactions, etc. However, the larger the discrepancies, the lower (higher) the probability of promotion (demotion).

Given this set of incentives the public official maximizes her welfare by choosing the standard level of service,  $Z^p$ , the level of transparency,  $T$ , and the customized level of service,  $Z_i$ , to household  $i$ . Specifically, the public official solves the following optimization problem:



$$\begin{aligned}
\max_{Z^p, Z_1, \dots, Z_n, T} V &= g(T) \sum_i b_i - \frac{1}{2} f_1 \sum_i (Z_i - Z^p)^2 - \frac{1}{2} f_2 (Z^* - Z^p)^2 \\
&\quad - \frac{1}{2} f_3 (T^* - T)^2 \\
st. b_i &= \frac{1}{2} (X_i - Z^p)^2 - \frac{1}{2} (X_i - Z_i)^2
\end{aligned} \tag{5}$$

where  $f_1, f_2, f_3 > 0$ . The first part of the utility function is the value of the total bribe. The remaining terms represent the incentives, which consist of three parts: (i) the difference between the customized level of service and the standard level of service,  $(Z_i - Z^p)$ , (ii) the difference between the norm set by the institution and the standard level of service provided by the public official,  $(Z^* - Z^p)$ , and (iii) the difference between the norm set by the institution and the level of transparency set by the public official,  $(T^* - T)$ . The parameters  $f_1, f_2$ , and  $f_3$  correspond to the importance of each part, respectively. The incentive structures (i.e.  $f_1, f_2$ , and  $f_3$  times the respective element) represent the loss in the expected lifetime earnings given a set of action (i.e.  $Z^p, Z_i$ , and  $T$ ). The fact that promotion prospects depend differently on various types of deviations from institutional norms is captured by the distinct  $f$ s.

The optimal levels (from the public official point of view) of  $Z^p, Z_i$ , and  $T$  are given by the first-order conditions of equation (5) provided that  $f_1$  and  $f_2$  are sufficiently large to satisfy the second-order conditions<sup>1</sup>

$$-g(T) \sum_i (X_i - Z^p) + f_1 \sum_i (Z_i - Z^p) + f_2 (Z^* - Z^p) = 0 \tag{6}$$

$$g(T)(X_i - Z_i) - f_1 (Z_i - Z^p) = 0 \tag{7}$$

$$g'(T) \sum_i b_i + f_3 (T^* - T) = 0. \tag{8}$$

Equation (6) states that at the optimal level of  $Z^p$  the increase in bribes from decreasing  $Z^p$  by one unit (the first term) is equal to the costs from such an increase (i.e., the increase in the variability and the deviation from the institutional norm, the second and the third terms, respectively). Equation (7) states that at the optimal level of  $Z_i$  the increase in bribes from providing another unit to household  $i$  (the first term) is equal to the costs of deviating from the standard level. Finally, equation (8) states that at the optimal level of  $T$  the increase in the value of bribes due to lower transparency (the first term) is equal to the costs from decreasing transparency (the second term).

Equations (6) through (8) can further be solved for the optimal levels of  $Z^p, Z_i$ , and  $T$ :

$$Z^p = Z^* - \frac{ng(T)^2}{f_2(g(T) + f_1) - ng(T)^2} (\bar{X} - Z^*) \tag{9}$$

<sup>1</sup>These second-order conditions are (i)  $ng(T) - f_1 n - f_2 < 0$ , and (ii)  $ng^2(T) - f_1 f_2 - f_2 g(T) < 0$ . When the second-order conditions do not hold, the solution is to set the standard level at the boundary, which is zero as the decision variables are non-negative.

$$Z_i = \frac{g(T)}{g(T) + f_1} X_i + \frac{f_1}{g(T) + f_1} Z^p \quad (10)$$

$$T = T^* + \frac{g'(T)}{f_3} \sum_i b_i \quad (11)$$

The standard level of service,  $Z^p$ , is given in equation (9). Note that the second-order condition ensures that the denominator is positive. Since  $\bar{X} > Z^*$  (due to  $\beta < 1$ , which implies that there are negative externalities in the provision of public services or the institution has limited resources), the public official sets  $Z^p$  below the norm,  $Z^*$ . Intuitively, the public official by reducing the standard level increases her leverage in providing customized services and hence her bribe receipts. Higher transparency,  $T$ , reduces the benefits from bribe and hence reduces the incentives to lower  $Z^p$ . Likewise, raising the cost of deviating from the norm,  $f_2$ , or the cost of deviating from the standard level,  $f_1$ , increases  $Z^p$ . The larger these costs, the closer  $Z^p$  to the norm,  $Z^*$ .

Equation (10) describes the customized service level,  $Z_i$ , provided to household  $i$ . This level is a weighted average of the preferred service level and the standard service level, where the weights are given by  $g(T)$  and  $f_1$ , respectively. The lower the transparency (higher  $g(T)$ ) or the smaller the costs of deviating from the norm,  $f_1$ , the closer  $Z_i$  to the household's preferences,  $X_i$ . In other words, when transparency is poor or the cost of providing different service levels is small, households can pay a bribe and get the service they wish. Transparency, on the other hand, creates a discrepancy between the nominal value of the bribe and its effective value to the public official and hence reduces the amount of favoritism that an individual can buy.

Finally, the level of transparency,  $T$ , is given in equation (11).  $T$  depends positively on the norm level,  $T^*$ , and the cost of deviating from this level,  $f_3$ , and it depends negatively on the amount of bribes,  $\sum b_i$ , and marginal change in the value of the bribe with respect to  $T$ ,  $g'(T)$ . Intuitively, the public official is forced to raise the transparency level to the institutional norm, if the cost of deviation is high or the payoff from favoritism is low, or the bargaining power of the household is high.

The amount of bribe,  $b_i$ , that household  $i$  would pay can be calculated by substituting the value of  $Z_i$  from equation (10) in equation (4):

$$b_i = \frac{1}{2} \left( 1 - \frac{f_1^2}{(g(T) + f_1)^2} \right) (X_i - Z^p)^2. \quad (12)$$

Equation (12) states that  $b_i$  depends on the difference between the preferred service level of household  $i$  and the standard level. Households who prefer more public services,  $X_i$ , would pay more bribe while households whose preferences are equal to the standard level would not pay any bribe. Likewise, an increase in the punishment to favoritism,  $f_1$ , or the transparency level,  $T$ , would reduce the amount of bribe that household  $i$  pays.

The total amount of bribe can be calculated simply by aggregating over all individuals:

$$\sum_i b_i = \frac{1}{2} \left( 1 - \frac{f_1^2}{(g(T) + f_1)^2} \right) \sum_i (X_i - Z^P)^2. \quad (13)$$

It is important to note that corrupt behavior itself does not impose a social cost; bribe payments can be regarded as simply a transfer from households to the public official. In this model, the social loss from corruption arises from the deviation of the standard service level,  $Z^P$ , from the socially optimal level,  $Z^*$ . In other words, the public official sets a lower level of service in order to increase the demand for her services.

As seen from equation (11) transparency does not depend directly on the costs of preferential treatment,  $f_1$ , or the cost of deviating from the institutional norm,  $f_2$ . These costs affect the level of transparency through their effect on corruption, but do not affect transparency once the effect of corruption is controlled. Similarly, the level of corruption in equation (13) does not depend on the cost of deviating from the transparency standards,  $f_3$ , nor on the standard level of transparency,  $T^*$ , or on the change in the value of bribe when transparency changes,  $g'(T)$ . Thus, although performance, corruption, and transparency are endogenous variables that are determined simultaneously because some variables are excluded from each equation it is possible to identify the system.

#### 4. Data and econometric estimations

We use a micro survey conducted in Bolivia to test the implications of our model. Bolivia is a good case in assessing civil servants' performance as a function of the incentives and constraints provided by the institutional environment for two reasons: First, the country has gone through a series of reforms to improve the public sector. Hence, it would be useful for policy purposes to evaluate the relative importance of determinants of public sector performance. Second, it represents many characteristics of an emerging market economy and a Latin American country, and consequently, can be considered as a useful case study for development economics.

For much of the twentieth century, Bolivia was one of the most under-developed countries in the region, plagued by political instability, economic stagnation, social unrest, and pervasive corruption (Wiggins et al., 2006). The country's illiteracy level is higher, and life expectancy is lower than other South American countries. The political instability can, in part, be attributed to the fractured state of Bolivian society, which is divided along geographic, ethnic, ideological, and class-based lines. Bolivia has introduced some major economic and institutional reform measures in the 1990s. Many of these reforms were aimed at enhancing transparency and accountability in public affairs. These reform measures include a constitutional reform, the introduction of the National Electoral Court and other autonomous institutions (such as the Banking Superintendence and the Central Bank), and a process of political decentralization. The institutional model for democratic governance was further developed with the Popular Participation Law of 1994 and the complementary Administrative Decentralization Law of 1995, which created local governments with new and significant fiscal and administrative responsibilities. These reforms also set up a series of institutional mechanisms to allow citizen participation and oversight. Despite these efforts, the results are mixed, as most of the reform measures have not been fully implemented. According to

governance indicators of the World Bank, it still has considerably more corruption and less effective government than other countries in the region.

#### 4.1. Data

The source of our data set is a survey of public officials in Bolivia, conducted by the World Bank as a part of a regional project performed in Latin American and Caribbean countries in the 1990s and 2000s. It covers public officials working in 110 public institutions including the top executive branch, ministries, line agencies, autonomous agencies, and local governments. Within each institution, a stratified random sample of at least 1% of all the staff was selected at each of the decision-making ranks. A local consulting firm was hired to conduct the survey. To achieve reliable results interviews were administered face to face with public officials. The response rate was above 90%. Respondents were protected by anonymity to encourage candid answers. In total, 1250 public sector officials participated in the interviews.

The survey includes a range of questions that focus on public sector attributes. We group these questions into the following categories:

- Performance of public agencies: The average of two indicators: the quality and the quantity of services. Higher numbers correspond to higher performance in service delivery.
- Corruption: The average of three indicators: the frequency of bribery, bribe-to-official income ratio, and the percent of budget diverted illegally. Higher numbers correspond to higher corruption.
- Transparency: The fraction of cases where the actions of public officials and the decision-making are transparent to external parties.
- Enforcement of rules: The proportion of cases for which guidelines and regulations on personnel, budget, and service management are monitored and enforced.
- Meritocracy: The percentage of cases for which personnel decisions are based on the level of education, professional experience, merit, and performance.
- Politicization: The proportion of cases for which decisions on personnel, budget, and service management are subject to political interference.
- Citizen voice: The existence of mechanisms that guarantee consumer feedback and complaints.
- Availability of resources: The proportion of cases for which physical, financial, and human resources of the agency are considered adequate.

The second set of questions is intended to reveal the personal characteristics of respondents:

- Education: Percent of public officials with university education.
- Wage satisfaction: The extent of respondent's satisfaction with wages and other work-related benefits.
- Honesty: The extent of respondent's satisfaction with motivations to be honest and trustworthy.

The answers are scaled from 1 to 7, based on the degree of the respondent's agreement or disagreement with the statement (where 1 corresponds to "strongly disagree" or "never", and 7 corresponds to "strongly agree" or "always"). We rescale the responds from 0 to 1. Then, we construct variables at the public official level by taking the average of questions in each category (a detailed description of the survey questions we use to construct each variable is reported in Appendix A). Next, we aggregate individual responds over each agency.

We also construct two dummy variables to capture the institutional characteristics of the public agencies. One is "decentralization", which is equal to one if the agency is a municipality. The other is "autonomy", which is equal to one if the agency is an autonomous institution, such as the Central Bank, Electoral Court, and Ombudsman.

The descriptive statistics are reported in Table 1. Our sample size is 89. Each observation in our econometric model is a public agency. Note that the average scores and standard deviations show that although the survey responses at the public official level are discrete and lie between 0 and 1 (by construction), the aggregated composite variables at the agency level are quite far from the end points. Thus, rather than using a limited dependent variable model, we opt for a linear regression. We do not prefer a multinomial logit/probit model, since there are many, albeit countable, distinct values for the dependent variables.

Simple correlations are reported in Table 2. Most correlations have the expected signs and statistically significant at 5%. The simple correlations among endogenous variables (performance, corruption, and transparency) are quite strong, suggesting that they are either highly related to each other or there are common factors affecting these variables. It is noteworthy that the simple correlation between wages and performance is insignificant. It is also interesting that decentralization and the various measures of governance are inversely related, such as meritocracy, lack of politicization, and the education of public officials. One exception is the voice variable, which is positively correlated with decentralization.

## 4.2. Econometric model

Equations (9), (11), and (13) yield the following simultaneous system of equations:

$$\begin{aligned} \text{Performance} &= h_1(\text{Transparency}, Z^*, f_1, f_2, g(\cdot)) \\ \text{Corruption} &= h_2(\text{Performance}, \text{Transparency}, f_1, g(\cdot)) \\ \text{Transparency} &= h_3(\text{Corruption}, T^*, f_3, g(\cdot)) \end{aligned} \quad (14)$$

**Table 1.** Descriptive statistics.

| Variable         | N  | Average | Standard Deviation | Minimum | Maximum |
|------------------|----|---------|--------------------|---------|---------|
| Performance      | 89 | 0.613   | 0.123              | 0.206   | 0.825   |
| Corruption       | 89 | 0.405   | 0.205              | 0.414   | 0.880   |
| Transparency     | 89 | 0.510   | 0.092              | 0.305   | 0.759   |
| Politicization   | 89 | 0.404   | 0.177              | 0.223   | 0.643   |
| Enforcement      | 89 | 0.498   | 0.105              | 0.265   | 0.775   |
| Meritocracy      | 89 | 0.490   | 0.116              | 0.234   | 0.722   |
| Voice            | 89 | 0.349   | 0.071              | 0.225   | 0.620   |
| Decentralization | 89 | 0.186   | 0.391              | 0.000   | 1.000   |
| Education        | 89 | 0.537   | 0.181              | 0.211   | 0.740   |
| Resources        | 89 | 0.507   | 0.112              | 0.236   | 0.777   |
| Honesty          | 89 | 0.467   | 0.163              | 0.263   | 0.650   |
| Wages            | 89 | 0.321   | 0.135              | 0.200   | 0.578   |
| Autonomy         | 89 | 0.093   | 0.292              | 0.000   | 1.000   |

**Table 2.** Simple correlations.

|                  | Perfor. | Corru.  | Trans.  | Polic.  | Enfor.  | Merit.  | Voice   | Decen.  | Educ.   | Resou.  | Honest | Wages   | Auton.  |
|------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------|---------|---------|
| Performance      | 1.000   | -0.553* | 0.723*  | -0.323* | 0.580*  | 0.452*  | 0.679*  | -0.036  | 0.251*  | 0.624*  | 0.030  | 0.128   | 0.312*  |
| Corruption       |         | 1.000   | -0.495* | 0.345*  | -0.312* | -0.444* | -0.398* | 0.299*  | -0.317* | -0.370* | -0.131 | -0.303* | -0.329* |
| Transparency     |         |         | 1.000   | -0.474* | 0.610*  | 0.291*  | 0.723*  | -0.283* | 0.230*  | 0.519*  | 0.102  | 0.319*  | -0.261* |
| Politicization   |         |         |         | 1.000   | -0.288* | -0.417* | -0.177  | 0.283*  | -0.402* | -0.211  | -0.194 | -0.371* | -0.460* |
| Enforcement      |         |         |         |         | 1.000   | 0.211*  | 0.264*  | 0.089   | 0.184   | 0.524*  | 0.002  | 0.056   | 0.102   |
| Meritocracy      |         |         |         |         |         | 1.000   | 0.449*  | -0.222* | 0.400*  | 0.275*  | 0.218* | 0.320*  | 0.287*  |
| Voice            |         |         |         |         |         |         | 1.000   | 0.195*  | 0.074   | 0.273*  | 0.203* | -0.122  | 0.123   |
| Decentralization |         |         |         |         |         |         |         | 1.000   | -0.547* | 0.097   | 0.136  | 0.068   | -0.153  |
| Education        |         |         |         |         |         |         |         |         | 1.000   | 0.168   | 0.149  | -0.038  | 0.322*  |
| Resources        |         |         |         |         |         |         |         |         |         | 1.000   | 0.059  | -0.036  | 0.326*  |
| Honesty          |         |         |         |         |         |         |         |         |         |         | 1.000  | -0.058  | 0.126   |
| Wages            |         |         |         |         |         |         |         |         |         |         |        | 1.000   | -0.119  |
| Autonomy         |         |         |         |         |         |         |         |         |         |         |        |         | 1.000   |

Note. \* denotes significant at 5%.

The system has three endogenous variables: public sector performance, corruption, and transparency. The exogenous variables are  $Z^*$ ,  $T^*$ ,  $f_1$ ,  $f_2$ ,  $f_3$ , and  $g(\cdot)$ .

We use three measures for the service norm,  $Z^*$ : resources of the institution, decentralization, and education. Conceptually, there is a direct link between the service norm and resources. Institutions with more resources would have a higher target for performance. Municipalities and line agencies are also likely to have different service norms due to differences in the nature of public services and their closeness to the public. Therefore, we expect decentralization to shape service norms, but we do not have any a priori expectation about the sign of the coefficient. Lastly, service standards of an institution may be shaped by the quality of employees working in that institution. Thus, we expect higher service norms in institutions that have better educated employees.

$T^*$  is the transparency norm, which is measured by three variables. First one is the education of public officials. We assume that transparency norm would be higher in institutions that have better-educated workers. The second variable is the *de jure* autonomy of the institution. Rules and regulations would require autonomous agencies to be more transparent. The last variable is decentralization. Municipalities that are closer to the public would also be more transparent in their operations.

$f_1$  (the cost of deviating from the standard service level),  $f_2$  (the cost of deviating from the service norm), and  $f_3$  (the cost of deviating from the transparency norm) are captured by four variables: The first two variables are the enforcement of rules and citizen voice, which are related to the probability of getting caught. The third variable is politicization, which is a proxy for the probability of getting punished. The fourth variable is a meritocracy, which captures the link between the following institutional norms and the prospects for promotion.

We also include two additional variables in the corruption equation: wage and honesty, which are cited in the literature with a negative influence on the incidence of corruption (Dimant & Tosato, 2017).

The considerations above yield the following model of simultaneous equations:

$$\begin{aligned}
 \text{Performance} &= \alpha_0 + \alpha_1 \text{Transparency} + \alpha_2 \text{Politicization} + \alpha_3 \text{Enforcement} \\
 &\quad + \alpha_4 \text{Meritocracy} + \alpha_5 \text{Voice} + \alpha_6 \text{Decentralization} + \alpha_7 \text{Education} \\
 &\quad + \alpha_8 \text{Resources} + \epsilon_1 \\
 \\ 
 \text{Corruption} &= \beta_0 + \beta_1 \text{Performance} + \beta_2 \text{Transparency} + \beta_3 \text{Politicization} \\
 &\quad + \beta_4 \text{Enforcement} + \beta_5 \text{Meritocracy} + \beta_6 \text{Voice} + \beta_7 \text{Decentralization} \\
 &\quad + \beta_8 \text{Honesty} + \beta_9 \text{Wage} + \epsilon_2 \\
 \\ 
 \text{Transparency} &= \gamma_0 + \gamma_1 \text{Corruption} + \gamma_2 \text{Politicization} + \gamma_3 \text{Enforcement} \\
 &\quad + \gamma_4 \text{Meritocracy} + \gamma_5 \text{Voice} + \gamma_6 \text{Decentralization} + \gamma_7 \text{Education} \\
 &\quad + \gamma_8 \text{Autonomy} + \epsilon_3
 \end{aligned}
 \tag{15}$$

$$\begin{aligned}
 \text{Transparency} &= \gamma_0 + \gamma_1 \text{Corruption} + \gamma_2 \text{Politicization} + \gamma_3 \text{Enforcement} \\
 &\quad + \gamma_4 \text{Meritocracy} + \gamma_5 \text{Voice} + \gamma_6 \text{Decentralization} + \gamma_7 \text{Education} \\
 &\quad + \gamma_8 \text{Autonomy} + \epsilon_3
 \end{aligned}$$



### 4.3. Estimation method

We use two methods to estimate the system. First, we estimate each equation separately using the seemingly unrelated regression (SURE) assuming that the variables that are suspected of being endogenous are in fact exogenous. We prefer SURE to the ordinary least squares (OLS), since the disturbance terms of the equations are likely to be correlated with each other due to reasons such as group level-fixed effects that influence decision-making or operations across agencies in a similar pattern.<sup>2</sup> When this is the case, SURE is more efficient than OLS. We also look for potential unobservable effects across regions (which are geographic units in the Bolivia and eight in total) by using a fixed-effects estimation approach, but F-test does not support the existence of any fixed effects (the minimum p-value of the test statistic is 0.72).

Second, we use an instrumental variable approach since our theoretical model suggests that three variables (performance, corruption, and transparency) may be subject to endogeneity. Two-stage least squares (2SLS) are by far the most-used estimator for the simultaneous equation problem. However, as discussed below, the instruments in our study are not quite strong. The weakness of instruments makes the estimators of 2SLS substantially biased in small samples with large variance (Bound, Jaeger, & Baker, 1995; Staiger & Stock, 1997). Another popular method, the generalized method of moments (GMM) also suffers from the same problem (Stock, Wright, & Yogo, 2002).

Several alternative estimation methods that are more robust to weak instruments have been proposed in the literature, mostly in the context of k-class estimators.<sup>3</sup> One example is limited information maximum likelihood (LIML) estimator. With independent and identically distributed (i.i.d.) disturbances, LIML estimator is consistent and asymptotically normally distributed (Bekker, 1994). The Fuller estimator is another example with similar properties, but even fewer outliers than LIML (Hahn, Hausman, & Kuersteiner, 2004). Unfortunately, both LIML and Fuller are inconsistent under heteroskedasticity (Bekker & Van der Ploeg, 2005). Moreover, LIML is sensitive to large small sample variability due to lack of moments in finite samples. A third robust method is the continuously updated GMM estimator, CUE.<sup>4</sup> It is efficient under heteroskedasticity since it allows for general non-spherical disturbances (Hausman, Menzel, Lewis, & Newey, 2007). However, CUE suffers from the lack-of-moment problem as well. Monte Carlo studies show that estimators with well-defined sample moments (Fuller and 2SLS) usually perform better, since lack of moments leads to wide dispersion in estimates in extremely weak instruments situations, i.e., when  $R^2$  of the first-stage regression is 0.1

<sup>2</sup>We use the Breusch–Pagan LM test to assess whether the disturbances are correlated across equations. The test statistic is 14.517 with a p-value of 0.00, rejecting the null hypothesis of no correlation. The correlation of residuals is  $-0.16$  between performance and corruption equations,  $0.27$  between performance and transparency equations, and  $-0.09$  between corruption and transparency equations.

<sup>3</sup>The k-class estimator of  $\beta$  is  $\hat{\beta}(k) = [Y'(I - kM_Z)Y]^{-1} [Y'(I - kM_Z)y]$  for a regression model of the form  $y = Y\beta + u$  and  $Y = Z\Pi + v$ , where  $y$  is the dependent variable,  $\beta$  is the coefficients of the structural equation,  $Y$  is the set of regressors (including both exogenous and endogenous variables) and  $Z$  is the set of instruments. Note that  $k=0$  corresponds to OLS,  $k=1$  corresponds to 2SLS,  $k=k_{LIML}$  corresponds to LIML, and  $k=k_{LIML} - \alpha/(n-K)$  corresponds to Fuller(a) where  $n$  is the sample size and  $K$  is the number of instruments.

<sup>4</sup>CUE introduced by Hansen, Heaton, and Yaron (1996) is the GMM-like generalization of LIML. The difference between the two-step GMM and continuously updated GMM is that in the former, a fixed weighting matrix is used in the calculation of estimators. In CUE, the weighting matrix is not fixed, but a function of the estimators and its estimation is done simultaneously with an estimation of other parameters via numerical methods.

or lower (Hahn et al., 2004). When  $R^2$  of the first-stage regression is above 0.3, on the other hand, finite sample problems associated with lack of sample moments cease to be a concern.

In our sample, as discussed below, our instruments are not very strong but first-stage  $R^2$  statistics are well above 0.3. The error terms in all three equations, on the other hand, exhibit heteroscedasticity, especially performance and corruption equations.<sup>5</sup> Based on these findings, we choose CUE as our preferred estimation method. However, we should note that when we repeat our estimation using LIML, Fuller, and 2SLS we do not observe any meaningful changes in our estimates, although the fitness of the model slightly declines in all three methods.

#### 4.4. Instruments

Before discussing the regression results, we introduce the instruments and the associated test statistics that assess the appropriateness of the instruments. The simultaneous equations model introduced in the previous section suggests that autonomy, honesty, and wages are the excluded instruments in the performance equation; education, resources, and autonomy are the excluded instruments in the corruption equation; resources, honesty, and wages are the excluded instruments in the transparency equation. Note that being an excluded instrument does not mean that it has no effect on the dependent variable. Its effect may be indirect through its influence on the endogenous regressors in that equation.

A variable must satisfy two conditions to be a valid instrument: (1) Relevance, i.e., being sufficiently correlated with the variable that is suspected of being endogenous, and (2) Exogeneity, i.e., being distributed independently of the error process. The first condition requires testing under-identification and weak-identification.<sup>6</sup> The second condition, on the other hand, necessitates testing the orthogonality condition.

For under-identification, two statistics are widely used: Anderson canonical correlations Lagrange multiplier and Kleibergen-Paap rk Lagrange multiplier. The first statistic assumes that the errors are homoscedastic, whereas the latter is also valid under heteroscedasticity (Kleibergen & Paap, 2006). We choose the latter due to heteroscedasticity in our sample. The results reported in the first row of Table 3 strongly reject the null hypothesis of under-identification. However, this result should be treated with caution, because the threshold for under-identification is very low.

For detecting weak instruments, various informal procedures are available, such as first-stage partial  $R^2$  and first-stage partial F-statistics on the excluded instruments (Stock et al., 2002). The former measures the contribution of excluded instruments to the explained variation in the endogenous variable, whereas the latter measures whether

<sup>5</sup>The Breusch-Pagan statistic, which is the standard test used for that purpose can only be used in the systems of equations if heteroskedasticity is present in the equation of interest and nowhere else in the system, i.e., the other structural equations in the system must be homoscedastic (Pagan & Hall, 1983). Hence, we use a more general test suggested by Pagan and Hall (1983) that relaxes this restriction. The chi-square test statistic is 18.447 with a p-value of 0.05 for performance equation, 31.049 with a p-value of 0.00 for the corruption equation, and 15.910 with a p-value of 0.10 for transparency equation. The null hypothesis is a constant variance of the error term.

<sup>6</sup>The difference between under identification and weak identification is that in the former the rank condition is violated and consequently the model is not identified. In the latter, the rank condition holds, the model is identified in a finite sample, but the amount of information available to estimate the parameters does not increase with the sample.

**Table 3.** Estimation results – first stage.

| Null Hypothesis  | Test statistics                                      | Dependent Variable: Performance | Dependent Variable: Corruption                | Dependent Variable: Transparency |
|--|--|---------------------------------|---|----------------------------------|
| <b>Instrument Relevance tests</b>  |  |                                 |   |                                  |
| Excluded instruments are irrelevant (under-identification)   | Kleibergen–Paap rk LM statistic                      | 10.835 (0.01)                   | 8.676 (0.01)                                  | 7.973 (0.05)                     |
| All coefficients of excluded instruments are zero  | Partial F statistics in the first stage regression   | 12.651††                        | Perform: 5.591†<br>Transp: 9.602††            | 5.371†                           |
|  | Partial R <sup>2</sup> in the first stage regression | 0.209                           | Perform: 0.186<br>Transp: 0.180               | 0.128                            |
| Excluded instruments are weak (weak-identification)  | Cragg–Donald Wald statistic                          | 7.217††                         | 5.322†  | 4.293†                           |
|  | Kleibergen–Paap rk Wald F statistic                  | 12.651††                        | 10.575††                                      | 5.371†                           |
| <b>Instrument Exogeneity (over-identification) test</b>  |  |                                 |   |                                  |
| Instruments are not correlated with the error term and correctly excluded from the structural equation | Sargan-Hansen J-statistic                            | 1.687 (0.43)                    | 0.006 (0.94)                                  | 2.459 (0.29)                     |
| <b>Test on exogeneity of variables treated as endogenous</b>   |  |                                 |   |                                  |
| Endogenous regressor tested is exogenous   | Difference-in-Sargan C statistic                     | 5.239 (0.02)                    | Perfor.: 2.158 (0.14)<br>Transp. 3.927 (0.05) | 8.160 (0.00)                     |

Note: P-values of LM, Sargan J, and Sargan C statistics are in parenthesis. Stock–Yogo critical values for partial F statistics are 6.46 for 10% and 4.36 for 15% maximal relative bias. Stock–Yogo critical values for weak identification tests (used for Cragg–Donald Wald and are Kleibergen–Paap rk Wald F statistics) are 5.44 for 10% and 3.81 for 15% maximal relative bias. These critical values are for single endogenous regressor, and they constitute more conservative thresholds (higher than they should be) in the case of two endogenous variables. †† denotes significance at 10% and † denotes significance at 15% according to Stock–Yogo critical values.

the coefficients of the excluded instruments are significantly different than zero. First-stage regression results are reported in Table 4. As a rule of thumb, Staiger and Stock (1997) reports that the first-stage F statistic must be large, typically exceeding 10 for inference to be reliable. In our case, some F-statistics are below but close to 10 (second row of Table 3) and partial R<sup>2</sup> is not very high (third row of Table 3), suggesting that our instruments may be relevant but not very strong.

First-stage methods used for assessing instrument relevance are informal, rely on arbitrarily chosen rules-of-thumbs, and do not distinguish between the many instruments problem and the weak instrument problem. Moreover, when there is more than one endogenous variable on the right-hand side, the F test and partial R<sup>2</sup> may not be good tests of relevance. As a more rigorous and formal test procedure, Stock and Yogo (2005) used the Cragg–Donald statistic, which is a generalization of the F statistic. The null hypothesis being tested is that the estimator is weakly identified in the sense that it is subject to bias that is unacceptably large.<sup>7</sup> The critical values of the test are determined by the number of instruments, the number of included endogenous regressors, and the size of estimator bias

<sup>7</sup>The Stock–Yogo weak instruments test identifies bias in two forms: maximal IV estimator bias and maximal Wald test size bias. The first one is based on the ratio of the bias of the estimator to the bias of OLS. The second one is based on the performance of the Wald statistic used for hypothesis testing. Following the common practice in the literature, we use the first form.

**Table 4.** Estimation results – first stage.

| Variable             | Dependent Variable:<br>Performance | Dependent Variable:<br>Corruption | Dependent Variable:<br>Transparency |
|----------------------|------------------------------------|-----------------------------------|-------------------------------------|
| Politicization       | -0.158**<br>(-2.34)                | 0.292**<br>(2.57)                 | -0.288**<br>(-2.52)                 |
| Enforcement          | 0.098<br>(1.21)                    | -0.130<br>(-1.54)                 | 0.223*<br>(1.94)                    |
| Meritocracy          | 0.137<br>(1.49)                    | -0.353***<br>(-3.68)              | 0.176<br>(1.51)                     |
| Voice                | 0.512***<br>(3.94)                 | -0.579***<br>(-3.05)              | 0.683***<br>(3.67)                  |
| Decentralization     | -0.024<br>(-1.22)                  | 0.044*<br>(1.80)                  | -0.010<br>(-0.37)                   |
| Education            | 0.173**<br>(2.38)                  | -0.088<br>(-1.22)                 | 0.159**<br>(2.09)                   |
| Resources            | 0.322**<br>(2.61)                  | -0.121<br>(-1.42)                 | 0.067<br>(0.78)                     |
| Honesty              | 0.017<br>(0.39)                    | -0.133<br>(-1.57)                 | 0.062<br>(1.34)                     |
| Wages                | 0.068<br>(0.95)                    | -0.202**<br>(-2.37)               | 0.105*<br>(1.77)                    |
| Autonomy             | -0.034<br>(-0.87)                  | -0.027<br>(-0.65)                 | -0.125***<br>(-3.02)                |
| <b>R<sup>2</sup></b> | 0.632                              | 0.540                             | 0.591                               |

Notes: Sample size is 89. t-statistics of the coefficient estimates are in parenthesis. \*\*\* Significant at 1%, \*\* significant at 5%, \* significant at 10%.

that is acceptable. Moreover, the critical values are different for different estimators because the estimators are not affected to the same degree by weak instruments.<sup>8</sup> Since the Cragg-Donald-based weak instruments test assumes homoscedastic errors, we also report Kleibergen-Paap Wald rk F statistic which is valid under non-i.i.d. errors. Our results presented in Table 3 (fourth and fifth rows) show that Cragg-Donald statistics are above the 15% critical values in all three equations, but above the 10% critical value only in the performance equation. Kleibergen-Paap Wald rk F statistics are above the 10% critical value in performance and corruption equations, whereas it is above the 15% critical value in transparency equation. Overall, we reject the weakness of instruments, but the evidence is not overwhelming.

Having discussed the relevance of instruments, we now turn to the second condition of being a valid instrument, i.e., exogeneity of instruments. We use Sargan-Hansen J-test for that purpose (Hansen, 1982). The null hypothesis that the instruments are valid cannot be rejected at the 5% level of significance (sixth row of Table 3).

Finally, we test the endogeneity of performance, corruption and transparency variables. Obviously, if the problem of endogenous regressors does not exist, there is no need to use the IV method since the OLS method would give us consistent and more efficient estimators than the IV method. The test statistics we use for that purpose are the difference of two Sargan-Hansen statistics: one for the equation with the smaller set of instruments, where the suspect regressor(s) is treated as endogenous, and one for the equation with the larger set of instruments, where the suspect regressor(s) is treated as exogenous. The null hypothesis is that the regressor(s) deemed endogenous is in fact

<sup>8</sup>Specifically, since LIML and CUE estimators are more robust to the presence of weak instruments than 2SLS and GMM, the critical values associated with these estimators tabulated by Stock and Yogo (2005) are tighter.

exogenous. The test statistics reported in the last row of Table 3 show that the exogeneity of transparency variable in the performance equation is rejected at 5% significance level, the exogeneity of transparency variable in the corruption equation is rejected at 5% significance level, and the exogeneity of corruption variable in the transparency equation is rejected at 1% significance level. We fail to reject the exogeneity of performance variable in the corruption equation, but the test statistic is close to being significant and theoretically one should expect causality from corruption to performance. Hence, we still treat that variable as endogenous to be cautious.

#### 4.5. Results

The estimation results corresponding to SURE and CUE are reported in Table 5.

Starting with the performance equation, we observe that once the influence of potentially important variables is considered via a multiple linear regression model, the partial correlations differ substantially from the simple correlations. This highlights the large bias in simple correlation and the danger in designing policies based on those statistics. Several regressors, such as politicization, enforcement of rules, and meritocracy, which are highly (pair-wise) correlated with performance have insignificant coefficients once other factors are controlled for. The significant variables in the SURE are transparency, voice, and resources with expected signs. Politicization and

**Table 5.** Estimation results.

| Explanatory Variable | Model  | Dependent Variable:<br>Performance |                    | Dependent Variable:<br>Corruption |                      | Dependent Variable:<br>Transparency |                      |
|----------------------|--|------------------------------------|--------------------|-----------------------------------|----------------------|-------------------------------------|----------------------|
|                      |  | SURE                               | CUE                | SURE                              | CUE                  | SURE                                | CUE                  |
| Performance          |  |                                    |                    | -0.200**<br>(-1.93)               | -0.085<br>(-0.72)    |                                     |                      |
| Corruption           |  |                                    |                    |                                   |                      | -0.345***<br>(-3.78)                | -0.302**<br>(-2.60)  |
| Transparency         |  | 0.372***<br>(3.48)                 | 0.259**<br>(2.41)  | -0.502***<br>(-3.97)              | -0.312***<br>(-3.19) |                                     |                      |
| Politicization       | f <sub>1</sub> , f <sub>2</sub> , f <sub>3</sub> | -0.126*<br>(-1.90)                 | -0.114*<br>(-1.79) | 0.195**<br>(1.99)                 | 0.267***<br>(2.84)   | -0.178*<br>(-1.85)                  | -0.235**<br>(-2.04)  |
| Enforcement          | f <sub>1</sub> , f <sub>2</sub> , f <sub>3</sub> | 0.066<br>(0.97)                    | 0.054<br>(1.03)    | -0.040<br>(-0.55)                 | -0.024<br>(-0.21)    | 0.197**<br>(2.21)                   | 0.173<br>(1.28)      |
| Meritocracy          | f <sub>1</sub> , f <sub>2</sub> , f <sub>3</sub> | 0.075<br>(0.67)                    | 0.166<br>(1.08)    | -0.267***<br>(-3.12)              | -0.310***<br>(-2.72) | 0.092<br>(1.09)                     | 0.119<br>(1.03)      |
| Voice                | f <sub>1</sub> , f <sub>2</sub> , f <sub>3</sub> | 0.474***<br>(4.28)                 | 0.457***<br>(3.18) | -0.317**<br>(-2.22)               | -0.397**<br>(-2.03)  | 0.508***<br>(2.87)                  | 0.541**<br>(2.50)    |
| Decentralization     | Z*, T*   | -0.022<br>(-1.26)                  | -0.012<br>(-0.24)  | 0.031<br>(1.56)                   | 0.036*<br>(1.87)     | -0.014<br>(-0.75)                   | 0.019<br>(1.12)      |
| Education            | Z*, T*   | 0.097*<br>(1.81)                   | 0.158**<br>(2.16)  |                                   |                      | 0.125*<br>(1.92)                    | 0.134*<br>(1.77)     |
| Resources            | Z*   | 0.246***<br>(2.95)                 | 0.293**<br>(2.41)  |                                   |                      |                                     |                      |
| Autonomy             | T*   |                                    |                    |                                   |                      | -0.078**<br>(-2.25)                 | -0.107***<br>(-2.81) |
| Honesty              |  |                                    |                    | -0.114*<br>(-1.83)                | -0.098<br>(-1.24)    |                                     |                      |
| Wages                |  |                                    |                    | -0.156**<br>(-2.37)               | -0.152**<br>(-2.11)  |                                     |                      |
| R <sup>2</sup>       |  | 0.673                              | 0.620              | 0.562                             | 0.515                | 0.573                               | 0.541                |

Notes: Sample size is 89. All variables lie in the [0,1] interval. t-statistics of the coefficient estimates are in parenthesis.

\*\*\* Significant at 1%, \*\* significant at 5%, \* significant at 10%.

education are also significant, albeit marginally. These results suggest that the performance of agencies improves substantially if these agencies are equipped with more transparency, citizen voice, and resources, as well as better educated public officials. Politicization of public institutions, on the other hand, is detrimental to their performance. When we address the potential endogeneity of transparency via the CUE method, the coefficient of transparency still has a substantial positive influence over performance, but its magnitude and significance diminish. This may be due to the presence of some unobservable factors that affect both performance and transparency in the same direction leading to a perception that the former would be positively correlated with the latter. Using an IV approach would correct this positive bias and provide a more reliable estimate.

Turning to the SUR estimates on corruption, we observe that performance, transparency, meritocracy, and wages have significant and negative coefficients, whereas politicization is positively associated with corruption. Honesty is borderline significant. Variables like decentralization and enforcement that are highly (pairwise) correlated with corruption have no substantial influence on corruption when other potentially relevant factors are considered. As discussed above transparency and performance may be subject to the reverse causality problem. Therefore, we repeat our estimation using the CUE method. Performance and honesty are no longer significant, whereas decentralization turns out to be marginally significant with a positive sign. The loss of performance as a significant variable is noteworthy, suggesting that the causality from that variable to corruption is the result of a feedback loop, which disappears once an IV technique is employed.

Finally, according to the SURE results, transparency is positively influenced by enforcement and voice, and negatively influenced by corruption, politicization, and autonomy. Education also has a positive influence on transparency, but its effect is somewhat marginal. Meritocracy and decentralization that have a high pair-wise correlation with transparency lose their significance once the influence of other factors is considered. When we address the endogeneity of corruption, that variable remains significant.

When there are several variables that are determined endogenously, the full effect of a change in any of the exogenous variable may be very different from the direct effect. For example, an increase in voice would improve transparency that would reduce corruption that would increase transparency even further. To estimate the full effect of policy variables one must calculate the reduced form equations (i.e., incorporating the linkage between various endogenous variables). First-stage results reported in [Table 4](#) can be used for that purpose. The effect of education and voice on performance increases as these variables strengthen transparency that, in turn, boost performance. Similarly, the effect of politicization on performance also becomes stronger because it causes more corruption and less transparency, which undermines performance. However, none of the public-sector management factors, such as rule enforcement, meritocracy, autonomy by fiat, and government pay, or individual characteristics such as honesty seems to have a significant impact on performance.

We discuss possible caveats and robustness tests in Appendix B, including the measurement errors in variables and alternative econometric specifications.

## 5. Conclusions

Drawing on an in-depth governance micro-survey within a country, we address empirically the question of the relative importance of the various determinants of governance. We find that commonly made inferences based on simple correlation can be highly misleading, given the high degree of multicollinearity between the various governance (and public sector management) determinants, as well as the endogeneity in these variables. In fact, if policy recommendations were to be made on the basis of simple correlations, undue emphasis would have been given to certain public sector management variables (such as relative wages, internal enforcement of rules, autonomy of agency by fiat, etc.), while undermining more important variables such as external voice, transparency, and the absence of politicization. The latter set of variables come out clearly significant and accounting for a much larger share of the variation than the former, more traditional public sector management type of variables.

We need to be particularly wary of implying that the above results would necessarily hold in other countries. Indeed, our claims at this early stage of this type of research with this new type of survey data ought to be modest. While in any country there would be a set of determinants that matter significantly more than others, these variables may vary from setting to setting. We plan to undertake a similar analysis of agency performance in other countries for which we have gathered data.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Notes on contributors

**Daniel Kaufmann** is the president and CEO of the Natural Resource Governance Institute. A recognized innovator and expert on governance and anticorruption around the world, he is currently a member of the Extractive Industries Transparency Initiative international board and serves in various international advisory boards, including the Organisation for Economic Cooperation and Development Secretary-General's High-Level Advisory Group on Anticorruption and Integrity and the Inter-American Development Bank's President High-Level Advisory Group on Transparency. He has also held several high-level positions at the World Bank. Kaufmann, a Chilean, received an M.A. and Ph.D. in economics at Harvard, and a B.A. in economics and statistics from the Hebrew University of Jerusalem.

**Gil Mehrez** received his Ph.D. in economics at University of California, Berkeley. He taught at Georgetown University and worked at the World Bank. He published extensively on monetary economics and governance.

**Tugrul Gurgur** received his Ph.D. in economics from University of Maryland, College Park. Previously, he worked at the World Bank and the Central Bank of Turkey. His research interests include development economics, governance, and econometrics.

## References

- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76, 169–217.



- Becker, G. S., & Stigler, G. J. (1974). Law enforcement, malfeasance, and compensation of enforcers. *Journal of Legal Studies*, 3, 1–18.
- Bekker, P. A. (1994). Alternative approximations to the distributions of instrumental variable estimators. *Econometrica*, 62, 657–681.
- Bekker, P. A., & Van der Ploeg, J. (2005). Instrumental variable estimation based on grouped data. *Statistica Neerlandica*, 59, 239–267.
- Besley, T., & McLaren, J. (1993). Taxes and bribery: The role of wage incentives. *Journal of Legal Studies*, 103, 119–141.
- Bound, J., Jaeger, D. A., & Baker, R. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak. *Journal of the American Statistical Association*, 90, 443–450.
- Brunetti, A., & Weder, B. (2003). A free press is bad news for corruption. *Journal of Public Economics*, 87, 1801–1824.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- Del Monte, A., & Papagni, E. (2007). The determinants of corruption in Italy: Regional panel data analysis. *European Journal of Political Economy*, 23, 379–396.
- Di Tella, R., & Schargrodsky, E. (2003). The role of wages and auditing during a crackdown on corruption in the city of Buenos Aires. *Journal of Law and Economics*, 96, 269–292.
- Dimant, E., & Tosato, G. (2017). Causes and effects of corruption: What has past decade's research taught us? A survey. *Journal of Economic Surveys*, 32, 335–356.
- Escaleras, M., Lin, S., & Register, C. (2010). Freedom of information acts and public sector corruption. *Public Choice*, 145, 435–460.
- Evans, P., & Rauch, J. (2000). Bureaucratic structure and economic performance in less developed countries. *Journal of Public Economics*, 75, 49–71.
- Fan, C. S., Lin, C., & Treisman, D. (2009). Political decentralization and corruption: Evidence from around the world. *Journal of Public Economic*, 93, 14–34.
- Ferraz, C., & Finan, F. (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, 191, 1274–1311.
- Fisman, R., & Gatti, R. (2002). Decentralization and corruption: Evidence across countries. *Journal of Public Economics*, 83, 325–345.
- Fuller, W. A. (1977). Some Properties of the limited information estimator. *Econometrica*, 45, 939–954.
- Goel, R. K., & Nelson, M. A. (2011). Measures of corruption and determinants of US corruption. *Economics of Governance*, 12, 155–176.
- Gray-Molina, G., de Rada, E. P., & Yáñez, E. (1999). Transparency and accountability in Bolivia: Does voice matter? Working Paper no. R-381, Inter-American Development Bank. Washington, D.C.
- Gurgur, T. (2016). Voice, exit and local capture in public provision of private goods. *Economics of Governance*, 17, 397–424.
- Gurgur, T., & Shah, A. (2015). Localization and corruption: Panacea or Pandoras box? *Annals of Economics and Finance*, 15, 109–136.
- Hahn, J., Hausman, J., & Kuersteiner, G. (2004). Estimation with weak instruments: Accuracy of higher-order bias and MSE approximations. *Econometrics Journal*, 7, 272–306.
- Hansen, L. P. (1982). Large sample properties of GMM estimators. *Econometrica*, 50, 1029–1054.
- Hansen, L. P., Heaton, J., & Yaron, A. (1996). Finite-sample properties of some Alternative GMM estimators. *Journal of Business and Economic Statistics*, 14, 262–280.
- Hausman, J., Menzel, K., Lewis, R., & Newey, W. (2007). A reduced bias GMM-like estimator with reduced estimator dispersion. Working Paper, Institute for Fiscal Studies, CeMMAP.
- Islam, R. (2006). Does transparency go along with better governance? *Economics and Politics*, 18, 121–167.
- Kaufmann, D., & Wei, S. (2000). Does grease money speed up the wheels of commerce? Policy Research Working Paper 2196. Washington, D.C.

- Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133, 97–126.
- Kolstad, I., & Wiig, A. (2009). Is transparency the key to reducing corruption in resource-rich countries. *World Development*, 37, 532.
- LaPorta, R., Silanes, F. L., Shleifer, A., & Vishny, R. (1999). The quality of government. *Journal of Law, Economics and Organization*, 15, 222–279.
- Lindstedt, C., & Naurin, D. (2010). Transparency is not enough: Making transparency effective in reducing corruption. *International Political Science Review*, 31, 301–322.
- Lui, F. (1985). An equilibrium queuing model of bribery. *Journal of Political Economy*, 93, 760–781.
- Menes, R. (1999). The effect of patronage politics on city government in American cities, 1900–1910. NBER Working paper 6975, Cambridge, MA.
- Mookherjee, D., & Png, I. P. L. (1995). Corruptible law enforcers: How should they be compensated? *Economic Journal*, 105, 145–159.
- Olken, B. A. (2009). Corruption perceptions vs. corruption reality. *Journal of Public Economics*, 93, 950–964.
- Pagan, A. R., & Hall, D. (1983). Diagnostic tests as residual analysis. *Econometric Reviews*, 2, 159–218.
- Parente, P. M. D. C., & Silva, J. M. C. S. (2012). A cautionary note on tests of overidentifying restrictions. *Economics Letters*, 115, 314–317.
- Porto, A., & Porto, N. (2000). Fiscal decentralization and voters' choices as control. *Journal of Applied Economics*, 3, 135–167.
- Prud'homme, R. (1994). On the dangers of decentralization. World Bank Policy Research Paper 1252, World Bank Institute, Washington D.C.
- Putnam, R. (1993). *Making democracy work: Civic tradition in modern Italy*. NJ, USA: Princeton University Press.
- Reinikka, R., & Svensson, J. (2005). Fighting corruption to improve schooling: Evidence from a newspaper campaign in Uganda. *Journal of the European Economic Association*, 2, 1–9.
- Staiger, D., & Stock, J. H. (1997). IV regression with weak instruments. *Econometrica*, 65, 557–586.
- Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. Chapter 5 In J. Rothenberg, J. H. Stock, & D. W. K. Andrews (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas* (pp. 80–108). NY: Cambridge University Press.
- Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20, 518–529.
- Svensson, J. (2003). Who must pay bribes and how much? *Quarterly Journal of Economics*, 118, 207–230.
- Tanzi, V. (1993). Corruption: Arm's-length relationships and markets. In G. Fiorentini & S. Peltzman (Eds.), *The Economics of Organized Crime* (pp. 161–180). NY: Cambridge University Press.
- Tiebout, C. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64, 416–424.
- Treisman, D. (2007). What have we learned about the causes of corruption from ten years of cross-empirical research? *Annual Review of Political Science*, 10, 211–244.
- Van Rijckeghem, C., & Weder, B. (2001). Bureaucratic corruption and the rate of temptation: Do Wages in the civil service affect corruption, and by how much. *Journal of Development Economics*, 65, 307–331.
- Wiggins S., Schejtman A., & Gray G. (2006). Bolivia case study. Working Paper Series 1. Development Studies Institute (DESTIN). London School of Economics & Political Science. United Kingdom.
- Yamamura, E., & Kondoh, H. (2013). Government transparency and expenditure in the rent-seeking industry: The case of Japan for 1998–2004. *Contemporary Economic Policy*, 31, 635–647.

## APPENDICES

### Appendix A. Construction of governance variables

The Public Officials Survey consists of more than 200 questions that are mostly related to different aspects of governance. Although it was possible to choose one representative question for each dimension of governance, we did not opt for this option for two reasons. First, choosing only one question is bound to be arbitrary since it is not based on objective criteria. Second, one question may be too “noisy” because of potential measurement errors or because it may fail to measure the aspect of governance we are interested in.

Instead, we preferred to group several questions of similar nature. One way of grouping is by taking the simple average. This is not, however, the best method since it gives each question the same weight. Instead, we used factor analytic techniques to detect the common structure in the information content of the questions. Thus, the aggregate variable represents only the information that is common to all sub-components. The reliability of all governance variables was checked using the Cronbach’s alpha test (Cronbach, 1951). Cronbach’s alpha is a measure used to assess the internal consistency of questions in describing a common concept. The alpha coefficient is never less than 0.80. In most of the cases, the coefficient is higher than 0.90.

Survey questions used to construct governance variables are as follows:

#### ENFORCEMENT of Rules in Personnel/Budget/Service Management

- There exists some formal procedure to evaluate the performance of the employees.
- The policy/guidelines/regulations of personnel management are well supervised (violations are always exposed).
- The policy/guidelines/regulations of personnel management are strictly enforced (violations are always punished).
- The policy/guidelines/regulations of budget management are well supervised (violations are always exposed).
- The policy/guidelines/regulations of budget management are strictly enforced (violations are always punished).
- The policy/guidelines/regulations of service management are well supervised (violations are always exposed).

#### POLITICIZATION of public institutions

- Decisions on personnel management are based on political connections or party affiliations.
- Decisions on budget management are based on political connections or party affiliations.
- Decisions on service management are based on political connections or party affiliations.

#### TRANSPARENCY of decisions in personnel/budget/service management

- Decisions on personnel management are done transparently (everybody knows who were designated, promoted, transferred, ascended, or received wage increases and why).
- Decisions on budget management are done transparently.
- Decisions on daily tasks are done transparently.
- Decisions on service delivery/performance of daily tasks are announced and opened to the internal of the institution (and also to the outside if they are applicable).
- There is a clear understanding of what our tasks and responsibilities are.
- I understand clearly what we have to do and that is why our performance is evaluated.

## **MERITOCRACY in public institutions**

- Percent of the cases the decisions on personnel management are based on professional experience.
- Percent of the cases the decisions on personnel management are based on merit.
- Percent of the cases the decisions on personnel management are based on the level of education.

## **VOICE**

- We all consider that citizens and users are our clients.
- Decisions on service delivery/performance are done based on users' complaint.
- Clearly defined mechanisms exist to ask users about their needs.
- Clearly defined mechanisms exist so that the users can express their preferences.

## **WAGES**

- Percent of employees satisfied with their wages.
- Percent of employees satisfied with their benefits (pension, health, etc.)

## **HONESTY**

- The probability that if a public official was overpaid by an administrative error, the public official will return the money given that there is 100% chance of not getting caught and the superiors are doing the same without getting caught.

## **RESOURCES**

- Quantity of resources of the agency is adequate.
- Quality of resources of the agency is adequate.
- Office supplies/Computers of the agency are adequate.
- Space/Offices of the agency are adequate.

## **Appendix B. Caveats and robustness tests**

### ***B.1. Measurement Errors in Variables***

The Public Officials survey is based on the perceptions of public officials about the institutions they are working in. Although this approach is useful in cases where hard data are difficult or impossible to produce, perceptions are subject to respondent bias and other measurement problems. Recent studies based on micro surveys suggest that people's perceptions do contain real information but reported perceptions may also be systematically biased due to respondents' characteristics, such as education, gender or race (Olken, 2009).

We classify measurement problems into two groups: (i) individual bias, and (ii) institutional bias.

First, public officials may overstate or understate some features of their institution due to differences in their perception caused by differences in their individual characteristics. Our sample in each institution is very diverse in terms of education, gender, age, and so on. Since we use the institutional average as an observation unit, these perception errors caused by individual characteristics are likely to cancel out each other and not carry over to the institutional level.

Second, it is possible that all individuals working in a particular agency may be more pessimistic or optimistic in their perceptions due to working conditions within their organization. This is a major concern particularly in cross-country studies in which a common reference point or criteria to measure qualitative variables might be impossible to find due to cultural differences between societies. We do not believe that this is the case in our data since all observations are from the same country, sharing a common culture, norms, and moral values. Therefore, it is reasonable to assume that each person more or less uses the same criteria to make a judgment about the conditions of her institution and differences in perceptions (if any) are individual-specific rather than institutional specific.

To test the validity of our arguments above we use a survey question, which should presumably be answered in the same way by all public officials: the corruption level in the Bolivian public sector in general. Deviation from the sample average captures the influence of individual characteristics as well as the institutional effect. We find that institutional bias is not statistically significant and hence conclude that the bias is very small or nil.

## **B.2 Model Specification and Instruments**

The instruments used in the first stage of our IV estimations are the entire set of exogenous variables of the model. The choice of adequate instruments for corruption and performance is not extensively addressed in the literature (see, however, Svensson, 2003) and even almost non-existent in the case of transparency.

It is possible that some of the exogenous variables that are used in our model are endogenous as well. For example, one could argue that wage which is assumed to be an exogenous variable in the corruption equation may be endogenous, if agencies with higher corruption pay lower wages. Using a variation of the Sargan–Hansen statistics, we test the possibility that some of the exogenous variables are indeed endogenous but find no evidence that any of the exogenous variables are indeed endogenous.<sup>9</sup>

Another possible specification error is omitted variable problem, i.e., we may incorrectly exclude some variables from the structural equations. These omitted variables may be endogenous or exogenous (excluded instruments). We examine each case separately below.

Regarding omitted endogenous variables, one obvious candidate is corruption. Our theoretical model suggests that corruption has no direct effect on performance, although there is ample empirical evidence that reports a causal link from the former to the latter (Dimant & Tosato, 2017). Public sector performance may also influence transparency if public officials having satisfied with the quality of their work are more inclined to open up and be transparent with the inner working of their agencies. The opposite may also be true, i.e., public agencies may be reluctant to be transparent if their performance is of poor quality. Hence, to address such feedback loops we re-estimate our model after adding corruption to the performance equation and performance to the transparency equation.

The second type of omitted variable problem is regarding the exogenous variables, i.e., exclusion restrictions. For example, it is possible that autonomy affects performance and hence it should not be excluded from the performance equation. In the previous section, we tested the validity of instruments (overidentifying restrictions) using Sargan–Hansen test and failed to reject the hypotheses that a variable was excluded from a given equation correctly. The problem with testing instrument exogeneity is that, unlike instrument relevance, it is virtually impossible to test it since the error term is unobservable. Researchers often use tests of overidentifying restrictions to assess the validity of the orthogonality conditions (moment conditions). However,

<sup>9</sup>We use the C statistic (also known as a “GMM distance” or “difference-in-Sargan” statistic) to test the exogeneity of exogenous variables. It is defined as the difference of the Sargan–Hansen statistic of the equation with the smaller set of instruments and that of the equation with the full set of instruments. The smaller set of instruments is valid under both the null and alternative hypotheses, whereas the full set of instruments is valid only under both the null hypotheses. Hence, if the difference-in-Sargan statistic is “large”, this indicates that the instruments in question are likely to be endogenous. The test statistic is distributed as chi-squared in the number of instruments tested.

they are not equivalent, the validity of the former is neither sufficient nor necessary for the validity of the latter. (Parente & Silva, 2012). Hence, as a robustness exercise, we try alternative models where excluded instruments are added to the model sequentially to check whether our conclusion would change.

The results are reported in Table 6–8 for performance, corruption, and transparency equations, respectively, along with the base results in the first columns of each table.

In the performance equation, we first add corruption as an explanatory (endogenous) variable. We test the endogeneity of corruption using difference-in-Sargan C statistic and find that corruption is an endogenous variable. Thus, we use the excluded instruments of the equation (autonomy, honesty, and wages) for IV estimation of corruption via the CUE method. Its coefficient is significant (albeit at 10%) with a negative sign, suggesting that corruption, in fact, has a direct influence on performance. Hence, contrary to our model, even after controlling for other factors corruption is likely to hamper public services. Partial F and partial  $R^2$  statistics indicate that our instruments in the performance equation are weakly associated with corruption, partly explaining the borderline significance of that variable. Note that transparency still has a significant coefficient, but its magnitude has diminished, as some part of its explanatory power is shared by the newly added variable, corruption. In the next two columns, we add two of our

**Table 6.** Alternative models for performance.

|   | BASE MODEL         | ALT. MODEL 1                                 | ALT. MODEL 2        | ALT. MODEL 3        |
|---|--------------------|--|---------------------|---------------------|
| Corruption                              |                    | −0.345*<br>(−1.85)                           |                     |                     |
| Transparency                            | 0.259**<br>(2.41)  | 0.217**<br>(2.04)                            | 0.269**<br>(2.20)   | 0.326*<br>(1.79)    |
| Politicization                          | −0.114*<br>(−1.79) | −0.093<br>(−1.31)                            | −0.138**<br>(−2.07) | −0.129**<br>(−2.02) |
| Enforcement                             | 0.054<br>(1.03)    | 0.048<br>(0.90)                              | 0.066<br>(1.04)     | 0.062<br>(1.11)     |
| Meritocracy                             | 0.166<br>(1.08)    | 0.192<br>(1.30)                              | 0.090<br>(0.65)     | 0.109<br>(0.83)     |
| Voice                                   | 0.457***<br>(3.18) | 0.482***<br>(3.76)                           | 0.493***<br>(3.90)  | 0.506***<br>(3.50)  |
| Decentralization                        | −0.012<br>(−0.24)  | −0.008<br>(−0.19)                            | −0.021<br>(−1.29)   | −0.018<br>(−0.97)   |
| Education                               | 0.158**<br>(2.16)  | 0.119*<br>(1.90)                             | 0.103*<br>(1.85)    | 0.171**<br>(2.45)   |
| Resources                               | 0.293**<br>(2.41)  | 0.327***<br>(2.84)                           | 0.292***<br>(2.87)  | 0.306***<br>(2.88)  |
| Wages                                   |                    |  | 0.085<br>(1.09)     |                     |
| Autonomy                                |                    |  |                     | −0.059<br>(−1.11)   |
| $R^2$                                   | 0.620              | 0.638  | 0.629               | 0.626               |
| Kleibergen–Paap rk LM statistic         | 10.835 (0.01)      | 7.141 (0.03)                                 | 9.342 (0.01)        | 7.764 (0.06)        |
| Partial F statistics in the first stage | 12.651††           | Corruption: 5.285†<br>Transp: 11.957††       | 12.886††            | 6.245†              |
| Partial $R^2$ in the first stage        | 0.209              | Corruption: 0.123<br>Transp: 0.212           | 0.171               | 0.130               |
| Cragg–Donald Wald statistic             | 7.217††            | 5.254†                                       | 7.737††             | 3.039               |
| Kleibergen–Paap rk Wald F statistic     | 12.651††           | 8.297††                                      | 12.886††            | 5.145†              |
| Sargan–Hansen J-statistic               | 1.687 (0.43)       | 1.002 (0.31)                                 | 0.319 (0.57)        | 0.459 (0.50)        |
| Difference-in-Sargan C statistic        | 5.239 (0.02)       | Corrup: 3.758 (0.06)<br>Transp: 6.175 (0.03) | 5.167 (0.02)        | 5.573 (0.01)        |

Notes: Models are estimated using CUE. Sample size is 89. All variables lie in the [0,1] interval. P-values of the coefficient estimates are in parenthesis. \*\*\* Significant at 1%, \*\* significant at 5%, \* significant at 10%. Stock–Yogo critical values for partial F statistics are 6.46 for 10% and 4.36 for 15% maximal relative bias. Stock–Yogo critical values for weak identification tests (used for Cragg–Donald Wald and are Kleibergen–Paap rk Wald F statistics) are 5.44 for 10% and 3.81 for 15% maximal relative bias. Stock–Yogo critical values are for single endogenous regressor, and they constitute more conservative thresholds (higher than they should be) in the case of two endogenous variables. †† denotes significance at 10% and † denotes significance at 15% according to Stock–Yogo critical values.

**Table 7.** Alternative models for corruption.

|   | BASE MODEL                          | ALT. MODEL 1                      | ALT. MODEL 2                     |
|---|-------------------------------------|-----------------------------------|----------------------------------|
| Performance                               | -0.085<br>(-0.72)                   | -0.028<br>(-0.50)                 | -0.012<br>(-0.26)                |
| Transparency                              | -0.312***<br>(-3.19)                | -0.275***<br>(-2.81)              | -0.378<br>(-1.41)                |
| Politicization                            | 0.267***<br>(2.84)                  | 0.261***<br>(2.89)                | 0.290***<br>(3.03)               |
| Enforcement                               | -0.024<br>(-0.21)                   | -0.016<br>(-0.20)                 | -0.038<br>(-0.70)                |
| Meritocracy                               | -0.310***<br>(-2.72)                | -0.319***<br>(-2.84)              | -0.337***<br>(-2.89)             |
| Voice                                     | -0.397**<br>(-2.03)                 | -0.390**<br>(-2.04)               | -0.411**<br>(-2.57)              |
| Decentralization                          | 0.036*<br>(1.87)                    | 0.036*<br>(1.89)                  | 0.037*<br>(1.91)                 |
| Honesty                                   | -0.098<br>(-1.24)                   | -0.079<br>(-1.40)                 | -0.072<br>(-1.17)                |
| Wages                                     | -0.152**<br>(-2.11)                 | -0.153**<br>(-2.29)               | -0.168**<br>(-2.43)              |
| Education                                 |                                     | -0.025<br>(-0.77)                 |                                  |
| Autonomy                                  |                                     |                                   | -0.035<br>(-1.08)                |
| R2  | 0.515                               | 0.517                             | 0.528                            |
| Kleibergen–Paap rk LM statistic           | 8.676 (0.01)                        | 5.007 (0.03)                      | 6.920 (0.01)                     |
| Partial F statistics in the first stage   | Perform: 5.591††<br>Transp: 9.602†† | Perform: 4.734†<br>Transp: 6.108† | Perform: 5.006†<br>Transp: 4.200 |
| Partial R <sup>2</sup> in the first stage | Perform: 0.186<br>Transp: 0.180     | Perform: 0.120<br>Transp: 0.147   | Perform: 0.161<br>Transp: 0.105  |
| Cragg–Donald Wald statistic               | 5.322†                              | 4.176†                            | 3.520                            |
| Kleibergen–Paap rk Wald F statistic       | 10.575††                            | 7.306††                           | 5.273†                           |
| Sargan–Hansen J-statistic                 | 0.006 (0.94)                        | -                                 | -                                |

Notes: Models are estimated using CUE. Sample size is 89. All variables lie in the [0,1] interval. P-values of the coefficient estimates are in parenthesis. \*\*\* Significant at 1%, \*\* significant at 5%, \* significant at 10%. Stock–Yogo critical values for partial F statistics are 6.46 for 10% and 4.36 for 15% maximal relative bias. Stock–Yogo critical values for weak identification tests (used for Cragg–Donald Wald and are Kleibergen–Paap rk Wald F statistics) are 5.44 for 10% and 3.81 for 15% maximal relative bias. Stock–Yogo critical values are for single endogenous regressor, and they constitute more conservative thresholds (higher than they should be) in the case of two endogenous variables. †† denotes significance at 10% and † denotes significance at 15% according to Stock–Yogo critical values.

excluded instruments, wages and autonomy sequentially to the model. It is possible that public agencies with better pay or autonomous structures may perform better. However, in both cases the coefficients of these variables are insignificant. As expected, excluding these variables from the instrument set leads to less precise estimates for the endogenous variable (transparency). Losing autonomy as an instrument, in particular, pushes transparency to borderline significance as the remaining instruments have a rather weak connection as demonstrated by the partial F and R<sup>2</sup> statistics. Moreover, Stock and Yogi weak identification test statistic is below the critical value 5.44 associated with 10% bias, suggesting weakness in instruments.

In the corruption equation, we move education and autonomy one by one from the list of excluded instruments to the list of included regressors. It is possible that public agencies with more educated employees or more autonomous structures suffer from less corruption. We find that the coefficients of these variables are rather insignificant. We do not observe any noteworthy change in the coefficients of endogenous variables (performance and transparency) when education is added to the model. However, when autonomy becomes an included regressor, the coefficient of transparency loses its significance, and the test statistics indicate that the remaining instruments (education and resources) have a rather weak association with the endogenous regressor (transparency). Partial F statistic points to a relative IV bias higher than 15% for transparency. Stock and Yogo weak identification test statistics are also quite low.



**Table 8.** Alternative models for transparency.

|   | BASE MODEL           | ALT. MODEL 1                                  | ALT. MODEL 2         | ALT. MODEL 3         |
|---|----------------------|---|----------------------|----------------------|
| Performance                               |                      | 0.321**<br>(2.50)                             |                      |                      |
| Corruption                                | −0.302**<br>(−2.60)  | −0.187**<br>(−2.12)                           | −0.514<br>(−1.47)    | −0.284**<br>(−2.52)  |
| Politicization                            | −0.235**<br>(−2.04)  | −0.178*<br>(−1.83)                            | −0.247**<br>(−2.52)  | −0.248**<br>(−2.30)  |
| Enforcement                               | 0.173<br>(1.28)      | 0.193<br>(1.42)                               | 0.197*<br>(1.88)     | 0.201*<br>(1.87)     |
| Meritocracy                               | 0.119<br>(1.03)      | 0.249*<br>(1.79)                              | 0.185<br>(1.57)      | 0.126<br>(1.40)      |
| Voice                                     | 0.541**<br>(2.50)    | 0.699***<br>(2.75)                            | 0.599***<br>(3.13)   | 0.522**<br>(2.37)    |
| Decentralization                          | 0.019<br>(1.12)      | 0.032*<br>(1.76)                              | 0.008<br>(0.60)      | 0.013<br>(0.94)      |
| Education                                 | 0.134*<br>(1.77)     | 0.130*<br>(1.79)                              | 0.158**<br>(2.34)    | 0.117<br>(1.59)      |
| Autonomy                                  | −0.107***<br>(−2.81) | −0.127***<br>(−3.45)                          | −0.125***<br>(−3.20) | −0.103***<br>(−3.08) |
| Wages                                     |                      |   | 0.073<br>(1.47)      |                      |
| Honesty                                   |                      |   |                      | 0.044<br>(1.18)      |
| R <sup>2</sup>                            | 0.541                | 0.578   | 0.545                | 0.547                |
| Kleibergen-Paap rk LM statistic           | 7.973 (0.05)         | 6.662 (0.04)                                  | 5.026 (0.08)         | 6.744 (0.03)         |
| Partial F statistics in the first stage   | 5.371†               | Performance: 4.987†<br>Corruption: 5.308†     | 3.128                | 5.398†               |
| Partial R <sup>2</sup> in the first stage | 0.128                | Performance: 0.107<br>Corruption: 0.155       | 0.069                | 0.145                |
| Cragg-Donald Wald statistic               | 4.293†               | 4.008†  | 2.473                | 4.201†               |
| Kleibergen-Paap rk Wald F statistic       | 5.371†               | 5.917††                                       | 3.128                | 5.398†               |
| Sargan-Hansen J-statistic                 | 2.459 (0.29)         | 0.775 (0.38)                                  | 1.688 (0.19)         | 2.422 (0.12)         |
| Difference-in-Sargan C statistic          | 8.160 (0.00)         | Perfor: 3.219 (0.07)<br>Corrupt: 7.025 (0.00) | 8.744 (0.00)         | 7.805 (0.00)         |

Notes: Models are estimated using CUE. Sample size is 89. All variables lie in the [0,1] interval. P-values of the coefficient estimates are in parenthesis. \*\*\* Significant at 1%, \*\* significant at 5%, \* significant at 10%. Stock-Yogo critical values for partial F statistics are 6.46 for 10% and 4.36 for 15% maximal relative bias. Stock-Yogo critical values for weak identification tests (used for Cragg-Donald Wald and are Kleibergen-Paap rk Wald F statistics) are 5.44 for 10% and 3.81 for 15% maximal relative bias. Stock-Yogo critical values are for single endogenous regressor, and they constitute more conservative thresholds (higher than they should be) in the case of two endogenous variables. †† denotes significance at 10% and † denotes significance at 15% according to Stock-Yogo critical values.

In the transparency equation, we first add performance as an explanatory (endogenous) variable to the model. Difference-in-Sargan C statistic suggests that this variable is close to being endogenous. Hence, we use the excluded instruments of the equation (resources, honesty, and wages) to estimate performance. Its coefficient in the CUE is highly significant suggesting that better performance leads to more transparency. The coefficient of corruption diminishes, but it is still significant as well. Then, we change our instrument set and add wages and honesty to the equation one by one. In the first case, the coefficient of corruption loses its precision as its standard deviation increases considerably. Test statistics suggest that the remaining instruments are rather weak, and the estimates are likely to suffer from weak instrument problem. When we add honesty to the equation, on the other hand, we do not observe any noteworthy change in the estimates. The coefficients of both wages and autonomy are insignificant, suggesting that there is no direct link from these variables to the dependent variable.

Overall, we do not find any evidence that excluded regressors that are used as instruments for endogenous variables are excluded from the associated equation incorrectly. We do find, however, from an empirical point of view, it would be appropriate to include corruption in the performance equation and performance in the transparency equation as endogenous variables, contrary to the implications of our theoretical model.