

Accepted Manuscript

Corruption causes inequality, or is it the other way around? An empirical investigation for a panel of countries

Laura Policardo, Edgar J. Sánchez Carrera

PII: S0313-5926(17)30227-8
DOI: <https://doi.org/10.1016/j.eap.2018.05.001>
Reference: EAP 225

To appear in: *Economic Analysis and Policy*

Received date : 19 October 2017

Revised date : 13 May 2018

Accepted date : 14 May 2018

Please cite this article as: Policardo L., Carrera E.J.S., Corruption causes inequality, or is it the other way around? An empirical investigation for a panel of countries. *Economic Analysis and Policy* (2018), <https://doi.org/10.1016/j.eap.2018.05.001>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Corruption causes inequality, or is it the other way around?

An empirical investigation for a panel of countries

Laura Policardo*

Edgar J. Sánchez Carrera[†]

Abstract

In past decades, a relevant strand of literature has found corruption to be an important determinant of income inequality. In this paper, it is argued that a reverse causal relationship between corruption and income inequality might exist. We claim that income inequality could in fact be responsible for fostering corruption, which may be a reaction to a perceived unfair income distribution. Looking at a panel of 50 countries from 1995 - 2015, we show that the direction of causality between corruption and income inequality is country-specific and may be bidirectional. Using a dynamic GMM model, we robustly find that income inequality positively affects corruption, while corruption does not appear to be significant in the determination of income inequality, therefore contradicting the existing empirical literature on this topic.

Keywords: Corruption perception; dynamic panel; income inequality; causality.

JEL Codes: C23; D30; K42; O1; O5.

*Italian Ministry of Economy and Finance. Via Pietrapiana 53, 50121 Florence, Italy. E-mail: laura.policardo@gmail.com
The opinions expressed in this publication are those of the authors and do not necessarily reflect the official opinion of the Italian Ministry of Economy and Finance.

[†]Department of Economics, Society, and Politics at The University of Urbino Carlo Bo, Italy. E-mail: edgar.sanchezcarrera@uniurb.it

Corruption causes inequality, or is it the other way around?

An empirical investigation for a panel of countries

Abstract

In past decades, a relevant strand of literature has found corruption to be an important determinant of income inequality. In this paper, it is argued that a reverse causal relationship between corruption and income inequality might exist. We claim that income inequality could in fact be responsible for fostering corruption, which may be a reaction to a perceived unfair income distribution. Looking at a panel of 50 countries from 1995 - 2015, we show that the direction of causality between corruption and income inequality is country-specific and may be bidirectional. Using a dynamic GMM model, we robustly find that income inequality positively affects corruption, while corruption does not appear to be significant in the determination of income inequality, therefore contradicting the existing empirical literature on this topic.

Keywords: Corruption perception; dynamic panel; income inequality; causality.

JEL Codes: C23; D30; K42; O1; O5.

1 Introduction

Corruption and its consequences on various economic indicators have been the focus of interest of many scholars during the last twenty years.

Despite arguments in favour of corruption exist, in particular with regard to the ability of corruption to help overcome bureaucratic rigidities and maintain allocation efficiency when there is competition between bribers (Bardhan, 1997), the misuse of public office for private gain in a manner that contravenes the rules of the game (so it is defined corruption) has been found responsible for losses in GDP growth (Mauro 1996; Leite and Weideman 1999; Tanzi and Davoodi 2000, Abed and Davoodi 2000), in the ratio of investments to GDP (Mauro 1996; Ades and Di Tella 1997; Tanzi and Davoodi 1997), in the ratio of public education and public health spending to GDP (Mauro 1998), in the ratio of tax revenues to GDP (Ghura 1998), in some measures of government revenues to GDP ratio (Tanzi and Davoodi 2000), in the level of inflation (Al-Marhubi, 2000) and finally in the amount of foreign direct investment (Habib and Zurawicki 2001, Dutta et al. 2017)¹.

In this paper, we investigate the effect of income inequality on corruption and their causal relationship. The existing empirical literature in economics and political sciences is indeed mainly focused at studying whether corruption may affect income inequality, but very few contributions consider the possibility that the reverse might be true as well, with the consequence of putting at serious risk of biasedness all empirical investigations which do not handle this potential problem of endogeneity.

On one hand, corruption may affect income inequality in several ways, directly, through biased tax systems favouring the wealth and well connected (Gupta et al, 2002; Dincer and Gunalp, 2005; Gymah-Brempong, 2002; Ullah and Ahmad, 2016), or indirectly, by means of other variables which, in turn, affect inequality, like for example, the level and effectiveness of social spending (Olken, 2007; Suryadarma, 2012) and unequal access to education and public services, with particular regard to health services (Azfar and Gurgur, 2005; Gupta et al., 2000). However, the debate on how corruption affects inequality comes often to very different conclusions and it is therefore inconclusive: Gupta et al. (2002), for example, show that high and rising corruption increases income inequality and poverty. They conclude that policies that reduce corruption will most likely reduce income inequality and poverty as well, while Andres and Ramlogan-Dobson (2011) show that lower corruption is associated with higher inequality in Latin America countries,

¹For a systematic review of the argument, see Akçay (2006).

and other scholars indeed find non-linearities in the corruption-inequality relationship (Fakir et al. 2017).

Several might be the causes of corruption. The main econometric literature finds protestant traditions, histories of British rule, income and (maybe) higher imports, and long exposure to democracy (but not actual levels of democracy) associated to lower levels of corruption (Treisman, 2000; Nur-Tegin and Czap, 2012). Inflation has instead been found positively correlated with corruption (Paldam, 2002), so as the membership to the OPEC (Montinola and Jackman, 2002). Moreover, freedom of the press, economic freedom, government decentralization and property rights are associated to lower levels of corruption (Dimant and Tosato, 2017, Saha and Ben Ali, 2017, Saha and Su, 2012). Transparency International indicates that possible simultaneous factors affecting corruption are due to: i) personal greed together with the decline of ethical sensitivity when working in public and private institutions, ii) inefficient regulations and controls, iii) lack of moral criteria in promotions (wages) on workplaces, and iv) lack of transparency at the institutional levels (Van Rijckeghem and Weder, 1997). We believe that several of the causes mentioned above (especially personal greed) might be fostered by income inequality: Champernowne and Cowell (1998) indeed claim that

“The most powerful motive for condemning inequality as unjust is a personal one: the resentful envy aroused by the spectacle of a wealthy few enjoying ease whilst one’s own kind have to toil and put up with constant inconvenience and hardship [...] (pag. 7)”.

That’s why we investigate causality between those two variables, because corruption may be the unjust individual response to a perceived unjust inequality, and the available main empirical contributions address mainly the reverse causal relationship. This fact, moreover, may put at risk of biasedness all the existing empirical investigations that do not consider this possibility, that is to say, that do not consider this potential serious problem of endogeneity between corruption and inequality.

All the economic contributions cited up to now give for granted that corruption is responsible for variations in income inequality, failing to check if the reverse might be true as well. Some attempts to verify whether income inequality may be responsible for variations in the level of corruption have been made in the sociological and recently in the economic literature: You and Khagram (2005) indeed find evidence that inequality may foster corruption through several channels: first of all, inequality affects people’s social norms about corruption and social beliefs about the legitimacy of rules and institutions,

making it easy to tolerate corruption as an acceptable behavior; secondly, the wealthy have both greater motivations and opportunities to engage in corruption, while the poor are more vulnerable to extortion and less able to monitor and hold the rich and powerful accountable. Bosco (2016), for example, finds evidence that social distress and poverty are responsible for increases in the level of corruption at any level of income.

Among scholars who face the problem of studying the causality between corruption and income inequality, Apergis, Dincer and Payne (2010) investigate the causality relationship within a multivariate framework for the U.S. over the period 1980 to 2004. Using cointegration techniques, they detect a long run relationship between corruption and income inequality and a bidirectional causality between those two variables. Chong and Gradstein (2007) investigate theoretically and empirically the relationship between inequality and institutional quality, placing its findings (of a double causality between these two variables) in the context of the conflicting evidence as to how corruption affects inequality.

Our paper is then aimed at proving that inequality may be a cause of corruption, and that the vice versa might not necessarily be true. We try to do that with a panel of 50 countries² which includes developed, developing and underdeveloped nations. The choice of the sample is due to the availability of data. No opportunistic choice of sample was made, except for the fact that we selected only countries that had at least 10 years during 1995-2015 with income inequality and corruption perception index simultaneously observed. This of course does not guarantee that in the regressions we'll have at least 500 observations, because other covariates might be unobserved when income inequality and/or corruption are.

This paper is structured as follows: section 2 introduces the data and shows some descriptive statistics. Section 3 presents some diagnostic tests and the econometric model. Section 4 illustrates the results, and section 5 concludes.

2 Dataset and descriptive statistics

The two variables, which are the focus of interest in this paper, are income inequality measured by the Gini index, and the Corruption Perception Index. Both variables are taken from the database of Teorell et

²Our set of countries are: Argentina, Armenia, Austria, Belarus, Belgium, Bolivia, Brazil, Colombia, Costa Rica, Cyprus, Czech Republic, Denmark, Dominican Republic, Ecuador, El Salvador, Estonia, Finland, France, Georgia, Greece, Honduras, Hungary, Iceland, Ireland, Italy, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Luxembourg, Mexico, Moldova, Netherlands, Norway, Panama, Paraguay, Peru, Poland, Portugal, Romania, Russia, Slovakia, Slovenia, Spain, Sweden, Thailand, Turkey, Ukraine, United Kingdom, Uruguay.

al. (2018), along with other covariates used in our regression models. Teorell et al. (2018) built up a large dataset of approximately 2,100 variables, based on a number of freely available data sources.

As previously mentioned, we downloaded from this dataset the indexes of income inequality and CPI, keeping the countries who had at least 10 years with simultaneous observations of Gini and CPI within 1995-2015. The variables downloaded, with the information about their original source are:

1. Income Inequality denoted by *GINI*, measured by the Gini index of households' income inequality³. World Bank, World Development Indicators.
2. Corruption Perception Index, denoted by *CPI*. It was created in 1995 by Transparency International and regularly updated every year. It ranks almost 200 countries on a scale of zero to 100, with zero indicating high levels of corruption and 100 indicating low levels of corruption. Despite the fact that criteria for the determination of the CPI score may vary slightly across countries and years (Lambsdorff J., 2006), we decided anyway to use this index of corruption because firstly it covers a substantial set of countries and years, secondly, because despite these pitfalls, it's been found to be highly correlated between two other proxies for corruption, like black market activity and overabundance of regulation (Wilhelm Paul G., 2002), and third because since it combines survey results from different organizations creating a "poll of polls", it increases reliability of the data by reducing the effect of any biases of individual surveys (Montinola and Jackman, 2002). We decided moreover not to use the International Country Risk Guide (ICRG) corruption index⁴ because, as erroneously thought by many authors, it does not determine a country's level of corruption, but the political risk involved in corruption. There are indeed different levels of tolerance of corruption in different countries: where corruption is more tolerated, the political risk is lower, but this does not mean that there is no corruption, it is just that the population is less upset than elsewhere with this practice. So, for these reasons we believe that CPI score could be the best measure of corruption available at the moment.
3. Real percapita GDP at constant 2005 US dollars. Gleditsch (2002).
4. Human capital index. From Penn World Tables 9.0. In this version of the Penn Tables, this index is a combination of Barro and Lee (2013) and Cohen and Leker (2014) average years of schooling.

³In some cases this index may indicate consumption expenditure inequality.

⁴This index is calculated by the PRS Group, and it easily downloadable at <http://www.prsgroup.com>

5. Government gross debt (expressed as percentage of GDP). International Monetary Fund.
6. Inflation index, expressed in averages for the year. International Monetary Fund.
7. Unemployment rate, expressed as percentage of total labor force. International Monetary Fund.
8. Population (millions). International Monetary Fund.
9. Total fertility rate (births per woman). World Bank, World Development Indicators.
10. Oil production value in 2000 US billion \$. Ross and Mahdawi (2015).
11. Life expectancy (both sexes, age 0-1 years). Institute for Health Metrics and Evaluation (IHME).
12. Unified Democracy Score (posterior mean). Pemstein et al. (2010).

Table 2 introduces the main summary statistics (median, standard deviation, minimum and maximum) for each enumerated variable used in our econometric model.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
year			1995	2015	674
Real Percapita GDP	15284.404	12747.001	162.96	59640.09	507
Life Expectancy	75.434	4.381	62.987	82.845	674
Government Gross Debt	47.299	28.57	3.664	180.938	639
Inflation	114.906	87.606	2.46	931.126	611
Population	28.967	42.087	0.294	204.47	614
Unemployment	8.393	4.589	0.488	31.6	644
Human Capital	2.878	0.45	1.832	3.734	624
Oil production value	8.763	32.395	0	303.861	654
CPI	47.065	21.619	16	97	674
Unified Dem. Score	0.885	0.690	-0.761	2.263	557
Fertility	1.964	0.616	1.126	4.331	674
Income inequality (GINI)	38.374	10.154	23.7	63	674

Figure 1 plots the average CPI score (Corruption perception index) against the average Gini index of income inequality, for the 50 countries included in our sample, during the period 1995 - 2015. Dots in the graph indicate each country's position in the corruption-inequality relationship.

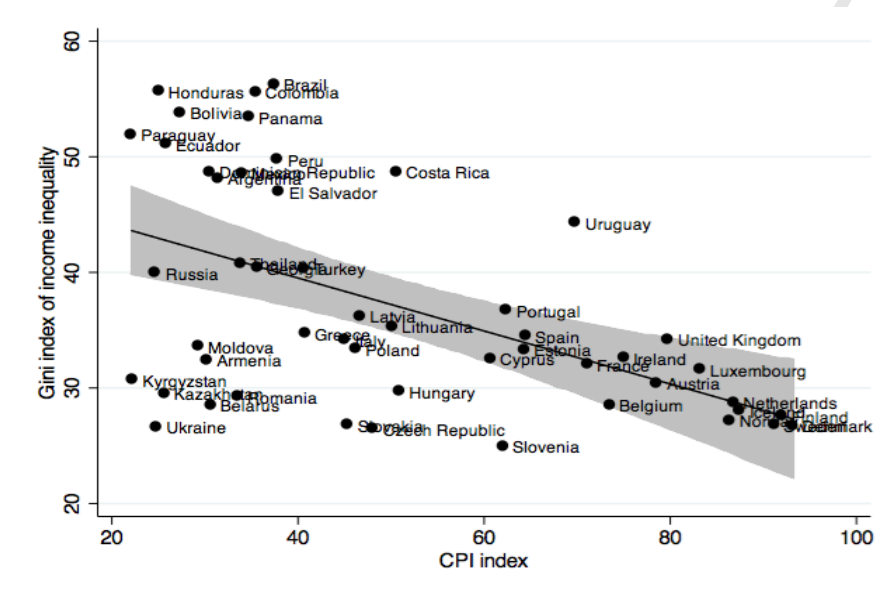


Figure 1: Relationship between average CPI and GINI over the period 1995-2015.

The dark line over the dots in figure 1 represents the estimated coefficient of a cross country regression seeing the mean of Gini index of income inequality for the period 1995-2015 for each country as the dependent variable, and the mean of CPI score as the regressor (plus a constant). The grey shadow over this line represents the 95% confidence interval. Since the CPI score increases as the perceived level of corruption decreases, figure 1 clearly shows that high levels of income inequality are associated to high levels of corruption, or - in this case - perceived corruption. We can notice from this figure that, apparently, at least two clusters exist: the first, represented by the Latin American Countries, and the second, by the ex-socialist countries. Despite the fact that in the first cluster the average level of income inequality is higher, while in the second cluster - for obvious historical reasons - is lower, we can however see that the relation between CPI index (corruption) and income appears to be persistently negative (positive), even though we consider separately those two clusters.

3 Diagnostic tests and model selection

Panel Unit root tests

In order to select the appropriate econometric model, some diagnostic tests are necessary. First of all, we check stationarity of all the variables.

Table 2: Unit root tests

Variable	Levin, Lin and Chu t*	Im, Peasaran and Shin W-stat	ADF - Fisher Chi-square	PP - Fisher Chi-square
CPI	-7.41428***	-2.44343***	150.717***	162.134***
GINI	-10.2036***	-4.19331***	192.276***	186.219***
Real percapita income	-14.2014***	-0.70501	131.652**	165.601***
Human Capital	116.756	-14.9398***	269.979***	288.251***
Government Gross Debt	-6.03951***	-1.08057	138.91***	119.598*
Inflation	-14.6001***	-2.4916***	141.415***	130.09***
Population	-38.3638***	-29.9944***	244.433***	211.187***
Unemployment	-10.9931***	-2.44782***	163.374***	126.643**
Oil production value	-9.20214***	-2.4768***	96.2725***	136.677***
Unified Dem. Score	-136.485***	-12.0935***	237.823***	303.135***
Fertility	-14.3038***	-8.26623***	173.354***	169.463***
Life Expectancy	-34.498***	-7.77843***	172.457***	188.437***

Legend: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Assuming intercept and trend, and selecting the number of lags according to the maximum AIC criteria, Levin, Lin and Chu (2002) test (in the following, LLC) rejects the hypothesis of common autoregressive parameters in all the variables but Human capital index. However, the alternative hypothesis of this test is that none of the groups have a unit root. For this reason, if LLC test for common unit root fails, it might be convenient to check whether individual unit root exists. Tests available for that are Im, Peasaran and Shin (2003), and Fisher AD and PP tests extended on panel datasets by Maddala and Wu (1999) and later by Choi (2001). Im, Peasaran and Shin (2003) test assumes that all units have the same nonstochastic components and runs separate ADF regressions for each unit. Then, it averages the likelihood ratios across units. The null of this test is that all units have unit root, but the alternative is that some units might have unit root. Fisher AD and PP tests have the same null and alternative as IPS, but it was calculated combining p-values of the single UR tests performed on each unit.

All the tests performed agree upon the fact that our corruption perception index and our variable representing income inequality are stationary. Human capital appears to be stationary, too, since at the end, three out of four tests for unit root reject the null of UR. Whether or not the variable representing Government indebtedness was stationary is not clear, since Im, Peasaran and Shin test finds it integrated of order 1 (our calculations prove that unit root tests performed on this variable differenced once show stationarity). Since this variable will be used as covariate only, this does not represent a problem, also because, as observed for human capital, in this case too three out of four tests suggest stationarity, and therefore we will assume that.

Panel Granger Causality test

Most of the literature available takes for granted that is corruption only that affects income inequality, and does not take into consideration the possibility that income inequality might fostering corruption too. The lack of this important consideration makes all the available empirical results at risk of biasedness or misrepresentation, because of this serious problem of endogeneity between income inequality and corruption, which is potentially left unexplored.

Therefore, an important question that must be addressed is the causality direction between inequality and corruption: is the inequality in income distribution generating more corruption or, on the contrary, corruption increasing inequality? In this section we try to address this question by applying two different Granger causality tests for panel data. There are a number of different approaches to test causality in a panel context. Most of them differ on the assumptions about homogeneity/heterogeneity of the coefficients across countries.

Our aim is to investigate the causality relationship between inequality and corruption using two different approaches:

1. The first is to treat the panel data as one large stacked set of data, and then perform the Granger Causality test in the standard way, that is to say by regressing one lag of the variable *GINI* and one lag of the variable *CPI* over actual values of *CPI* and then over actual values of *GINI*, and a dummy for each of the country included in the dataset. If the lagged value of *GINI* significantly affects actual values of *CPI* then we say that *GINI* causes *CPI*, while if the lagged value of *CPI* significantly affects actual values of *GINI*, then we say that *CPI* causes *GINI*. This method assumes that all

coefficients are same across all cross-sections, except for the individual constant;

2. A second approach is suggested by Dumitrescu and Hurlin (2012). They propose an extreme opposite assumption, allowing all coefficients to be different across cross-sections. Dumitrescu and Hurlin's simple test of homogeneous non-causality assumes, under the null, that there is no causal relationship for any of the units of the panel and considers a heterogeneous panel data model with fixed coefficients (in time). It also specifies the alternative hypothesis as heterogeneous causality, which assumes that there is a causal relationship from x to y for a subgroup of individuals only (possibly, all).

Table 3 shows the results for the first test of the Granger causality when homogeneity of coefficients is assumed among countries.

Table 3: Causality test assuming homogeneity of coefficients. The t-statistics are relative to the coefficient attached to one lag of CPI for the null a) and one lag of GINI for the null b).

Null Hypothesis:	t-statistics	Prob.
a) CPI does not Granger Cause GINI	-.0167845	0.224
b) GINI does not Granger Cause CPI	-.1426596	0.007

Source: own calculations. A p-value smaller than 0.05 implies a rejection of the null hypothesis

Note that the hypothesis that CPI does not Granger-cause GINI is rejected at 1%. It means that there is a direction of causality from corruption to income distribution.

Dumitrescu and Hurlin (2012) pointed out that imposing homogeneity of coefficients β_i , the causality test-statistic leads to fallacious inference, since it was evident that an homogeneous specification of the relation between the variables does not allow to interpret the causality relation when economic behavior differs between countries. For instance, assume that there is a causality relation for a set of N countries, for which the coefficients β_i are strictly identical. What conclusion can be drawn when introducing into the sample a set of N_1 countries for which, in contrast, there is no relation of causality? They conclude that whatever the value of the ratio N/N_1 is, the test of the causality is nonsensical. Therefore, they propose to test the Homogeneous Non Causality (HNC) hypothesis by taking into account both the heterogeneity of the regression model and that of the causal relation. Table 3 shows the results of the test according to Dumitrescu and Hurlin's hypothesis.

Dumitrescu and Hurlin's test clearly suggests that –assuming heterogeneity of the coefficients between countries– corruption causes income inequality in at least one country, and also income inequality causes

Table 4: Causality test assuming heterogeneity of coefficients.

Null Hypothesis:	W-Stat.	Zbar-Stat.	Prob.
CPI does not homegeneously cause GINI	2.12084	2.29616	0.0217
GINI does not homogeneously cause CPI	2.51388	3.41022	0.0006

Source: Own calculations. One lag was considered. A p-value smaller than 0.05 implies a rejection of the null hypothesis

corruption in at least one country, among the 50 analyzed, so there is a bidirectional causality. However, the significance is stronger in the GINI \rightarrow CPI direction, because its p-value is 0.0006 against a p-value of 0.0217 of the CPI \rightarrow GINI direction. This may imply that in an eventual regression where GINI is the dependent variable, the significance of CPI might be weaker with respect the significance of GINI in a regression where CPI is considered as a dependent variable.

Econometric model

In the present study we consider a dynamic panel data model, in the sense that it contains (at least) one lagged dependent variable. We apply the Arellano and Bond(1991) generalized method of moments (GMM) approach to estimate the relationship between GINI and CPI. They suggest estimating equation 1 and 2 in first differences, because this allows to remove the potential endogeneity problem due to the omission of country-specific effects, and use a two-step procedure (using as instruments the second and third lag of the dependent variable) since one way fixed effect model of Dynamic panel data tend to generate correlation between regressors and the error (Nickell, 1981). This procedure guarantees (as our tests will show) that the residuals are independent and identically distributed and therefore our instruments are highly correlated with the lagged dependent variable (and its difference) but uncorrelated with the composite error process.

In practice we specify the functional form of the equations for GINI and CPI as follows:

$$GINI_{i,t} = \alpha_{i1} + \beta_{11}GINI_{i,t-1} + \beta_{21}CPI_{i,t} + \beta_{*1}X_{i,t} + u_i + \epsilon_{1i,t}, \quad (1)$$

$$CPI_{i,t} = \alpha_{i2} + \beta_{12}CPI_{i,t-1} + \beta_{22}GINI_{i,t} + \beta_{*2}X_{i,t} + u_i + \epsilon_{2i,t}. \quad (2)$$

where β_{*1} and β_{*2} are vectors of coefficients and $X_{i,t}$ is a matrix of covariates which, for equation 1, includes

real per capita GDP, along with its square value, human capital index and oil production value. In order to check for robustness of our results, we also add, one by one to the matrix of regressors, fertility, life expectancy, inflation and unemployment. For equation 2 instead, our matrix of covariates is represented by real per capita GDP, human capital, government gross debt, oil production value, population and a trend. Here, too, in order to check for robustness of our results, we add two variables, that is to say an index of democracy (represented here by the unified democracy score) and life expectancy.

The reason why we choose to include income and income square among the covariates for equation 1 is due to the potential existence of the so-called Kuznets (1955) curve, while the inclusion of human capital for explaining variations in income inequality is due to the fact that is commonly found in the literature that human capital, intended as scholarization, is an important determinant for reduction in income inequality (Glomm and Ravikumar, 1992). Abundance in natural resources is instead found negatively correlated with many desirable indicators of human development (the so-called resource curse), like economic growth, and life expectancy, and positively correlated with income inequality, at least for economies highly based on natural resources exploitation (Parcero and Papyrakis, 2016). As additional regressors for robustness check we add, in equation 1, fertility, since it may be responsible for increases in income inequality (De la Croix and Doepke, 2003), life expectancy, inflation (which may exacerbate inequality because it decreases the purchasing capacity of the weaker population strata), and finally unemployment, which usually hits people less educated, and therefore it may increase inequality.

Among the determinants of corruption in equation 2 we find, instead, real GDP per capita, human capital, income inequality, a measure of the extent of government in the economic activity (here, gross government debt), a value of oil production, which, in this case too, is often correlated negatively with many indicators of human development (potentially corruption, then), population, and a time trend.⁵ For robustness check, we also add progressively and one by one a variable of life expectancy and an indicator of quality of democracy.

Taking first differences, these two models become:

$$\Delta GINI_{it} = \beta_{11}\Delta GINI_{it-1} + \beta_{21}\Delta CPI_{i,t} + \beta_{*1}\Delta X_{i,t} + \epsilon_{1i,t} \quad (3)$$

⁵Gross government debt is included in this regression because it is reasonable to assume that the more resources a government handles, the more incentives for corruption arises.

$$\Delta CPI_{it} = \beta_{12}\Delta CPI_{i,t-1} + \beta_{22}\Delta GINI_{i,t} + \beta_{*2}\Delta X_{i,t} + \epsilon_{2i,t} \quad (4)$$

and these are the models that will be estimated later. As previously said, we use a two-step GMM estimator where, as instruments for the lagged dependent variable, we use the second and third lag of the same dependent variable. As it is possible to notice, models for GINI and CPI differ slightly, because it is reasonable to assume that the data generating processes and the variables affecting them are different. In the equation for CPI indeed we include a trend (as it is clear in equations 2 and 4). For GINI the inclusion of a trend does not seem to have a relevant effect, and it was not considered in the model. Moreover, while we consider GDP squared as a regressor in the equation determining the level of inequality, therefore allowing for a potential Kuznets effect (equation 1 and 3), in the equation for corruption per capita GDP was left linear (equation 2 and 4). Apart from these details, the two models are similar, because the dependent variables have been regressed over one lag of the dependent variable, over income, over the level of human capital, the oil production value and finally over GINI (CPI) in the equation for corruption (income inequality).

4 Estimation results

This subsection presents the estimation results of equation (3) and (4) applying dynamic panel data analysis (Arellano-Bond based on two-step GMM-Differenced estimator with robust standard errors). Results are presented in Tables 5 and 6.

Table 5 shows that the level of income inequality depends positively on its past value. Sargan (1975) test does not reject the null of validity of over-identifying restrictions, that is to say that instruments used are valid instruments, uncorrelated with the error term, and the excluded instruments are correctly excluded from the estimated equations. Moreover, no serial correlation in the residuals is detected. Both Arellano Bond AR(1) and AR(2) tests for serial autocorrelation in first differenced residuals indeed accept at the standard 5% level the null of no first-order and second-order autocorrelation for all the five models ran. No cross-sectional dependency is moreover detected since Pesaran (2004, 2015) CD test for cross-sectional dependence accepts at the standard 5% level the null of independency among cross-sections (in this case, countries). Variations in real per capita GDP and its square value seem to affect significantly the dynamics of income inequality, in a way that resembles the dynamics of the so-called Kuznets curve (an inverse u-

Table 5: GMM estimation of equation (3). Dependent variable: GINI

	Model 1	Model 2	Model 3	Model 4	Model 5
$GINI_{i,t-1}$	0.0569 (0.51)	0.0967 (0.75)	0.0372 (0.21)	0.166 (1.12)	0.182 (1.13)
$CPI_{i,t}$	-0.0187 (-0.61)	-0.0136 (-0.39)	-0.0107 (-0.31)	-0.0287 (-0.74)	-0.0195 (-0.51)
Real percapita $GDP_{i,t}$	-0.000503** (-2.65)	-0.000548** (-2.83)	-0.000532* (-2.54)	-0.000564** (-2.63)	-0.000172 (-0.62)
Real percapita $GDP^2_{i,t}$	5.84e-09* (2.06)	6.07e-09* (2.38)	5.49e-09* (2.10)	5.49e-09* (2.11)	1.56e-09 (0.47)
Human Capital Index $_{i,t}$	-12.54*** (-4.04)	-11.01** (-2.90)	-9.612* (-2.16)	-8.337* (-2.00)	-7.888 (-1.62)
Oil prod. value $_{i,t}$	0.00982 (1.72)	0.0100 (1.89)	0.0121** (2.77)	0.0108* (2.11)	0.00951* (2.07)
Fertility $_{i,t}$		2.337 (1.61)	1.925 (1.24)	2.461 (1.35)	1.561 (0.93)
Life expectancy $_{i,t}$			-0.271 (-1.32)	-0.254 (-1.21)	-0.536* (-2.29)
Inflation $_{i,t}$				0.00379* (2.09)	0.00440* (1.99)
Unemployment $_{i,t}$					0.170* (2.23)
_cons	79.11*** (6.30)	68.85*** (4.53)	87.81*** (3.43)	78.10** (3.21)	92.32*** (4.01)
N	354	354	354	323	319
Sargan test	29.11 [p=0.4070]	30.14 [p=0.3565]	32.27 [p=0.2636]	32.94 [p=0.1990]	32.15 [p=0.2264]
Wald spec. test	57.44 [p=0.000]	45.89 [p=0.000]	53.67 [p=0.000]	88.16 [p=0.000]	149.40 [p=0.000]
AB test for no autocorr. in first diff. errors					
Order 1:	-1.5177 [p=0.1291]	-1.6073 [p=0.1080]	-.94436 [p=0.3450]	-1.8021 [p=0.0715]	-1.8784 [p=0.0603]
Order 2:	-.09433 [p=0.9248]	-.06454 [p=0.9485]	-.29987 [p=0.7643]	-.14026 [p=0.8885]	-.45071 [p=0.6522]
CD test for cross - sectional independence	-1.875 [p=0.061]	-1.228 [p=0.220]	-.927 [p=0.354]	-.531 [p=0.595]	-.267 [p=0.789]

t statistics in parentheses; p-values in []
* ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$)

shape) for models 1-4. However, this result does not appear to be robust with respect to the simultaneous inclusion of other covariates, namely, fertility, life expectancy, inflation and unemployment. In model 5 indeed neither GDP nor GDP square appear to be significant, therefore letting the reader think that other factors may possibly determine the dynamics of income inequality (see, for example, Policardo et al., 2016). Variations in the level of scholarization (human capital index) appear to affect negatively and significantly the dynamics of income inequality in models 1-4. An increase in the level of human capital index is indeed associated to a decrease in income inequality. However, in model 5 this variable loses its significance, letting other factors determine the variation in income distribution. The value of oil production increases income inequality in models 3, 4 and 5, and therefore, even though this result is not robust, is coherent with the common view in the literature according to which countries with natural resources tend to perform worse in terms of human and social development.

Fertility does not seem to affect income inequality, while inflation and unemployment are responsible for increasing income inequality because they reasonably affect the strata of the population which is weaker, with limited possibilities to increase their incomes as a consequence of rampant inflation or with higher probability to lose their job in case of recession (i.e. the people that are often more affected by unemployment are usually the less educated and specialized).

It is worth noticing that in this equation, the corruption perception index is not significant for the determination of income inequality in all the 5 models. This is in line with results obtained with our tests of causality, since the standard Granger causality test that assumes homogeneity of coefficients finds the direction of causality from GINI to CPI and not vice versa. Dumitrescu and Hurlin's version of Granger test, despite the fact it shows that there is at least in one country the direction of causality from CPI to GINI, its statistical significance is weaker with respect to the opposite direction of causality, from GINI to CPI for at least one country.

Table 6 presents the estimated coefficients of equation 4. Corruption perception is persistent through time, since CPI lagged once and twice are significant at 5% level in all the three models.

In our sample, real percapita GDP appears to decrease the perceived corruption⁶. Basically, what is often found in the literature is a negative relationship between economic development (expressed by GDP)

⁶It can be useful to remind that CPI index is high for countries with low levels of perceived corruption, while it is low for countries with high values of perceived corruption. This means that a positive coefficient in the regression implies an increase in this variable (and therefore a reduction of corruption) as a consequence of an increase in the respective covariate.

Table 6: GMM estimation of equation (4). Dependent variable: CPI

	Model 1	Model 2	Model 3
$CPI_{i,t-1}$	0.246** (2.83)	0.242** (2.82)	0.240** (2.79)
[1em] Real percapita $GDP_{i,t}$	0.000375** (3.28)	0.000384*** (3.49)	0.000392*** (3.35)
$GINI_{i,t}$	-0.296** (-2.77)	-0.304** (-2.96)	-0.308** (-3.03)
Human Capital $_{i,t}$	-19.71 (-1.78)	-18.77 (-1.65)	-18.52 (-1.68)
Government Gross Debt $_{i,t}$	-0.0560*** (-4.46)	-0.0562*** (-4.40)	-0.0579*** (-4.67)
Oil prod. value $_{i,t}$	-0.0191** (-2.92)	-0.0195** (-3.01)	-0.0195** (-3.13)
Population $_{i,t}$	-0.0156 (-0.09)	-0.0354 (-0.21)	-0.0249 (-0.16)
year $_{i,t}$	0.155 (0.67)	-0.175 (-0.61)	-0.202 (-0.72)
Life Expectancy $_{i,t}$		1.030* (2.09)	1.064* (2.11)
Unified dem. score $_{i,t}$			-0.335 (-0.43)
_cons	-208.2 (-0.47)	375.1 (0.71)	426.0 (0.83)
N	329	329	329
Sargan test	19.17197 [p=0.8637]	19.47316 [p=0.8521]	19.79393 [p=0.8393]
Wald spec. test	83.63 [p=0.000]	94.43 [p=0.000]	98.08 [p=0.000]
AB test for no autocorr. in first diff. errors			
Order 1:	-1.7712 [p=0.0765]	-1.7331 [p=0.0831]	-1.7242 [p=0.0847]
Order 2:	-.13614 [p=0.8917]	-.12785 [p=0.8983]	-.12473 [p=0.9007]
CD test for cross- sectional independence	-.818 [p=0.413]	-1.202 [p=0.229]	-1.284 [p=0.199]

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

and corruption, and our results corroborate these common findings.

Income inequality is instead found to consistently increase corruption; when income inequality is high, i.e. poverty is widespread in the society, several mechanisms may be responsible for increasing corruption: the first is the increase of petty corruption due to the fact that poor people, in pursuing wealth, engage more often in unfair and illegal activities, as a possible consequence of a common adverse influence of social norms about corruption and beliefs about the legitimacy of rules and institutions (You and Khagram, 2005). Another way through which income inequality may be responsible for increases in corruption is that the wealthy have both greater motivation and opportunity to engage in corruption, while the poor are more vulnerable to extortion and less able to monitor and hold the rich and powerful accountable (You and Khagram, 2005).

In this equation, human capital seems not to affect significantly the level of perceived corruption, while government indebtedness (which is used here as a proxy for government dimension in the economic activity) increases opportunity for corruption and therefore its perception by the public opinion.

As expected, in our model the value of oil production is found to increase corruption. Our result is therefore consistent with what is commonly found in the literature as the “natural resources’ curse”, that is to say, the evidence according to which countries whose economy is based on the exploitation of natural resources tend to perform worse in terms of social development (and here, in terms of corruption, too). It is also consistent with Montinola and Jackman (2002) results, according to which corruption is positively affected by the membership of the OPEC (Oil Production and Exporting Countries).

Population and time trend appear not to affect the level of perceived corruption. Our index of democratization appears not significant in the model. This is consistent with the view of Treisman (2000), according to which the current degree of democracy does not affect corruption, but probably only a long exposure to democracy does. Life expectancy is negatively correlated with corruption. Kaufmann et al. (1999) indeed find a negative relationship between corruption and life expectancy, although they studied the reverse causal relationship between those two variables.

All the three models are well specified since Sargan tests do not reject the null of validity of over-identifying restrictions with p-values ranging around 0.8, globally significant since Wald tests reject the null of global non-significance and moreover, no serial correlation in the residuals is detected, because both Arellano Bond AR(1) and AR(2) tests for serial autocorrelation in first and second differenced residuals

indeed accept at the standard 5% level the null of no first and second order autocorrelation for all the three models. Moreover, Pesaran (2004, 2015) CD test for cross-sectional independence in the residuals accepts at the standard 5% level the null of independency among countries in all the estimated models.

5 Concluding remarks

In this paper, we tackle the possibility, little studied in the empirical literature, that inequality causes corruption, showing moreover that corruption does not produce per se income inequality. Using a panel of 50 countries during the period 1995 through 2015, with a two-step Arellano-Bond dynamic GMM estimation, we show that increases in income inequality are responsible for increasing corruption (here, perceived corruption), while corruption does not statistically impact inequality. This result is also robust with respect to the inclusion of additional variables potentially affecting corruption.

Our tests of causality are indeed coherent with this view: while the standard Granger causality test which assumes homogeneity of coefficients among countries suggests that is income inequality that causes corruption and not vice versa, Dumitrescu and Hurlin's test is less categorical, admitting the possibility that in some countries corruption might cause income inequality. However, we reasonable believe that the number of countries for which corruption causes income inequality is smaller that the number of countries for which inequality causes corruption, because the significance of the test associated to the direction of causality "corruption causes inequality" is smaller than the one associated to the relation "income inequality causes corruption".

So, along to what is commonly found in the econometric literature, we find strong evidence that income inequality is potentially an important determinant of corruption, or maybe, of petty corruption (the kind of corruption which is more pervasive in all people's day life, and which is more likely to affect the index of perception of corruption).

The reasons why we believe that income inequality is responsible for increasing corruption are several, and the main is that when poverty is widespread and people underpaid, the incentives to pursue wealth (even in an unfair and illegal way) increase, and with them corruption. In this sense, further research should focus on how inequality affects people's social norms about corruption and beliefs about the legitimacy of rules and institutions. Since corruption can be driven by imitative behavior or may be already a norm

of behavior, and this is an important research question. Future research could also study who has more incentive to support corruption, between rich and poor, since wealthy people have both greater motivation and opportunity to engage in corruption, while the poor are more vulnerable to extortion and less able to monitor and hold the rich and powerful accountable.

References

- [1] Abed G. and Davoodi H., 2000. Corruption, Structural Reforms and Economic Performance in the Transition Economies. IMF Working Paper No.132. Washington: International Monetary Fund.
- [2] Ades A., and Di Tella R., 1997. The New Economics of Corruption: A Survey and some New Results. *Political Studies* 45: 496-515.
- [3] Andres A., and C. Ramlogan-Dobson, 2011. Is Corruption Really Bad for Inequality? Evidence from Latin America, *The Journal of Development Studies* 47(7), pp. 959-976.
- [4] Akçay S., 2006. Corruption and Human Development. *Cato Journal* 26(1): pp. 29-48.
- [5] Al-Marhubi, F. A., 2000. Corruption and Inflation. *Economics Letters* 66: 199-202.
- [6] Apergis N., Dincer O. and J. Payne, 2010. "The relationship between corruption and income inequality in U.S. states: evidence from a panel cointegration and error correction model", *Public Choice* 145, pp. 125-135.
- [7] Arellano, M. and S. Bond., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58. pp. 277 ? 297.
- [8] Azfar O. and Gurgur, T., 2005. Does Corruption Affect Health and Education Outcomes in the Philippines? Available at SSRN: <https://ssrn.com/abstract=723702> or <http://dx.doi.org/10.2139/ssrn.723702>
- [9] Bardhan P., 1997. Corruption and development: a Review of Issues. *Journal of Economic Literature* 35: pp. 1320-1346.
- [10] Barro, Robert J. and Jong-Wha Lee, 2013. A new data set of educational attainment in the world, 1950-2010. *Journal of Development Economics* 104: 184?198.

- [11] Bosco, B., 2016. Old and new factors affecting corruption in Europe: Evidence from panel data. *Economic Analysis and Policy* vol. 51, pp. 66-85.
- [12] Champernowne D.G. and F.A. Cowell, 1998. "Economic inequality and Income distribution". Cambridge University Press. ISBN 0 521 5855 2.
- [13] Choi, I. 2001, Unit root tests for panel data. *Journal of International Money and Finance* 20: 249-272.
- [14] Chong A., and M. Gradstein, 2007. "Inequality and Institutions", *The Review of Economics and Statistics* 89(3), pp. 454-465.
- [15] Cohen, Daniel and Laura Leker, 2014. Health and Education: Another Look with the Proper Data, mimeo Paris School of Economics.
- [16] De la Croix, David, and Matthias Doepke. 2003. Inequality and Growth: Why Differential Fertility Matters. *American Economic Review*, 93(4): 1091-1113.
- [17] Dimant, E. and Tosato, G. (2017). Causes And Effects of Corruption: What has Past Decade's Empirical Research Taught us? A Survey. *Journal of Economic Surveys*, DOI: 10.1111/joes.12198
- [18] Dumitrescu E. and Hurlin C., 2012. Testing for Granger Non-causality in Heterogeneous Panels. *Economic Modelling* 29:1450-1460.
- [19] Dincer C. O. and Gunalp B., 2005. Corruption, Income Inequality, and Growth: Evidence from U.S. States. *SSRN Electronic Journal*. DOI: 10.2139/ssrn.690381.
- [20] Dutta N., Kar S. and Saha S., 2017. Human Capital and FDI: How does Corruption affect the relationship? *Economic Analysis and Policy* 56, pp. 126-134.
- [21] Fakir, Adnan M.S. Ahmad, Azraf Uddin Hosain, K.M. Masnun Hossain, Mostafa Rafid Gani, Ridhim Sadman, 2017. The comparative effect of corruption and Piketty's second fundamental law of capitalism on inequality. *Economic Analysis and Policy* vol. 55, pp. 90-105.
- [22] Ghura, D. (1998) Tax Revenue in Sub-Saharan Africa: Effects of Economic Policies and Corruption. IMF Working Paper. Washington: International Monetary Fund.

- [23] Gleditsch, K. S. (2002). Expanded trade and gdp data (version 6.0). *Journal of Conflict Resolution*, 46 (5), 712-724.
- [24] Glomm G. and B. Ravikumar, 1992. Public Versus Private Investments in Human Capital: Endogenous Growth and Income Inequality. *The Journal of Political Economy* vol. 100(4), pp. 818-834.
- [25] Gupta, S.; Davoodi, H.; and Tiongson, E., 2000. Corruption and the Provision of Health Care and Education Services. IMF Working Paper No. 116. Washington: International Monetary Fund.
- [26] Gymah-Brempong K., 2002. Corruption, Economic Growth and Income Inequality in Africa. *Economics of Governance* 3:183-209.
- [27] Gupta, S.; Davoodi, H.; and R. Alonso-Terme, 2002. "Does corruption affect income inequality and poverty?", *Economics of Governance* 3, pp. 23-45.
- [28] Habib, M., and Zurawicki, L., 2001. Country-Level Investments and the Effect of Corruption: Some Empirical Evidence. *International Business Review* 10: 687-700.
- [29] Kaufmann, D., Kraay A., and Zoido-Lobaton P., 1999. Governance Matters. World Bank Policy Research Department Working Paper No. 2196. Washington: World Bank.
- [30] Kuznets, Simon, 1955. Economic Growth and Income Inequality. *American Economic Review* 45, pp. 1-28.
- [31] Im K. S., M. H. Pesaran, and Y. Shin, 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115: 53-74.
- [32] Lambsdorff J., 2006. The validity and Precision of subjective indicators (CPI), in C. Sampford, A. Shocklock, C. Connors, and F. Galtung, eds., *Measuring Corruption*, Ashgate, Adelshot, pp. 81-99.
- [33] Leite C. and Weideman J., 1999. Does Mother Nature Corrupt? Natural Resources, Corruption and Economic Growth. IMF Working Paper No.85. Washington: International Monetary Fund.
- [34] Levin A., C.F. Lin, and C.S. J. Chu, 2002. Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics* 108: 1-24.

- [35] Li, H., Squire, L., Zou, H., 1998. Explaining International and Intertemporal Variations in Income Inequality. *The Economic Journal* 108, 26-43.
- [36] Maddala, G. S., and S. Wu. 1999. A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* 61: 631-652
- [37] Mauro P., 1996. The Effects of Corruption on Investment, Growth and Government Expenditure. IMF Working Paper No. 98. Washington: International Monetary Fund
- [38] Mauro P., 1998. Corruption and Composition of Government Expenditure. *Journal of Public Economics* 69: 263-79.
- [39] Montinola J.R., and Jackman R.W., 2002. Sources of Corruption: a cross-country study. *British Journal of Political Science* 32(1), p. 147-170.
- [40] Nickell, Stephen, 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica* vol 49(6), pp. 1417-1426.
- [41] Nur-Tegin K. and Czap H.J., 2012. Corruption: Democracy, Autocracy, and Political Stability. *Economic Analysis and Policy* 42(1), pp. 51-66.
- [42] Olken, Benjamin, 2007. Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* Vol. 115(2) : pp. 200-249.
- [43] Paldam M., 2002. The Cross-Country Pattern of Corruption: Economics, Culture and the Seesaw Dynamics. *European Journal of Political Economy* Vol. 18(2), pp. 215-240.
- [44] Parcero O.J. and Papyrakis E., 2016. Income Inequality and the Oil Resource Curse. *Resource and Energy Economics* vol. 45, pp. 159-177.
- [45] Pemstein D., Meserve S.A. and Melton J., 2010. Democratic Compromise: a Latent Variable Analysis of Ten Measures of Regime Type. *Political Analysis* Vol. 18(4), pp. 426-449.
- [46] Pesaran, M. Hashem, 2004. General Diagnostic Tests for Cross Section Dependence in Panels. IZA Discussion Paper No. 1240.

- [47] Pesaran, M. Hashem, 2015. Testing Weak Cross-Sectional Dependence in Large Panels. *Econometric Reviews* Vol. 34(6-10), pp.1089-1117.
- [48] Policardo L., Punzo L., and Sanchez Carrera E.J., 2016. Unsustainable inequality: is there a turning point? Università degli Studi di Siena, WP No. 728.
- [49] Ross, M. and Mahdawi P., 2015. Oil and Gas Data, 1932-2014. Harvard Dataverse.
- [50] Saha S. and Ben Ali M.S., 2017. Corruption and Economic Development: New Evidence from the Middle Eastern and North African Countries. *Economic Analysis and Policy* 54, pp 83-95.
- [51] Saha S. and Su J.J., 2012. Investigating the Interaction Effect of Democracy and Economic Freedom on Corruption: A Cross-Country Quantile Regression Analysis. *Economic Analysis and Policy* 42(3), pp 389-396.
- [52] Sargan, J. D., 1975. Testing for misspecification after estimating using instrumental variables. Mimeo. London School of Economics.
- [53] Solt F., 2011. Standardizing the World Income Inequality Database. *Social Science Quarterly* 90(2):231-242. SWIID Version 3.1, December 2011
- [54] Suryadarma D., 2012. How Corruption Diminishes the Effectiveness of Public Spending on Education in Indonesia. *Bulletin of Indonesian Economic Studies* Vol. 48(1), pp. 85-100.
- [55] Tanzi V. and Davoodi H., 1998. Corruption, Public Investment, and Growth, in H. Shibata et al. (eds.), "The Welfare State, Public Investment, and Growth", Springer-Verlag, Tokyo, pp. 41-60.
- [56] Tanzi V. and Davoodi H., 2000. Corruption, Growth and Public Finances. IMF Working Paper No. 116. Washington: International Monetary Fund.
- [57] Teorell, Jan, Stefan Dahlberg, Sren Holmberg, Bo Rothstein, Natalia Alvarado Pachon Richard Svensson. 2018. The Quality of Government Standard Dataset, version Jan18. University of Gothenburg: The Quality of Government Institute, <http://www.qog.pol.gu.se> doi:10.18157/QoGStdJan18
- [58] Treisman D., 2000. The Causes of Corruption: a Cross National Study. *Journal of Public Economics* Vol. 76(3), pp. 399-457.

- [59] Ullah M.A. and Ahmad E., 2016. Inequality and Corruption: Evidence from Panel Data. *Forman Journal of Economic Studies* Vol. 12, pp. 1-20
- [60] Van Rijckeghem, C. and Weder, B. (1997) Corruption and the rate of temptation: do low wages in the civil service cause corruption. *International Monetary Fund Working Paper*, IMF, Washington, DC.
- [61] Wilhelm, Paul G., 2002. International Validation of the Corrupt Perception Index: Implications for Business Ethics and Entrepreneurship Education. *Journal of Business Ethics and Entrepreneurship Education* 35(3): 177-189.
- [62] You, Jong-Sung and Sankeev Khagram, 2005. A comparative Study of Inequality and Corruption. *American Sociological Review* 70(1), pp. 136-157.

