Economics Bulletin

Volume 39, Issue 4

Is the effect of corruption on entrepreneurial activity nonmonotonic? A semi-parametric panel data analysis

Michael Polemis
University of Piraeus

Abstract
We investigate the impact of corruption on entrepreneurial activity in the OECD countries by using a semi-parametric panel data model developed in Robinson (1988). We argue that this relationship exhibits a nonmonotonic shape with multiple turning points. Our findings are robust to different measures of corruption and alternative dimensions of entrepreneurship.

The author is indebted to John P. Conley (Editor) and Stephen Bazen (Associate Editor) for giving him the opportunity to revise and thus improve the paper. Special thanks also go to two anonymous referees for their constructive comments and suggestions that enhanced the merit of this study. All the remaining errors belong to the author. The usual disclaimer applies.
Citation: Michael Polemis, (2019) "Is the effect of corruption on entrepreneurial activity nonmonotonic? A semi-parametric panel data analysis", *Economics Bulletin*, Volume 39, Issue 4, pages 2976-2989
Contact: Michael Polemis - mpolemis@unipi.gr.
Submitted: August 18, 2019. **Published:** December 24, 2019.

1 Introduction

There is no doubt that institutions do play an important role in stimulating economic growth (Acemoglu *et al* 2005, and Aghion *et al*. 2007). Bearing this in mind, control of corruption is of paramount importance to enhance institutional quality and foster entrepreneurial activity (Mauro 1995, and Berdiev and Saunoris 2018). The study of this relationship constitutes an important issue since the prevalence of entrepreneurial activity, which is caused by corruption may downturn economic growth (Swaleheen 2012). Specifically, a high level of corruption may facilitate the shadow economy, accelerating unproductive entrepreneurship (Baumol 1990) or by decreasing growth-oriented entrepreneurship (Aidis and Mickiewicz 2006).

While the majority of existing literature assumes that the linkage between corruption and entrepreneurial activity is monotonic and negative (see among others Berdiev and Saunoris 2018, and Goel and Saunoris 2019), some studies claim that the shape of this relationship might be concave or convex (Anokhin and Schulze 2009, Rose-Ackerman, 2001, and Mohamadi *et al* 2017). This argument is based on the theoretical ground that institutional trust may rise slowly in response to improvements in the control of corruption suggesting a nonmonotonic relationship (Anokhin and Schulze 2009, Aidis *et al*. 2012, and Dreher and Gassebner 2013).

However, the exact shape of the curve (convex or concave) remains an open question. On the one hand, it stands to reason that the effects of a unit of improvement in the control of corruption may have greater impact when corruption is high than when it is low. If this is the case, one might expect the relationship between the control of corruption and entrepreneurial activity to be positive and convex (Anokhin and Schulze 2009). The decreasing part of the curve justifies that when inefficient governments are in power, entrepreneurs can use corrupt practices to bypass highly regulated, wasteful, and ambiguous regulative arrangements (Mohamadi *et al*. 2017). On the other hand, institutional trust may rise slowly in response to improvements in the control of corruption (see Rose-Ackerman 2001, 2004). This means that it needs some time to accumulate. As clearly stated by Anokhin and Schulze (2009), "*In the early stages of an improved corruption climate, venturing and other forms of innovative activity will be viewed by others as an experiment. Domestic rates of entrepreneurial activity might then be expected to rise slowly at first, but accelerate as early experiments pay off and institutional trust rises*". In such a case, the empirical relationship takes the form of a concave curve. Specifically, the increasing part of the curve denotes that if governments persist in controlling corruption, societal trust emerges, which in turn favors entrepreneurship.

To investigate the exact shape of the relationship between corruption and entrepreneurial activity we used two widely used corruption measures (control of corruption and corruption index) as our basic variables over four different dimensions of entrepreneurial activity (i.e., total early stage entrepreneurship, nascent entrepreneurship, new business owners and established business owners). The reason for using various types of entrepreneurial activity is to check for the robustness of our findings since many of the existing studies examine the effect of corruption to specific categories such as nascent entrepreneurship (see for example Mohamadi *et al*. 2017).

Contrary to the existing studies, which presume strong assumptions (i.e., known functional form, specific distribution of the error term), we employ a flexible semi-parametric double residual estimator developed by Robinson (1988) to properly avoid potential form misspecification. The reason for using this estimator to guide our research is that it leads to small bias and outperforms other semi-parametric techniques (i.e., Yatchew's differencing estimator) as argued in Verardi and Debarsy (2012).

Our findings support the existence of nonmonotonic effects between corruption and all the different types of entrepreneurial activity. Specifically, a nonlinear convex "*U shaped*" curve arises for nascent and new entrepreneurs, while the control of corruption exhibits a concave curve for other entrepreneurs. The specification test results further support the superiority of the semi-parametric model.

The rest of the paper is organised as follows: Section 2 describes the sample and the semi-parametric methodology applied (double residual estimator). Section 3 discusses the parametric and the semi-parametric results along with the necessary robustness checks, while Section 4 concludes the paper providing some policy implications.

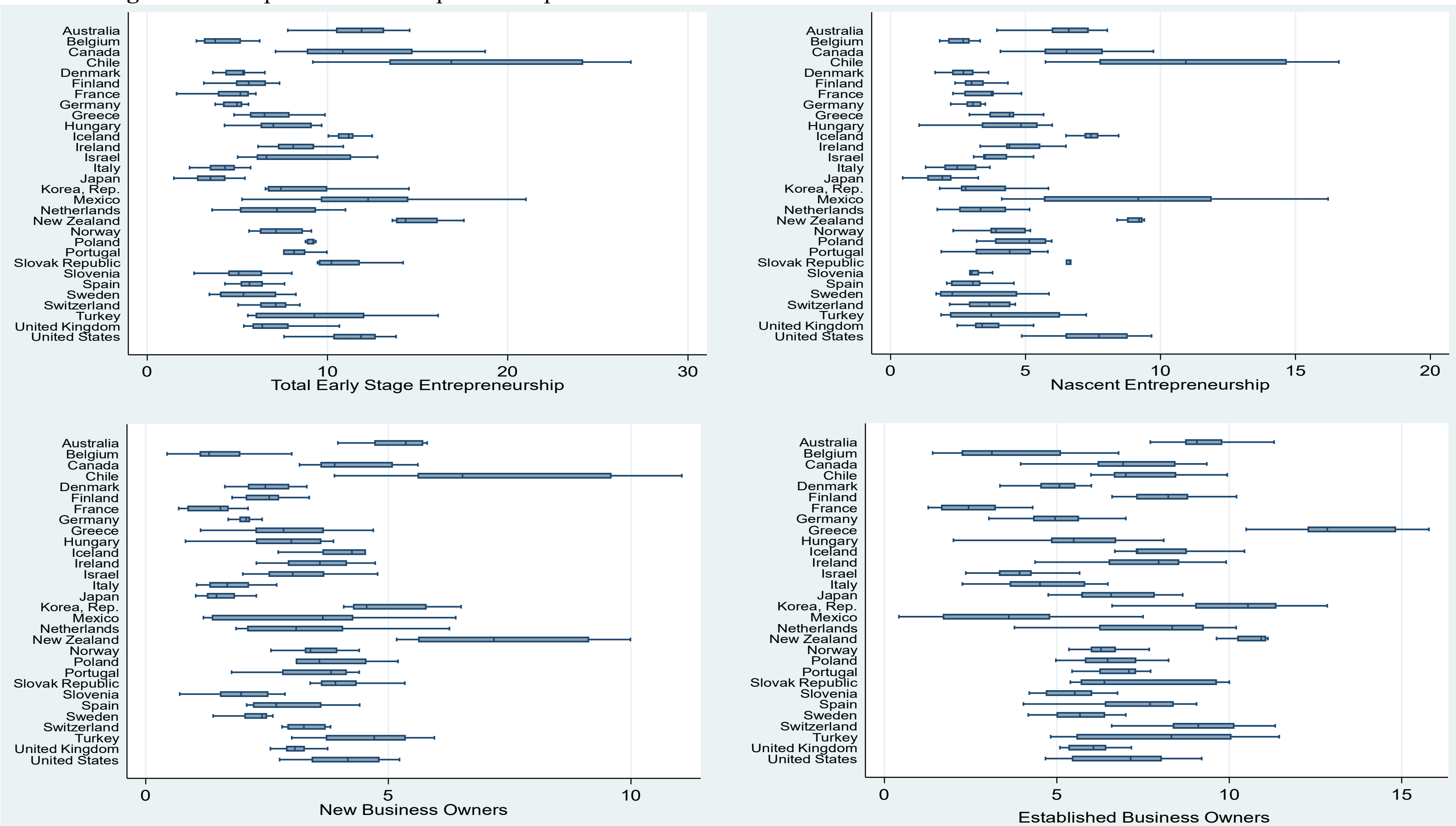
2 Data and methodology

2.1 Sample and descriptive statistics

The sample consists of an unbalanced panel data set of 30 OECD countries over the period 2002–2017. We employ two alternative corruption measures, widely used by the literature (Berdiev and Saunoris 2018, and Swaleheen 2012), namely the control of corruption (CORWGI) taken from the Worldwide Governance Indicators database and the corruption index (CORIRG) drawn from the International Country Risk Guide. The former ranges from -2.5 (weak control) to 2.5 (strong control) while the latter takes values within the interval 0 (more corruption) to 6 (less corruption). The rest of the variables except for the index of economic freedom, are taken from the World Bank. Table A.1 in the Appendix reports the descriptive statistics of the sample variables, while Figures (1a&b) present the Boxplots of the dependent variables and the corruption indicators broken down by OECD country.

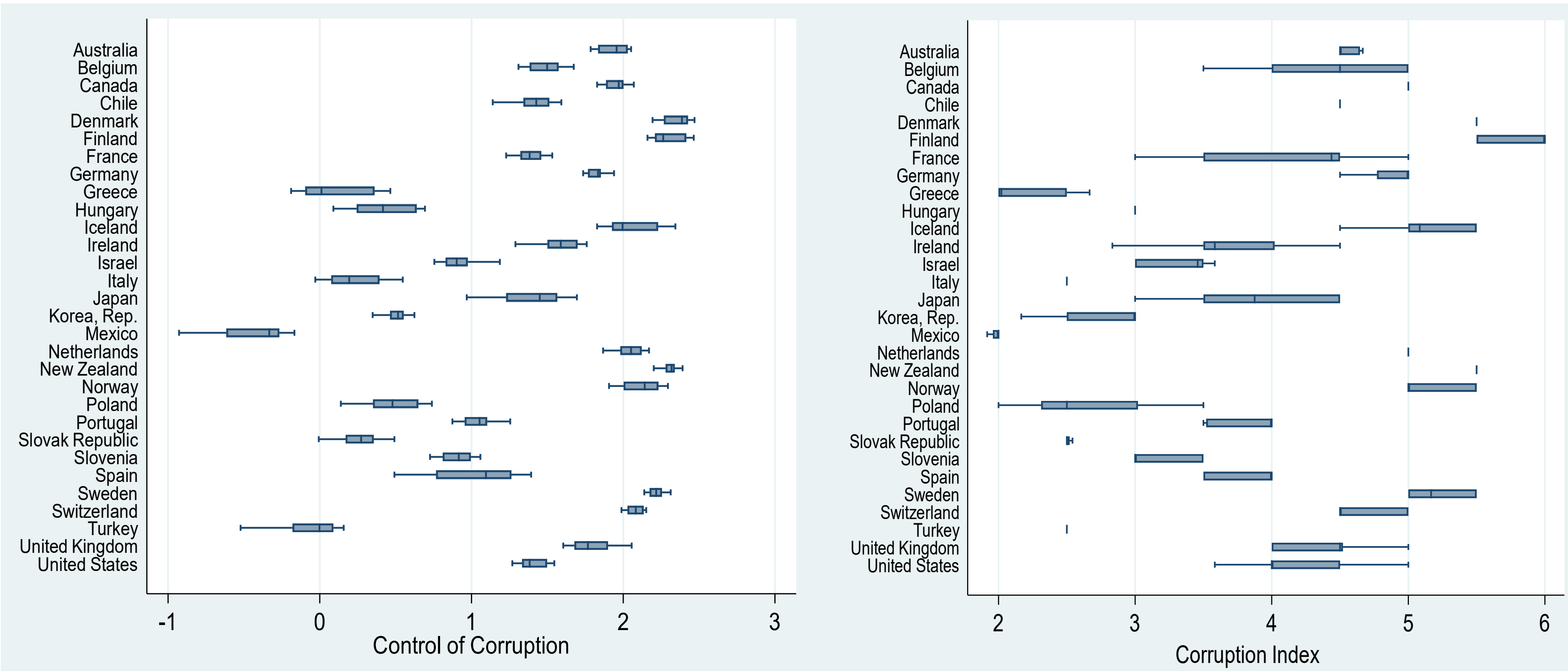
As it is evident from Figure (1a), countries such as Chile, Mexico, New Zealand and Turkey exhibit high variability in terms of entrepreneurial activity, while Canada, Australia, United Kingdom and United States present almost the least variability among the sample countries. Moreover, the distribution of entrepreneurial activity in nearly all if its four dimensions (TEA, NASC, EST and NEW) in Chile, Greece, New Zealand and Slovak Republic has a bit of a tail or it is skewed to the right (i.e., towards the large numbers of entrepreneurial activity). Figure 1b, presents the distribution of the two corruption measures namely the control of corruption indicator (left panel) and the corruption index (right panel). From the careful inspection of the relevant figure, we argue that the countries with the most variability in terms of control of corruption (see left panel of the figure) is Spain, Greece, Hungary and Mexico, while the opposite holds for the Netherlands, United Kingdom, United States, Norway, France, Germany, Sweden and Switzerland. Lastly, for countries like Japan, Poland, Ireland and Sweden the distribution of the corruption index (CORIRG) is skewed to the right indicating low levels of corruption (see right panel of the figure). The opposite holds for France and Greece, while Mexico and Slovak Republic exhibit the lowest variability in terms of corruption.

Figure 1a: Boxplots of the entrepreneurship variables



Notes: Outside values (outliers) have been excluded from the figure.

Figure 1b: Boxplots of the corruption indicators



Notes: Outside values (outliers) have been excluded from the figure.

2.2 Methodology

Let the partially linear model be given by the following equation:

$$y_i = a + x_i^T \beta + f(\psi_i) + \varepsilon_i \qquad i = 1,...,\text{N} \qquad (1)$$

where y_i is the value taken by the dependent variable for country *i*, x_i^T is the vector of exogenous linear regressors, $f(\psi_i)$ is an unknown function of ψ_i entering the model in a nonlinear way, α is a constant term and ε_i is the i.i.d error term. [1] Following Robinson (1998), we apply a conditional expectation to both sides of Eq. (1) as follows:

$$E\left(y_i \middle| \psi_i\right) = a + E\left(x_i^T \middle| \psi_i\right)\beta + f(\psi_i) + \underbrace{E\left(\varepsilon_i \middle| \psi_i\right)}_{0} \qquad (2)$$

By taking the difference of the two equations, we have:

$$\underbrace{y_i - E\left(y_i \middle| \psi_i\right)}_{\varepsilon_1} = \underbrace{\left[x_i^T - E\left(x_i^T \middle| \psi_i\right)\right]}_{\varepsilon_2}\beta + \varepsilon_i \qquad (3)$$

In the case that conditional expectations are unknown, they have to be estimated by relying on some consistent estimators $y_i = m_y(\psi_i) + \varepsilon_{1i}$ and $x_{ki} = m_{xk}(\psi_i) + \varepsilon_{2ki}$, where k = 1,…,K is the index of the explanatory variables entering the model parametrically (Verardi and Debarsy 2012). After replacing them in Eq. (3), the double residual estimator $\widehat{\beta}$ is given as:

$$\widehat{\beta} = \left(\widehat{\varepsilon}_2'\widehat{\varepsilon}_2\right)^{-1}\left(\widehat{\varepsilon}_2'\widehat{\varepsilon}_1\right) \qquad (4)$$

After having estimated the parameter vector $\widehat{\beta}$, we can fit the nonlinear relation between ψ_i and y_i by estimating Eq. (5) nonparametrically:

$$y_i - x_i^T \widehat{\beta} = \alpha + f(\psi_i) + \varepsilon_i \qquad (5)$$

To check for the appropriateness of a parametric polynomial alternative approximation, Hardle and Mammen (1993) developed a test statistic that compares the nonparametric and parametric regression fits by using squared deviations between them (see Verardi and Debarsy 2012). With this test, we are able to check if the nonparametric part of the relation may be better approximated by a polynomial functional form.

Specifically, Hardle and Mammen (1993) propose a testing procedure based on square deviations between the nonparametric kernel estimator $\widehat{m}(z_i)$ with bandwidth h and a parametric regression $\widehat{f}(z_i,\theta)$. The relevant specification tests assess if the nonparametric fit can be approximated by a parametric adjustment of order k. The test statistic they propose is given as:

$$\text{T}_n = N\sqrt{h}\sum_{i=1}^{n}\left\{\widehat{f}(z_i) - \widehat{f}(z_i,\theta)\right\}^2 \pi(\cdot) \qquad (6)$$

where $\widehat{f}(z_i)$ is the nonparametric function estimated in Eq. (5) $\widehat{f}(z_i,\theta)$ is an estimated parametric function, h is the bandwidth used, and π (·) is a weighting function for the squared deviations between fits. To obtain critical values, Hardle and Mammen (1993) suggest using the wild bootstrap (Verardi and Debarsy 2012). According to the relevant test, an absence of rejection of the null hypothesis means that the polynomial adjustment is suitable. In such a case, the parametric model is accepted revealing that the polynomial adjustment is at least of the degree that has been tested.

[1] For presentational simplicity we omitted subscript t.

The dependent variables of the y vector contain the four dimensions of entrepreneurship activity, namely total early stage entrepreneurial activity TEA (% of 18-64 population), nascent entrepreneurship rate NASC (% of 18-64 population), new business ownership rate NEW (% of 18-64 population) and established business ownership rate EST (% of 18-64 population). The x vector includes the list of the covariates. These are the government final consumption expenditure to GDP (GOV) as a measure of government size, the GDP per capita as a proxy for total income, the level of economic freedom (FREED), the population density (POP), and the sum of exports and imports over GDP (TRADE) as a proxy for trade openness. Finally, ψ includes the two corruption measures (CORWGI and CORIRG) allowing for possible nonlinearities.

3 Results and discussion

Following the spirit of Goel and Saunoris (2019), we estimate the baseline (parametric) model as follows:

$$y_i = x_i \alpha + Z_i \beta + \gamma_i + \delta_t + \varepsilon_i \quad (7)$$

where y is a vector including the four dependent variables, namely TEA, NASC, NEW and EST. Vector x consists of the two corruption measures expressed in first and second polynomial order (CORWGI, $CORWGI^2$, CORIRG and $CORIRG^2$). Moreover, Z denotes the vector of the rest of the covariates (GOV, GDP, POP, FREED and TRADE). The model also accounts for country and year fixed effects (γ_i, δ_t) to allow for time invariant unobserved heterogeneity.[2]

Similarly with other empirical studies (see also Papaioannou, 2017, and Dai *et al*. 2014), we address endogeneity concerns and reverse causality, by adopting the 2SLS method. The relevant method proceed as follows: In the first stage, we regress the endogenous variable (control of corruption and corruption index) on government effectiveness (GEFF).[3] In the second stage, we use the predicted (fitted) values of the corruption measures (CORWGI and CORIRG) in levels and quadratic form drawn from the first stage regressions as the main independent variables along with the rest of the covariates (see Eq. 7). In this case, the used instrument (i.e., government effectiveness) is excludable and properly addresses the reverse causality and possible endogeneity bias of the model.

The reason for choosing the specific variable to serve as a proper instrument in the estimation of corruption-entrepreneurship nexus is twofold. First, the relevant variable is strongly correlated with the potentially endogenous regressor (corruption measure), acting as a genuinely exogenous instrument. Second, it only influences the dependent variables (TRA, NASC, NEW and EST) through the potentially endogenous independent variables (CORWGI and CORIRG). One could also use an alternative approach to deal with reverse causality by utilizing lagged values of the endogenous

[2] We have used STATA ver. 15 and the command "*semipar*" to estimate our model (Verardi and Debarsy 2012). The use of the fixed effects were justified by applying the Hausman test where the null hypothesis was clearly rejected. To preserve space, the results are available upon request.

[3] The relevant instrument reflects perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. Government effectiveness/efficiency and corruption refer to two different phenomena, as corruption is understood as an abuse of public positioning for private gain (Rodriguez *et al*. 2005). This distinction is in line with what is known as the "*grease the wheels*" hypothesis (see Mohamadi *et al*. 2017). The variable ranges from -2.5 (weak) to 2.5 (strong) governance performance. This variable is also drawn from the Worldwide Governance Indicators database.

variables. However in this case, a reasonable theory as to why the relevant instruments (i.e., lagged values of corruption) can be excluded is needed to properly interpret the results.

Table 1 presents the parametric results. Regarding the control of corruption we observe that in all of the specifications (see Table 1; Columns 1,3,5 and 7) the estimated coefficients alternate in sign starting from negative (first order) to positive (second order). This suggests a nonmonotonic quadratic ("*U-shaped*") relationship. The pattern of the curve implies that the influence is positive for low and high degrees of institutional trust. The rest of the covariates are in most cases statistically significant and properly signed.

However, when we employ the other measure of corruption (CORIRG) we confirm that its "*U-shaped*" nonmonotonic influence is not statistically significant in all of the specifications.[4] This finding suggests the absence of a nonlinear (convex) effect of corruption on entrepreneurial activity like the one we showed earlier with the control of corruption measure (CORWGI). Based on the above, we argue that from an economic perspective there is an asymmetry since measures to reduce corruption exhibit a non-linear effect on entrepreneurship, but not the extent of corruption itself.

[4] The results are available upon request.

Table 1: Parametric results

	Dependent variable: *Total early stage entrepreneurship*	**Dependent variable:** *Nascent entrepreneurship*	**Dependent variable:** *New business owners*	**Dependent variable:** *Established business owners*
		Specification		
Regressors	**(1)**	**(2)**	**(3)**	**(4)**
Control of Corruption	-4.632***	-3.853**	-1.581*	-1.115
	(1.638)	(1.583)	(0.880)	(1.406)
Control of Corruption *(squared)*	1.426**	1.352**	0.521*	0.0104
	(0.594)	(0.558)	(0.284)	(0.491)
Government Size	0.320**	0.0866	0.127*	0.378***
	(0.142)	(0.0969)	(0.0918)	(0.0931)
Population Density	0.0458***	0.0225*	-3.45e-05	0.0156*
	(0.0145)	(0.0157)	(0.0104)	(0.00968)
Economic Freedom	0.000812	-0.0170	0.0102	0.0312
	(0.0626)	(0.0481)	(0.0310)	(0.0486)
Income	0.000298***	0.000201***	6.92e-05*	-2.82e-05
	(6.35e-05)	(5.25e-05)	(4.11e-05)	(4.30e-05)
Trade Openness	-0.0262*	-0.0115	-0.0159*	0.0278***
	(0.0146)	(0.0102)	(0.00980)	(0.0100)
		Diagnostics		
Observations	380	335	335	380
Adjusted R^2	0.190	0.127	0.137	0.129
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
F-test (p-value)	0.0000	0.0001	0.0069	0.0000
Kleibergen-Paap LM statistic	17.255***	12.607***	12.607***	17.255***
	[0.000]	[0.0004]	[0.0004]	[0.000]
Cragg-Donald Wald F statistic	24.504	13.874	13.874	24.504
	{5.53}	{5.53}	{5.53}	{5.53}
D-W-H test	19.42648***	30.13773***	19.16361***	3.93159*
	[0.00006]	[0.00000]	[0.00007]	[0.10322]

Notes: Time dummies are included but not reported. Kleibergen-Paap LM statistic and Cragg-Donald Wald F statistic denote the under identification and weak identification tests respectively where rejection of the null hypothesis indicates that the model is properly identified. P-values in brackets, while number in {} denotes the critical values for Cragg-Donald F statistic and i.i.d. errors for 25% maximal IV size. D-W-H is the Durbin-Wu-Hausman test for endoegenity. The null hypothesis denotes that any endogeneity among the regressors would not have deleterious effects on OLS estimates. A rejection of the null indicates that endogenous regressors' effects on the estimates are meaningful, and IV techniques are required (Davidson and MacKinnon 1993). Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

In addition, several precautions are taken in order to avoid the problem of instrument proliferation. Tests reported in the bottom of Table 2 clearly show that our instruments are exogenous and they do not suffer from weak identification problem (see Cragg-Donald Wald F statistic). Further, the instrumental variable technique is necessary while our model is not under-identified (see Kleibergen-Paap LM statistic) since in nearly all of the specifications the null hypothesis is rejected. Taken together, we argue that the parametric model is properly identified.

In order to justify the use of the instrumental variable (IV) method over the standard OLS estimators, we check for endogeneity by employing the Durbin-Wu-Hausman (D-W-H) test (see Davidson and MacKinnon 1993). The latter, computes a test for endogeneity in a regression estimated via IV, the null hypothesis for which states that an OLS estimator of the same equation would yield consistent estimates (Baum *et al*. 2003). In other words, any endogeneity among the regressors would not have deleterious effects on OLS estimates (Hausman 1978). Contrary, a rejection of the null indicates that endogenous regressors' effects on the estimates are meaningful, and IV techniques are required.[5] As it is evident from Table 1 (see last row), in all of the specifications the null hypothesis is rejected justifying the use of the IV approach (2SLS).

Table 2: Specification test results

	Specification			
Polynomial order (k)	**(1)**	**(2)**	**(3)**	**(4)**
0	2.803** [0.02]	3.733*** [0.00]	2.051** [0.04]	1.926* [0.05]
1	2.473** [0.04]	1.981** [0.04]	1.935* [0.05]	1.954* [0.05]
2	2.778** [0.02]	1.965* [0.05]	3.348*** [0.00]	2.046** [0.04]
3	1.008 [0.30]	3.922** [0.01]	4.190*** [0.00]	0.548 [0.74]

Notes: P-values in brackets. 100 bootstrap replicates were used. *** p<0.01, ** p<0.05, * p<0.1. Acceptance (rejection) of the null hypothesis means that the parametric (nonparametric) polynomial model of order k (= 0,1,2,3) is suitable.

In the next stage we check whether the estimated parametric approximation is appropriate, by applying the specification test proposed by Hardle and Mammen (1993). As it turns out this assumption is clearly rejected (see Table 2). This means that the parametric model does not capture the nonlinear effects generated by the inclusion of the quadratic term in all of the specifications.[6] Therefore, we proceed to estimate the semi-parametric model by allowing the two corruption measures to enter nonparametrically.

Figure 2 plots estimates of the impact of control of corruption (horizontal axis) on entrepreneurial activity along with 95% confidence bands. It is evident that the relationship between the control of corruption and all of the four dimensions of entrepreneurial activity is nonlinear and statistically significant since the confidence

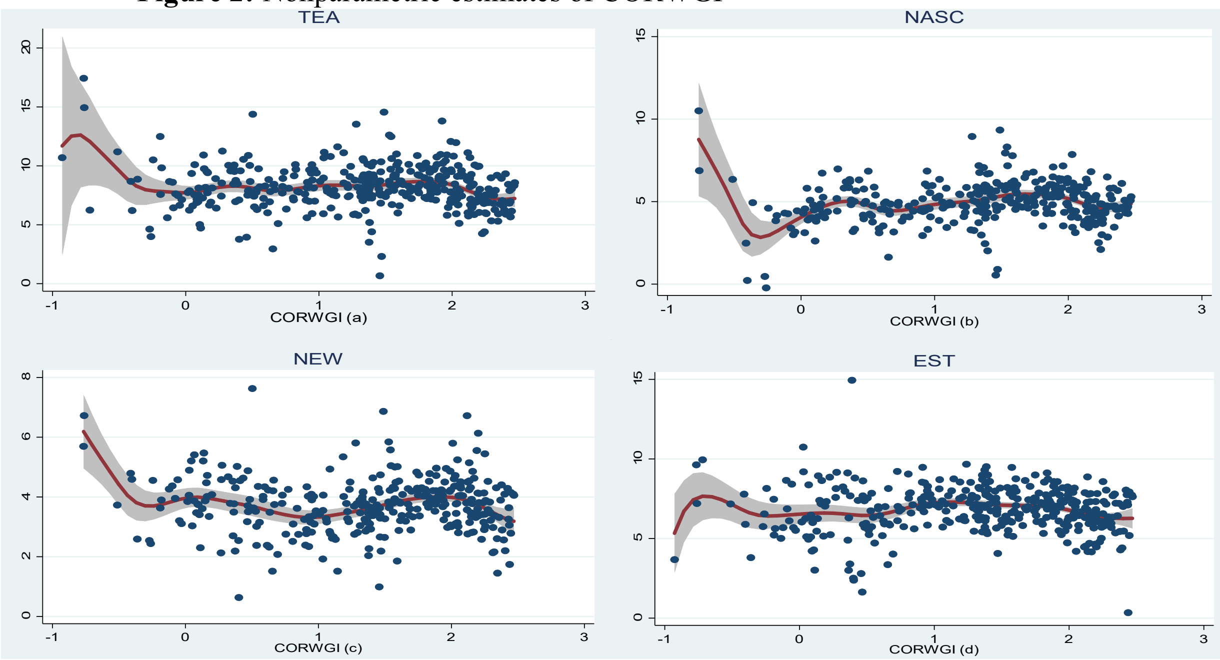
---

[5] This D-W-H test is numerically equivalent to the standard Hausman test in which both forms of the model must be estimated. Under the null, it is distributed Chi-squared with m degrees of freedom, where m is the number of regressors specified as endogenous in the original instrumental variables regression (see Baum *et al*. 2003).

[6] When under a cubic approximation, the null is no longer rejected in two specifications (see Columns 1 and 7).

bands do not include zero. From the careful inspection of Fig. (1a), it is illustrated that the maroon line uncovers a somewhat "*hump-shaped*" curvature. This means that at initial stages of an improved corruption environment, business venturing will be viewed by early stage entrepreneurs as an experiment (Anokhin and Schulze 2009).

Figure 2: Nonparametric estimates of CORWGI

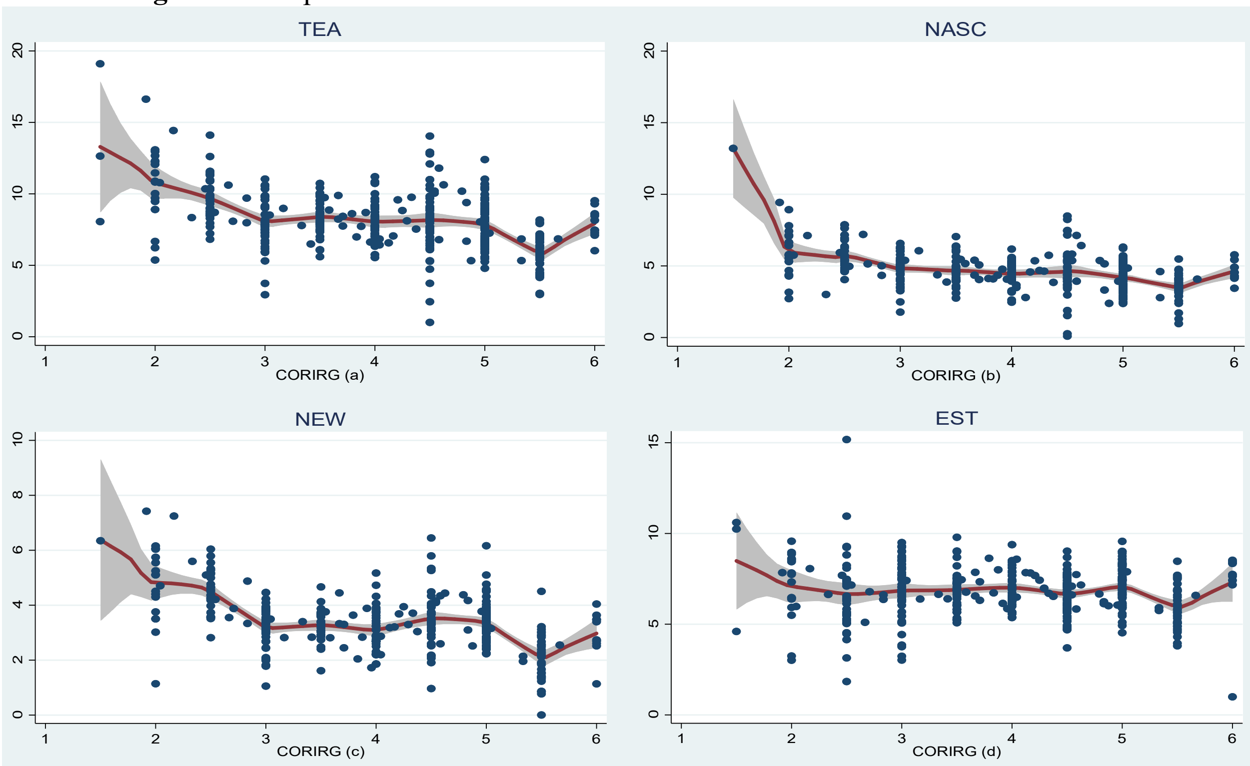


Notes: The dots in the graph represent the estimated partial residuals for entrepreneurial activity in the semi-parametric model. The maroon curve illustrates the semi-parametric estimation of *f(CORWGI)*. The Gaussian kernel weighted local polynomial fit was used, while the bandwidth is chosen by least squares cross validation. Shaded area denotes the 95% confidence bands.

As a consequence, entrepreneurial activity might then be expected to rise slowly at first, but accelerate as control of corruption rises. This evidence is in alignment with the study of Rose-Ackerman (2001) supporting a concave curve. Similar findings apply for the established entrepreneurs (see Fig.2d).

The opposite holds for panels (b) and (c) where the shape of the examined relationship resembles more of a convex ("*U-shaped*") curve especially for nascent entrepreneurs (panel b) and to a lesser extent for new business owners (panel c). In the latter case, the empirical findings reveal that the impact of control of corruption on entrepreneurial activity appears to be very weak and tentative even though mildly nonlinear. This finding which also appears in Anokhin and Schulze (2009) and Mohamadi *et al*. (2017), portrays that for nascent entrepreneurs, the effects of a unit of improvement in the control of corruption have greater impact when corruption is high than when it is low. Lastly, when we employ the other corruption measure (CORIRG), nonlinear curvatures as well as the existence of statistical significance between corruption and entrepreneurship activity still remains intact (Figure 3). These relationships further support the robustness of our findings.

Figure 3: Nonparametric estimates of CORIRG



Notes: The dots in the graph represent the estimated partial residuals for entrepreneurial activity in the semi-parametric model. The maroon curve illustrates the semi-parametric estimation of *f(CORIRG)*. The Gaussian kernel weighted local polynomial fit was used, while the bandwidth is chosen by least squares cross validation. Shaded area denotes the 95% confidence bands.

4 Conclusions

The paper documents the impact of control of corruption on entrepreneurial activity in the OECD countries by using instrumental variable parametric estimates along with a semi-parametric panel data model developed in Robinson (1988). Specifically the relevant study uses panel data for 30 OECD countries over the period 2002–2017 to analyse the relationship between a number of measures of entrepreneurial activity (total early stage entrepreneurship, nascent entrepreneurship, new business owners and established business owners) and two measures of corruption (control of corruption and corruption index).

The empirical findings of the semi-parametric model postulate that the relationship between measures of corruption and entrepreneurship has a nonmonotonic shape and is robust to alternative dimensions of entrepreneurship. The results incur significant policy implications since a prospective entrepreneur needs a long time period to cultivate the efficiency gains generated by the control of corruption. In other words, substantial changes in the level of institutional quality are less likely to accumulate and pay off rapidly.

Appendix

Table A.1: Summary statistics by country

	CORWGI	CORIRG	TEA	NASC	NEW	EST	GOV	GDP	FREED	POP	TRADE	GEFF
Country	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**	**Mean (Std)**
Australia	1.937	4.609	11.547	6.283	5.211	9.294	17.780	51,725	80.806	2.853	41.502	1.709
	(0.098)	(0.182)	(2.042)	(1.548)	(0.635)	(0.961)	(0.401)	(3,252)	(1.971)	(0.208)	(1.858)	(0.122)
Belgium	1.490	4.385	4.112	2.685	1.533	3.614	23.079	43,993	69.681	358.485	152.502	1.599
	(0.111)	(0.570)	(1.154)	(0.700)	(0.684)	(1.665)	(0.999)	(1,511)	(1.586)	(12.038)	(12.644)	(0.221)
Canada	1.953	4.888	11.846	6.747	4.245	6.961	20.402	48,209	78.306	3.721	65.759	1.824
	(0.075)	(0.257)	(3.905)	(1.734)	(0.916)	(1.667)	(0.883)	(1,839)	(2.150)	(0.182)	(4.391)	(0.075)
Chile	1.403	4.231	18.820	11.290	7.223	7.338	11.930	12,767	77.775	22.85	67.364	1.168
	(0.153)	(0.648)	(5.880)	(3.938)	(2.230)	(1.467)	(1.165)	(1,708)	(0.824)	(1.157)	(6.912)	(0.118)
Denmark	2.356	5.5	5.003	2.649	2.512	4.911	25.508	59,131	75.925	131.108	96.215	2.078
	(0.091)	(0.0)	(0.869)	(0.587)	(0.525)	(0.834)	(1.097)	(1,848)	(2.332)	(3.519)	(8.249)	(0.176)
Finland	2.298	5.822	5.570	3.201	2.422	7.709	22.830	45,934	73.45	17.602	76.224	2.099
	(0.106)	(0.239)	(1.112)	(0.601)	(0.7904)	(2.230)	(1.533)	(1,983)	(0.901)	(0.362)	(5.280)	(0.147)
France	1.382	4.114	4.638	3.364	1.360	2.578	23.434	40,914	62.343	118.170	56.832	1.508
	(0.085)	(0.698)	(1.237)	(1.078)	(0.516)	(0.983)	(0.600)	(1,099)	(1.948)	(3.030)	(3.991)	(0.137)
Germany	1.824	4.828	4.791	2.986	2.027	4.971	18.801	42,340	71.406	234.634	77.622	1.606
	(0.061)	(0.328)	(0.558)	(0.425)	(0.266)	(0.931)	(0.618)	(2,952)	(1.778)	(2.267)	(8.946)	(0.100)
Greece	0.103	2.231	6.793	4.145	2.970	12.926	20.542	25,727	57.993	85.098	56.927	0.534
	(0.235)	(0.269)	(1.410)	(0.996)	(1.08)	(3.076)	(1.181)	(2,870)	(2.802)	(0.906)	(6.614)	(0.192)
Hungary	0.421	3.0	7.060	4.296	2.793	5.422	21.060	13,496	65.662	110.798	151.585	0.701
	(0.215)	(0.0)	(2.133)	(1.453)	(0.981)	(2.064)	(1.008)	(1,076)	(1.573)	(1.814)	(19.676)	(0.168)
Iceland	2.065	5.161	11.270	7.353	4.327	7.921	24.003	44,910	73.4875	3.139	86.006	1.706
	(0.182)	(0.3475)	(1.137)	(0.891)	(1.224)	(1.202)	(0.659)	(3,562)	(2.2431)	(0.166)	(12.319)	(0.246)
Ireland	1.582	3.765	8.200	4.756	3.454	7.446	16.274	54,531	79.462	64.543	178.110	1.499
	(0.132)	(0.457)	(1.335)	(0.875)	(0.790)	(1.640)	(2.369)	(8,465)	(2.523)	(4.026)	(26.411)	(0.120)
Israel	0.930	3.291	8.089	4.262	3.087	4.047	23.424	30,351	66.681	350.166	69.868	1.267
	(0.140)	(0.249)	(2.817)	(1.692)	(0.895)	(1.017)	(1.358)	(2,658)	(2.812)	(31.536)	(8.503)	(0.101)
Italy	0.238	2.497	4.201	2.556	1.723	4.568	19.348	35,909	62.156	200.935	52.999	0.477
	(0.185)	(0.010)	(0.857)	(0.721)	(0.510)	(1.301)	(0.578)	(1,512)	(1.629)	(4.258)	(4.245)	(0.165)
Japan	1.413	3.890	3.485	1.880	1.555	6.644	19.128	45,100	70.95	350.081	29.882	1.497
	(0.205)	(0.625)	(1.214)	(0.869)	(0.427)	(1.270)	(0.9509)	(1,772)	(2.892)	(1.109)	(5.150)	(0.209)
Korea Republic	0.502	2.783	8.831	3.373	5.325	10.043	14.235	21,515	69.537	509.517	85.277	1.079
	(0.082)	(0.302)	(2.865)	(1.371)	(1.780)	(2.039)	(1.047)	(3,009)	(1.989)	(11.644)	(15.746)	(0.113)
Mexico	-0.429	1.971	12.286	9.071	3.409	3.594	11.361	9,538	65.787	58.241	61.716	0.185
	(0.237)	(0.281)	(4.663)	(3.789)	(1.707)	(2.222)	(0.765)	(423)	(1.375)	(3.837)	(8.338)	(0.099)
Netherlands	2.043	5.0	7.075	3.405	3.270	7.649	24.322	50,519	74.825	492.001	134.793	1.841
	(0.097)	(0.0)	(2.384)	(1.033)	(1.324)	(1.961)	(1.496)	(2,097)	(1.234)	(9.464)	(15.690)	(0.114)
New Zealand	2.309	5.5	14.962	9.050	7.377	10.650	18.378	34,440	81.762	16.459	57.603	1.795
	(0.0496)	(0.0)	(1.793)	(0.453)	(2.180)	(0.690)	(0.987)	(1,869)	(0.737)	(0.901)	(3.093)	(0.099)
Norway	2.118	5.177	7.374	3.997	3.650	6.417	21.243	88,592	69.15	13.359	69.459	1.898
	(0.122)	(0.239)	(1.227)	(0.903)	(0.674)	(0.901)	(1.780)	(2,314)	(2.329)	(0.696)	(2.305)	(0.071)
Poland	0.487	2.656	8.704	4.930	3.578	6.396	18.314	12,243	63.362	124.379	83.121	0.586
	(0.180)	(0.526)	(1.862)	(1.052)	(1.430)	(2.178)	(0.393)	(2,192)	(3.763)	(0.257)	(12.194)	(0.137)
Portugal	1.041	3.825	7.566	4.155	3.465	6.796	19.582	22,116	64.018	114.383	71.032	1.092
	(0.101)	(0.235)	(2.109)	(1.464)	(0.922)	(0.812)	(1.186)	(562)	(1.008)	(0.996)	(7.924)	(0.131)
Slovak Republic	0.258	2.596	10.818	7.038	4.120	7.285	18.890	15,978	66.862	112.195	162.277	0.823
	(0.133)	(0.200)	(1.718)	(1.252)	(0.771)	(1.894)	(0.812)	(2,694)	(3.360)	(0.505)	(21.887)	(0.082)
Slovenia	0.896	3.177	5.255	3.002	1.992	5.405	18.953	23,087	60.975	100.997	130.801	1.024
	(0.110)	(0.393)	(1.361)	(0.634)	(0.658)	(0.883)	(0.898)	(1,642)	(2.173)	(1.406)	(16.510)	(0.094)
Spain	1.025	3.828	5.792	2.912	2.977	7.305	18.673	30,775	68.368	90.702	57.426	1.191
	(0.309)	(0.234)	(0.940)	(0.715)	(0.816)	(1.360)	(1.369)	(1,104)	(1.571)	(3.562)	(5.069)	(0.309)
Sweden	2.210	5.239	5.534	3.103	2.250	5.651	25.447	52,033	71.487	22.924	85.069	1.938
	(0.070)	(0.250)	(1.675)	(1.525)	(0.363)	(0.891)	(0.717)	(3,018)	(1.484)	(0.953)	(4.506)	(0.108)
Switzerland	2.081	4.677	7.023	3.665	3.290	9.076	11.7580	73,682	80.162	197.694	111.039	1.974
	(0.058)	(0.239)	(1.014)	(0.833)	(0.398)	(1.559)	(0.371)	(3,302)	(1.175)	(9.779)	(12.333)	(0.109)
Turkey	-0.058	2.429	9.547	4.291	4.547	8.021	13.794	11,238	59.637	94.141	48.794	0.212
	(0.177)	(0.165)	(3.737)	(2.080)	(1.023)	(2.500)	(0.965)	(2,089)	(4.953)	(6.525)	(2.743)	(0.135)
United Kingdom	1.795	4.416	6.981	3.698	3.177	5.928	19.530	40,081	77.187	258.574	55.960	1.646
	(0.144)	(0.361)	(1.500)	(1.060)	(0.500)	(0.637)	(1.015)	(1,518)	(2.054)	(9.031)	(4.391)	(0.133)
United States	1.458	4.158	11.375	7.510	4.117	6.867	15.267	49,324	78.1	33.560	27.164	1.564
	(0.202)	(0.392)	(1.844)	(1.519)	(0.765)	(1.510)	(0.821)	(2,249)	(2.198)	(1.337)	(2.760)	(0.081)

Notes: CORWGI denotes the control of corruption. CORIRG stands for the corruption index. TEA denotes the total early stage entrepreneurial activity (% of 18-64 population). NASC is the nascent entrepreneurship rate (% of 18-64 population). NEW denotes the new business ownership rate (% of 18-64 population). EST denotes the established business ownership rate (% of 18-64 population). GOV denotes the government final consumption expenditure as a percentage of GDP, GDP denotes the GDP per capita as a proxy for total income, FREED is the level of economic freedom. POP denotes the population density. TRADE is the sum of exports and imports over GDP and finally GEFF denotes the level of government effectiveness.

References

Acemoglu, D., Johnson, S., and Robinson, J.A. (2005) "Institutions as a fundamental cause of long-run growth" *Handbook of economic growth*, Philippe Aghion and Steven Durlauf (ed), vol. 1, chapter 6, 385-472.

Aidis, R., Estrin, S., Mickiewicz, T.M., (2012) "Size matters: entrepreneurial entry and government" *Small Business Economics* **39(1)**, 119–139.

Aghion, P., Alesina, A., and Trebbi, F. (2007) "Democracy, Technology, and Growth" NBER working paper number 13180.

Aidis, R., Mickiewicz, T., (2006) "Entrepreneurs, expectations and business expansion: lessons from Lithuania" *European-Asian Studies* **58(6)**, 855–880.

Anokhin, S., and Schulze, W.S. (2009) "Entrepreneurship, innovation, and corruption" *Journal of Business Venturing* **24**, 465–476.

Baum, C. F., Schaffer, M. E., Stillman, S., (2003) "Instrumental variables and GMM: Estimation and testing" *Stata Journal* **3(1),** 1-31.

Baumol, W., (1990) "Entrepreneurship: productive, unproductive, and destructive" *Journal of Political Economy* **98(5)**, 893–921.

Berdiev, A.Z., and Saunoris, J.W. (2018) "Corruption and Entrepreneurship: Cross-country evidence from formal and informal sectors" *Southern Economic Journal*, **84(3)**, 831–848.

Dai, M., Liu Q., and Serfes, K. (2014) "Is the effect of competition on price dispersion nonmonotonic? Evidence from the U.S. airline industry" *Review of Economics and Statistics*, **96(1)** 161–170.

Davidson, R. and MacKinnon, J (1993) "*Estimation and Inference in Econometrics"* New York: Oxford University Press.

Dreher, A., Gassebner, M., (2013) "Greasing the wheels? The impact of regulations and corruption on firm entry" *Public Choice* **155(3–4)**, 413–432.

Goel, R.K., and Saunoris, J.W. (2019) "International corruption and its impacts across entrepreneurship types" *Managerial and Decision Economics* **40**, 475-487.

Hardle W., and Mammen E. (1993) "Comparing nonparametric versus parametric regression fits" *Annals of Statistics* **21**, 1926-1947.

Hausman, J (1978) "Specification Tests in Econometrics" *Econometrica* **46(6)**, 1251-1271.

Mauro, P (1995) "Corruption and Growth" *The Quarterly Journal of Economics* **110**, 681–712.

Mohamadi, A., Peltonen, J., and Wincent, J (2017) "Government efficiency and corruption: A country-level study with implications for entrepreneurship" *Journal of Business Venturing Insights* **8**, 50–55.

Papaioannou, S.K (2017) "ICT diffusion across US and EU industries: Does upstream regulation matter?" *Economics Letters* **156**, 102–105.

Robinson, P. (1988) "Root-n-consistent semi-parametric regression" *Econometrica* **56**, 931-54.

Rodriguez, P., Uhlenbruck, K., Eden, L (2005) "Government corruption and the entry strategies of multinationals" *Academy of Management Review* **30(2)**, 383–396.

Rose-Ackerman, S (2001) "Trust, honesty, and corruption: reflection of the state-building process" *European Journal of Sociology* **42**, 27–71.

Rose-Ackerman, S (2004) "The challenge of poor governance and corruption" Copenhagen consensus challenge paper. Copenhagenconcensus.com.

Swaleheen, M (2012) “Curbing corruption for higher growth: The impact of persistence” *Economics Letters* **116**, 255-257.
Verardi, V., and Debarsy, N (2012) “Robinson’s square root of N consistent semiparametric regression estimator in Stata” *Stata Journal* **12(4)**, 726-735.