EC831: Groupwork Contribution Statement

Group Number: 15

Registration Number	Contribution
HUZHE49005	I selected our research topic, set up the empirical model, collected and processed all the data, ran regressions, reported the empirical results and communicated with the tutor.
NIEQU21007	I collected a lot of information about the economic impact of corruption on the subject of our report, summarized and described the various impacts of corruption in developing countries, and communicated with other team members on time.
MAOJI46208	I participated in the selection and discussion of the subject, had some discussions on the collection and processing of data, proposed some improved methods, proposed some new exploration in regression analysis, and discussed with team members. At the same time, I focused on analyzing the beneficial economic effects of corruption on our research topic and critical review of Glaser and Sachs (2006).

Group 15

Members: Hu Zhentao (HUZHE49005)

Mao Jinghua (MAOJI46208)

Nie Quan (NIEQU21007)

Examine the relationship between corruption and economic growth

Does corruption hurt growth?

Word count: 6002

Abstract

This study aims to investigate the impact of corruption on economic growth across the states of the United States. We test the assertion by Glaser and Saks (2006) that the influence of corruption on economic growth in the U.S. is both negative and statistically insignificant. Utilizing cross-sectional data spanning a broader time period, we employ Ordinary Least Squares (OLS) regression and two-stage least squares (2SLS) regression to mitigate the effects of endogeneity. Our findings reveal that the negative relationship between corruption and economic growth becomes statistically insignificant after concerning the endogeneity of corruption levels, partly consistent with the conclusions of Glaser and Saks (2006). Moreover, by incorporating historical newspaper data and the distance from state capitals to Boise as instrumental variables, our research makes a novel methodological contribution to the study of corruption.

Contents

1. Introduction

The phenomenon of corruption, deeply rooted in human history, continues to captivate scholarly attention. Though the definitions of corruption were once diverse, they converged in recent years, understood as the misuse of public power for private benefit. This paper also adopts this widely accepted definition.

Despite advancements in conceptual clarity, empirically unravelling the economic ramifications of corruption remains a complex endeavor marred by contradictions. Existing literature implies that this paradox is partly due to the validity of the statistical methods employed to evaluate corruption levels. Most research relies on opinion surveys as the corruption indexes, restrained by their subjective nature, failing to provide convincing conclusions. As an alternative, Glaser and Saks (2006) use the number of state corrupt officials convicted by the federal court as the proxy for corruption levels in US states. This approach offers an objective criterion and mitigates concerns about the association with the quality of the judicial system and the level of corruption, as the federal court is relatively isolated from local corruption. This paper claims that the negative association between corruption and economic growth becomes insignificant when more controlled variables are included.

However, Glaser and Saks (2006) only ran an OLS regression with a short period of data without conducting tests to solve the endogeneity concern of the explanatory variable---- US states level of corruption. To address these gaps, this research aims to elucidate the relationship between corruption and economic growth in the United States by verifying the convictions put forth by Glaser and Saks (2006). To achieve this goal, we also adopted the conviction number as the explanatory variable but introduced data for a longer period. Additionally, we will optimize econometric techniques by introducing two instrumental variables---- the frequency of the keyword "corruption" and each state's name appearing in the historic US newspapers simultaneously and the distance from each state's capital to Boise. As the robust tests, we used the Boylan–Long corruption perception index and fixed effect model with panel data. We found that, unlike Glaser and Saks (2006), even more controlled variables are concluded in the OLS regression, the negative association between corruption and economic growth is still significant. However, when the data adopts the average value instead of the figure of the starting year and more econometric methods are applied to alleviate the endogeneity concern, this association becomes insignificant.

Our paper's main contribution lies in introducing reliable instrumental variables for corruption. It also indicates that theoretical explanations of why the US illustrates a unique relationship between corruption and economic growth compared to other developed countries are worth developing.

The remaining parts of this paper are structured as follows: Section 2 reviews the definition of corruption and the empirical study of the relationship between corruption and economic growth. Section 3 is a critical review of Glaser and Saks (2006). Section 4 describes and illustrates the data used in our empirical research. Section 5 discusses the methodology of our empirical study. Section 6 reports the results. Section 7 points out this paper's limitations and the last section is a conclusion.

2. Literature Review

2.1 Definition of corruption

In the classical ages, corruption was used to describe a dysfunctional polity. Plato describes corruption as a regime serving the interest of the rulers instead of being guided by the law (common good). This moralistic definition is also adopted by totalitarian regimes. They accuse people of being corrupt when they deviate from the official ideology (Friedrich, 2017). However, such moralistic definitions have been gradually replaced by revisionist definitions. They treat corruption as a series of value-free behaviours of exercising public power (Farrales, 2005).

Though the revisionist definitions became mainstream, disagreement on which aspect to emphasize was still prevailing in the 1960s. Heckelman (1970) concludes these definitions to be either public-office-centered, market-centered or public-interest-centered. Among them, the public-office-centered definitions focus on the abuse of public power to grab private interest. The market-centred definition describes a corruptive civil servant who regards his public office as a profit-maximizing business (Klaveren, 1970). The public-interest-centred definition emphasises corruption advances special interest at the cost of the public interest. Nevertheless, as Farrales (2005) points out, the public-office-centered definition has been widely accepted nowadays. The primary reason is that it is the most conservative definition that can be treated as a "least common denominator". Another reason is that later quantitative research and corruption indexes tend to adopt the public-office-centred definition. Our paper will also adopt this public-office-centered definition.

2.2 Corruption's economic impact on different regions

Much research suggests that corruption "sands the wheels of growth" in developing countries. For instance, Machoski and Araujo (2020) developed a model to evaluate the impact of corruption in public health on economic growth in Brazilian municipalities. They collect the data from the Office of the Comptroller General (CGU). It records corruption cases in the health and sanitation sectors among 180 Brazilian municipalities in 2009. They use the Ordinary Least Squares (OLS) and Quantile Regressions and find a negative result of public health corruption on economic growth in Brazilian municipalities. Though their interest is in public health corruption, as CGU's data concludes all kinds of corruption, they also conduct a preliminary analysis on the correlation between all types of corruption and the economic growth in these 180 Brazilian municipalities. The conclusion is similar to the situation in the public health sector that municipalities with higher corruption levels have lower economic growth rates. Also, Alfada (2019) analyzes the corruption data from the Indonesia Corruption Eradication Commission (KPK) to estimate the level of corruption in 19 Indonesian provinces. He finds that corruption has a deteriorating effect on economic growth and the negative effect is more significant when the corruption level is higher.

However, other papers indicate that corruption "greases the wheel" of economic growth and the anti-corruption measures tend to produce worse economic outcomes. Research carried out by Nam et al. (2020) on firms in Vietnam during the period from 2011 to 2015 proposes that non-native companies could better their performance and increase their lifespan by exploiting corruption. This could happen during the developmental stages of overseas companies, where they bribe government officials to ease the acquisition of lucrative projects and lower the cost of essential equipment in their production operations. For example, enterprises may opt not to purchase firefighting or environmental protection equipment that meets the required standards by bribing government officials. The resulting decrease in investment could then be used to expand the enterprise and create jobs, thus promoting regional economic growth. Similarly, Chen et al. (2021) discovered that the anti-corruption campaign initiated in China's Heilongjiang province in 2004 resulted in a negative economic effect on local and foreign businesses, as well as the overall economic development of the province.

The existing research's conclusions on the relationship between corruption and economic growth in developed countries are also diverse. Del Monte and Papagni (2001) performed an econometric analysis to evaluate public expenditure corruption's impact on economic growth in Italy. They include 28-year dynamic panel data among 20 Italian regions. It includes the level of public expenditure corruption (indicated by the number of convicted officials against public administration), real GDP growth rate, expenditure on public investment, and the amount of private investment, etc. They find corruption in public expenditure has a long-run negative impact on economic growth because this sort of corruption decreases the efficiency of public investment and leads to less money put into private sectors.

On the other hand, Glaeser and Saks (2006) proceeded with an empirical model to examine whether corruption would deter economic growth in developed countries. The authors took advantage of the number of corrupt state officials convicted by the United States Department of Justice to evaluate the level of corruption in each American state. They found that though one standard deviation of corruption would lead to a decline of one-fourth standard deviation in economic activities, this effect becomes insignificant when other control variables are included. In addition, Pluskota (2020) revealed a non-linear and non-monotonic correlation between corruption and economic growth in developed countries. The author collected the economic growth rates of 24 developed European countries and the World Bank's Corruption Index from 1996 to 2017 and found an open downward parabolic relationship between economic development and corruption. When the level of corruption has not reached a certain maximum value, there is a negative correlation between corruption and economic development. Beyond the maximum, there is a positive correlation between corruption and economic growth. However, many scholars suspect the World Bank's Corruption Index and other International Corruption Indexes' (ICIs) validity. As Anderson and Heywood (2009) and Donchev and Ujhelyi (2009) point out, ICIs are composite indexes that aggregate a series of surveys mainly responded to by Western leaders and experts. As a result, they tend to measure corruption perceptions, but not necessarily corruption experience. Beyond that, the ICIs' number of sub-indicators is different for different countries and the same country in different years (Gilman, 2018). It harms the ICIs' impartiality and long-term consistency.

2.3 Explanations of corruption's diverse economic impact on different regions

Aside from research errors, why does corruption have such diverse influences on different regions? Ang (2020) classifies corruption into four types: petty theft, grand theft, speed money and access money. Petty theft refers to the stealing or misuse of public funds among low-level bureaucrats. Grand theft refers to political elites' embezzlement or misappropriation of large sums of public money. Speed money means the petty bribes paid to low-level bureaucrats to speed things up. Access money is a collusion between political elites and tycoons. The tycoons share the profit and the political elites grant exclusive access and valuable privileges. He contends that the first two corruption types damage economic growth, the speed money enhances efficiency but leads to extra expenses and access money stimulates the economy but generates distortions, risks and inequality. Then he concludes that different regions have their primary types of corruption which makes a difference between the correlation of corruption and economic growth. However, the distinctions between different types of corruption are not clear. As a result, this theory is hard to be verified or falsified by further quantitative research. Also, the scope of this theory's arguments is limited to China which casts its generalizability into doubt.

Another reasonable explanation is the threshold theory. As Bose et al. (2007) point out, the correlation between corruption and economic growth is not monotonic. Corruption is likely to benefit economic growth when the corruption is lower than the threshold and it harms the economic growth when the level of corruption exceeds the threshold. But a significant flaw exists in its empirical research that it adopts the Corruption Perception Index (CPI) from 1995 to 2000 to evaluate the level of corruption. However, the CPI is not comparable before 2012 because the data included different components and time periods. (Gründler and Potrafke, 2019).

3. A critical review of Glaser and Saks (2006)

The article by Glaser and Saks (2006) investigates the determinants and consequences of corruption in the United States using a dataset of federal corruption convictions in each U.S. state. This work provides an empirical study of the factors associated with corruption and their impact on economic development in the United States, providing insights that complement the broader international discourse on corruption and governance.

Glaser and Saks find that states with higher levels of education and higher levels of income exhibit lower levels of corruption. This relationship persists even when historical factors are taken into account, suggesting that education plays a key role in reducing corruption. In addition, the authors find that corruption is positively correlated with income inequality and racial division, while there is no significant correlation with government size.

The analysis also reveals a weak negative correlation between corruption and state economic development, but this correlation becomes insignificant when more controlled variables are included. It uses the number of state-level corrupt officials convicted in federal court published by the Department of Justice, a more specific measure of corruption (as opposed to public opinion polls, which are commonly used in cross-national studies). In addition, unlike local justice systems, federal courts can be considered exogenous to corruption at the state level.

However, when studying the relationship between corruption and economic growth, the explanatory variable only includes the data from 1976 to 1980, a short period. All other controlled variables adopt their value of the starting year—1980 which may fail to represent the feature of the whole period from 1980 to 2000. In the meantime, the empirical methodology is questionable. It only conducts an OLS regression without any endogeneity test of the explanatory variable--- the conviction number. In addition, this article examines data from a specific time period, but corruption and its determinants may evolve. Analyses of recent trends or longitudinal changes in the determinants of corruption can provide insights into how the relationship identified by Glaser and Saks has held up over time.

To solve the drawbacks above and verify this paper's causal convictions between corruption and economic growth, we will make improvements on both data selection and econometric methods. Our study will include the latest data and statistics over a longer time horizon. Unlike this paper, which used only OLS regression, we will introduce instrumental variables of corruption and apply a fixed effect model to solve the endogeneity concern.

Since the main purpose of our paper is to alleviate the endogeneity concern of the explanatory variable----level of corruption, the paragraphs below will discuss the econometric methods that have been used to mitigate this endogeneity concern. For example, Tao et.al introduces a time-vary Difference-in-Differences (DID) model and selects whether a province has senior officials under investigation as the treatment effect to evaluate the level of corruption. Klašnja (2015) attempts to solve the endogeneity by Regression Discontinuity Design (RDD). More specifically, in Romania, mayors' salaries connect to the population threshold. So, the cities with a population right under and above the threshold are similar except for the mayors' salaries. Also, the existing literature has shown that higher salaries lead to lower corruption (Becker and Stigler, 1974; Besley, 2004). Thus, this paper adopts the discontinuity of mayors' salaries as the proxy for the level of mayoral corruption.

For instrumental variables of corruption level, however, most paper adopts the corruption level in the past few years or some International Corruption Indexes. Nevertheless, they are still likely to be endogenous to the omitted variables, thus failing to fulfil the exclusion restriction. To make up for this shortcoming, our paper introduces two new instrumental variables based on Ramirez (2014) and Geol and Nelson (2010). More details of these two IVs will be discussed in section 5.2.

4. Data

4.1 Variable Description and Data Sources

The dependent variable, Ln (Real state gross product 1990-2020), is the natural logarithm of real state gross product in 2020 divided by the equivalent figure in 1990. This data is collected from the US Bureau of Economic Analysis. Following the footsteps of Glaser and Saks (2006), we adopted the number of state-level corrupt officials per 100,000 people convicted by the federal court from 1990 to 2020 as the main explanatory variable. This data is accessed from the Justice Department's "Report to Congress on the Activities and Operations of the Public Integrity Section".

We use two instrumental variables. The first instrumental variable follows Ramirez (2014), which is the average frequency of the word "corruption" and each state's name appearing in one historical newspaper simultaneously per 100,000 people from 1900-1940. The main data of this instrument is from https://www.newspapers.com/, which collects over 900 million historical newspapers. We entered "corruption" and each state's name as the keyword to estimate the level of corruption of a state in a period of historical time. We exclude Indiana and Oregon for the reason of outliers. We also rule out Alaska and Hawaii for they didn't join the US until 1959. In addition, as the state of Washington is confounded by Washington DC, we do not include it in our regression. Thus, the sample size is 45.

Following Geol and Nelson (2010), the other instrumental variable we adopt is the distance between each state's capital to Boise. We get this data from Google Maps. We excluded Hawaii and Alaska as they are classified into the West region but they are distant from western America.

To verify Glaser and Saks (2006), we include all the controlled variables in "Table 6" of this paper (which has been attached to the Appendix as Table 9), but we adjusted the period of some variables. The main adjustment is that Table 9 includes conviction data from 1976-1980, while we covered the conviction numbers from 1990-2020. The variable Gross State Product 1990 refers to the total gross state product in 1990. Income and population are the median household income and total population by state. All three variables' values above are collected from the Bureau of Economic Analysis. South, Mideast, and Midwest are dummy variables to show the region of a state. The criteria of this division follow U.S. Census Bureau regions and divisions. Income inequality is the Gini coefficient by state. Share black and Urban are the proportion of Black people and the urbanization rate by state. All six variables (South, Mideast, Midwest, Income equality, share black and Urban) above are from the United States census. Population refers to the population size by state. Share gov. employment is the total government employees by state and share with 4+ years of college is the percentage of people who have a bachelor or higher degree by state. These three data are from the Federal Reserve Economic Data. To solve the concern that the number of convictions by the federal court is decided by both a state's level of corruption and the enforcement efforts, we take federal employment by state into account. This data is also from the Federal Reserve Economic Data.

As the robustness check, we use the Boylan–Long corruption perception index, a perception corruption index, to replace the number of convictions as the explanatory variable. This index is based on a 1992 survey in which reporters were asked to score 1 to 7 among a series of questions on their living state's corruption level compared to other states. "1" refers to the reporters who think their living state is the least corrupt state in the US and vice versa. More details about this index can be found in Boylan and Long (2003).

We also collected panel data to conduct a fixed-effect regression. The dependent variable, Ln (real GDP growth 1990-2020), is the natural logarithm of the annual real GDP growth rate among states in the US from 1990 to 2020. As this data is only available after 2006, we collected the real GDP by state in the US from the Bureau of Economic Analysis and the annual inflation rate from the World Bank. Based on these two data sources, we calculated the values of this dependent variable. All other variables in the fixed-effect model have been mentioned above. The only difference is all these variables' values are annual data from 1990-2020. All these data are collected from the Federal Reserve Economic Data and Bureau of Economic

Analysis. However, we included fewer variables compared to the IV regression as some annual data are not accessible which will be elaborated in section 6.1.1.

4.2 Descriptive Statistics

Table 1.1 Summary of cross-sectional data

Variable	N	Mean	SD
One-year data			
Ln (Gross state product 1990)	50	11.09691	1.041261
Share with 4+ years of college 1990	50	0.19756	0.0374668
Share black 1990	50	0.09534	0.0924933
Income inequality	50	0.4692	0.0182768
Gov employees 1990	50	369.0493	388.1743
Urban 1990	50	68.18	14.67142
Federal employees 1990	50	60.01598	64.56294
Log (Population) 1990	50	8.029656	1.018083
Racial fraction 2020	50	0.4708695	0.1510188
Urban 2010	50	73.59	14.56517
Perception index 1992	46	3.486848	1.156422
Average data			
Ln (Real state gross product 1990-2020)	50	1.270865	.2295173
Corruption 1990-2020	50	0.2944091	0.1675616
Share with 4+ years of college 2006-2020	50	0.26346	0.0465522
Gov employees 1990-2020	50	422.6404	437.7379
Federal employees 1990-2020	50	53.10477	53.40785
Unemployment 1990-2020	50	5.448371	1.040337
Data of Instrumental variable			
Newspaper frequency	45	1.281245	2.456035
Distance	47	7.60329	0.6055373

Table 1.2 Summary of panel data

Variable	N	Mean	SD
Panel data			
Annual real GDP growth1990-2020	1550	0.0205331	0.0350085
Ln (Annual real GDP growth1990-2020)	1211	-3.788257	0.9892343
Corruption 1990-2020	1531	16.96538	21.76399
Gov employees	1550	422.6404	436.1106
Federal employees	1550	53.10477	53.34043
Unemployment	1550	5.448371	1.888318
Midian household income	1550	46383.3	13345.8
Population	1550	5847.351	6481.672

As shown in Table 1.1 and Table 1.2, "One-year data" is the data by state in a certain year, "Average data" is the average value of a statistic by state in a period, "Panel data" is the annual data of a statistic by state in a period from 1990-2020 and "Data of Instrumental variable" refers to the average frequency of "corruption" and state's name appeared in newspapers from 1900-1940 and each state capital's distance to Boise.

5. Empirical Model

5.1 Baseline model

Our study aims to test whether empirical data supports a causal relationship between corruption and economic growth. More specifically, we will verify whether the conviction "In the US, corruption's impact on economic growth is negative but insignificant" in Glaser and Saks (2006) is valid. So, our baseline model follows Glaser and Saks (2006) OLS regression but introduces the latest data. The baseline model is:

$$Y = \alpha + \beta_1 \text{Conviction Rate}_{1990\text{-}2020} + \beta_2 \text{State Characteristics}_{1990} + U$$

Where Y refers to the natural logarithm of the real gross state product in 2020 divided by the equivalent figure in 1990. "Conviction Rate $_{1990\text{-}2020}$" is the average annual number of corrupt state officials convicted by the federal court per 100,000 people by state from 1990-2020. State Characteristics include a region dummy variable and 8 other controlled variables (i.e., income equality, urbanization rate, state population, etc.) and U is the error term. We focused on whether the explanatory variable β_1 is negative but insignificant.

5.2 Instrumental variables and Two-stage least squares regression

To solve the endogeneity concern especially the revere causality between GDP growth rate and corruption, our first attempt is to use a two-stage least squares (2SLS) regression.

The first instrumental variable of corruption we introduce refers to Ramirez (2014) which utilizes the historical newspaper search engine to track the frequency of "corruption" and related keywords' frequency in mainstream US newspapers to evaluate the level of historical corruption in the US. Based on that we added the keyword of each state's name to figure out the historic level of corruption in each state. We choose this as the instrumental variable for states' historic level of corruption is highly likely to correlate with their current corruption level and it appears that this instrument can only affect the economic growth rate through the current level of corruption. Nevertheless, more empirical tests will be conducted in section 6.2 to prove the exclusion restriction.

Also, as Geol and Nelson (2010) find, states belonging to the western part of America share significantly lower levels of corruption. This finding is also verified by our data. As shown in Table 2, states belonging to the West region have approximately 28% fewer conviction numbers than those states that do not belong to this region. In this way, the other instrumental variable we adopt is each state capital's distance to Boise, the capital of Idaho, which is located in the centre of western America. We expect this variable to be significantly positively correlated with the explanatory variable as western states are more adjacent to Boise.

Table 2. The average conviction numbers of states that belong or do not belong to the West region from 1990 to 2020

West	N	Mean
0	37	0.312086
1	13	0.244098
Total	50	0.294409

Note: "0" represents states that do not belong to Western America, and "1" represents states that belong to Western America.

In the first stage of the 2SLS, the Conviction Rate $_{1990\text{-}2020}$ is treated as the dependent variable and "frequency", "Log $_{distance}$", the region dummy variable and all other 8 controlled variables are independent variables. In the second stage, these two IVs, the dummy variable and other controlled variables are regressed to estimate the coefficient of the endogenous variable β_1.

5.3 Fixed effect model and the corruption perception index

As the robustness check, we adopt a two-factor fixed effect model for panel data from 1990 to 2020, as follows:

$$Y_{it}= \beta_0+\beta_1 C_{it}+\beta_2 X_{it}+\theta_i+V_t+U_{it}$$

Where "i" is the state and "t" is the year. For all states and years, Y_{it} is the annual growth rate of the real regional gross domestic product (GDP), C_{it} represents corruption levels measured by conviction numbers per 100,000 people, X_{it} is a vector containing several covariates, including a region dummy variable, and other controlled variables such as the aggregate state government employees, median household income and unemployment rate, etc. As the IV regression above, this model also aims to test whether C_{it} has a negative but insignificant correlation with the independent variable.

Geol and Nelson (2010) find that the result of the cross-state corruption study in the US alters when different measures of corruption are adopted. In response to this concern, another robustness test we take is to use the Boylan–Long corruption perception index as the explanatory variable. Like 5.1 and 5.2, in this robust test, we adopt the OLS and 2SLS regression model and include the same controlled variables. Again, we focus on whether the explanatory variable keeps an insignificantly negative relationship with economic growth.

6. Empirical findings

6.1 Baseline Results

In all regressions in Table 3, we use the natural logarithm of real state gross product in 2020 divided by the equivalent figure in 1990 independent variable, and the average number of convictions per 100,000 people as the explanatory variable. The Gini coefficient in US states in 1990 is not accessible, instead, we use this data in 2022 as the value of "Income inequality". As the unemployment rate fluctuates over the years, we adopt the average unemployment rate by state from 1990-2020. Except for these two variables, all other controlled variables adopted data in 1990. The definition and sources of these variables have been illustrated in section 4.

Table 3. Relation between corruption and real GDP growth 1990-2020 (Adopting starting year's data)

VARIABLES	(1) Ln (Real state gross product 1990-2020)	(2) Ln (Real state gross product 1990-2020)	(3) Ln (Real state gross product 1990-2020)
Corruption 1990-2020	-0.480**	-0.405**	-0.317*
	(0.184)	(0.186)	(0.181)
Ln (Gross state product 1990)		-0.734***	-0.634***
		(0.185)	(0.187)
South	-0.0878	-0.00917	-0.136
	(0.0793)	(0.112)	(0.115)
Northeast	-0.285***	-0.252**	-0.278***
	(0.0879)	(0.0962)	(0.100)
Midwest	-0.151*	-0.0974	-0.249**
	(0.0814)	(0.0884)	(0.100)
Share with 4+ years of college		2.073*	0.700
		(1.030)	(1.184)
Share black		-0.165	0.108
		(0.477)	(0.523)
Income inequality		0.0364	0.240
		(2.315)	(2.390)
Government employees		0.000117	0.000158
		(0.000141)	(0.000253)
Ln (Population 1990)		0.688***	0.632***
		(0.176)	(0.183)
Urban 1990		0.00526*	0.00589
		(0.00286)	(0.00354)
Federal employees			-0.000397
			(0.00140)
Unemployment rate			-0.0979***
			(0.0327)
Racial fraction			0.0162
			(0.391)
Constant	1.528***	3.275**	3.306**
	(0.0719)	(1.232)	(1.240)
Observations	50	50	50
R-squared	0.284	0.528	0.625

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The first and second columns of this table follow Table 9. The difference is we utilize a longer period and the latest data. For the first column, the result is quite similar to Table 9. Namely, states with a higher average number of corruption convictions in

1990-2020 are associated with lower average real GDP growth during this period. This effect is significant and large that a 0.1 increase in conviction numbers is correlated with a 4.8% decrease in average real GDP growth.

However, unlike Table 9, which indicates that when more controlled variables are introduced, average conviction numbers' impact on real GDP growth becomes small and insignificant. Even though we include these control variables, this correlation is almost as significant and large as column (1). In column (3), when we took the federal government employees' number by the state in 1990, the average unemployment rate from 1990-2020 and the racial fraction in 2020 into account, the association remains significant at a 10% level, though the coefficient drops by around 28%.

For the concern that the data of the starting year may not represent situations of the whole period, Table 4 concludes with more average data. Specifically, "Government employees" and "Federal employees" adopt the average numbers from 1990 to 2020. For the data accessibility concern, "Share with 4+ years of college" is from 2006-2020. "Population" is the average population by state from 1990-2020. "Urban" is the urbanization rate in 2010 according to the 2010 US Census. Also, we replace "Share Black" with "Racial fraction" calculated by the 2020 US Census.

Table 4. Relation between corruption and real GDP growth 1990-2020 （Adopting average or mid-year data）

VARIABLES	(1) Ln (Real state gross product 1990-2020)	(2) Ln (Real state gross product 1990-2020)	(3) Ln (Real state gross product 1990-2020)
Corruption 1990-2020	-0.480**	-0.154	-0.0695
	(0.184)	(0.129)	(0.127)
Ln (Gross state product 1990)		-0.843***	-0.768***
		(0.105)	(0.102)
Northeast	-0.285***	-0.135*	-0.156**
	(0.0879)	(0.0765)	(0.0721)
Midwest	-0.151*	-0.00271	-0.100
	(0.0814)	(0.0619)	(0.0659)
South	-0.0878	0.0218	-0.0325
	(0.0793)	(0.0616)	(0.0632)
Share with 4+ years of college		2.134***	1.575**
		(0.547)	(0.669)
Racial fraction		-0.198	-0.138
		(0.230)	(0.213)

Income inequality		0.0944	-0.00833
		(1.598)	(1.522)
Government employments		0.000184**	0.000259*
		(8.22e-05)	(0.000139)
Ln (Population)		0.763***	0.737***
		(0.103)	(0.0978)
Urban 2010		0.00686***	0.00684***
		(0.00241)	(0.00222)
Federal employees			-0.00106
			(0.00112)
Unemployment rate			-0.0638***
			(0.0227)
Constant	1.528***	-1.928**	-1.802**
	(0.0719)	(0.747)	(0.695)
Observations	50	50	50
R-squared	0.284	0.780	0.823

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We found that after the adjustment of the data selection, the correlation between average conviction numbers and average real GDP growth became weaker and less significant. In column (2), the coefficient is only one-third of the equivalent coefficient in Table 3 and the correlation is insignificant. When more controlled variables are included like column (3) illustrated, the correlation becomes smaller and insignificant. These diverse baseline results indicate the necessity for an endogeneity test of the explanatory variable.

6.2 Two-stage least squares regression

For the 2SLS regression, "Newspaper frequency" and "Ln (Distance to Boise)" are two instrumental variables. "Newspaper frequency" is the average number of both the word "corruption" and a certain state appearing in one newspaper per 100,000 annually. The period is from 1900-1940. "Ln (Distance to Boise)" refers to the natural logarithm of each state capital's distance to Boise in kilometres. All other variables adopt the same data source as Table 4.

Table 5. Relation between corruption and real GDP growth 1990-2020 (IV regression)

VARIABLES	(1) second stage frequency Ln (Average real GDP growth1990-2020)	(2) second stage frequency and distance Ln (Average real GDP growth1990-2020)
Corruption1990-2020	0.2948	0.2525

	(0.278)	(0.200)
Ln (Gross state product 1990)	-0.8304***	-0.8309***
	(0.108)	(0.107)
Northeast	-0.0696	-0.0720
	(0.069)	(0.067)
Midwest	-0.0951	-0.0927
	(0.065)	(0.062)
South	-0.0284	-0.0270
	(0.061)	(0.059)
Racial fraction	-0.0500	-0.0550
	(0.199)	(0.197)
Unemployment	-0.0989***	-0.0979***
	(0.021)	(0.020)
Urban 2010	0.0090***	0.0090***
	(0.002)	(0.002)
Gov employees	0.0003***	0.0003***
	(0.000)	(0.000)
Federal employees	-0.0016	-0.0015
	(0.001)	(0.001)
Income inequality	-1.1949	-0.9904
	(1.994)	(1.683)
Share with 4+ years of college	1.2177**	1.1900**
	(0.592)	(0.571)
Ln (Population)	0.8027***	0.7990***
	(0.106)	(0.102)
Constant	-1.6033**	-1.6225**
	(0.757)	(0.737)
Observations	45	44
R-squared	0.812	0.816
OverID test		0.0689
		(0.793)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

When we only include "Newspaper frequency" as the instrumental variable, as shown in the first column of Table 10 (Seen in the Appendix), the first-stage F statistic is strong. It demonstrates that states with higher corruption levels in history are likely to be more corrupt today. Then, as shown in Table 10 Column (2), we include both instrumental variables in the first stage of regression. As we expected, both instruments are significantly related to the explanatory variable.

In the first column of Table 5, where we use corruption's frequency in historic newspapers as an instrument for the current corruption level. We find that a 0.1 increase in corruption numbers relates to a 2.95% increase in real GDP growth. However, this positive relationship is insignificant. When we include both instruments, as displayed in column (2), this insignificant positive relationship remains. The overidentification test attached at the bottom of this table shows that the instrumental variables are irrelevant to the error term in the second-stage regression. Thus, these two instruments do not have independent impacts on the dependent variable and meet the exclusion restriction.

However, when we utilize different newspaper search engines such as Proquest and Library of Congress to evaluate the level of historic corruption by state, the results in the first stage regression are different from what we have reported in Table 10 Column (1). The differences are possibly due to the diverse mechanisms of the newspaper search engines and their data sources. Also, these differences raise doubts about whether the frequency of "corruption" and a certain state name in Newspaper.com can reflect US historic corruption by state. All of these above indicate that research on the mechanism of these newspaper engines is valuable for studies based on keyword searching.

6.3 Robustness Check

6.3.1 Fixed Effect Regression

Given dummy variables of the region are not compatible with the fixed effect model, variables like "South", "Midwest", and "Northeast" are not included. Also, "Racial fraction", "Urban", "Income equality", etc. are excluded as these annual data from 1990 to 2020 are not accessible. Instead, we introduced "Midian household income" which is annual the median household income by state. All other variables follow the same criteria above except that they are panel data.

Table 6 Relation between corruption and real GDP growth 1990-2020 (Panel data)

VARIABLES	(1) Ln (Annual real GDP growth1990-2020)	(2) Ln (Annual real GDP growth1990-2020)	(3) Annual real GDP growth1990-2020	(4) Annual real GDP growth1990-2020
Corruption1990-2020	-0.00254750	-0.00201428	-0.00005158	-0.00000189
	(0.002)	(0.002)	(0.000)	(0.000)
Gov employees 1990-2020		-0.00321490**		0.00014552***
		(0.002)		(0.000)
Federal employees		-0.00396613		-0.00028584**

		(0.004)		(0.000)
Unemployment		0.24860748***		0.00890397***
		(0.033)		(0.001)
Midian income		0.00000355		0.00000018
		(0.000)		(0.000)
Population		0.00020668***		0.00000644***
		(0.000)		(0.000)
Constant	4.29284016***	2.73378613***	0.00159822	0.08175624***
	(0.189)	(0.516)	(0.004)	(0.014)
Observations	1,193	1,193	1,531	1,531
R-squared	0.267	0.309	0.358	0.410
Number of states	50	50	50	50

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Columns (1) and (2) in Table 6 show that corruption conviction numbers and the natural logarithm GDP growth rate only have a very small and insignificant relationship with a 0.1 increase in corruption conviction numbers associated with 0.02% to 0.025% decrease in the state GDP growth rate.

For the natural logarithm of Growth GDP growth rate excluded all the data of negative economic rate, 339 samples are excluded. What's worse, these samples are not randomly distributed as all of them represent negative economic growth rates. To solve this sample bias, we utilize "State GDP growth rate" as the dependent variable as columns (3) and (4) show. Similar results are presented that the negative relationship between corruption and state GDP growth rate is small and insignificant.

6.3.2 Perception Corruption Index

As shown in Table 7, we adopt the Boylan–Long corruption perception index as the explanatory variable and all other variables utilize the same data as Table 3.

Table 7 Relationship between corruption index and Real GDP growth rate (Adopting starting year's data)

VARIABLES	(1) Ln (Average real GDP growth1990-2020)	(2) Ln (Average real GDP growth1990-2020)	(3) Ln (Average real GDP growth1990-2020)
Perception index	-0.0545*	-0.107***	-0.0852**
	(0.0304)	(0.0322)	(0.0316)
Ln (Gross state product 1990)		-0.708***	-0.667***

		(0.179)	(0.182)
South	-0.109	0.0456	-0.0664
	(0.0854)	(0.111)	(0.117)
Northeast	-0.304***	-0.135	-0.164
	(0.107)	(0.105)	(0.111)
Midwest	-0.188**	-0.151*	-0.261**
	(0.0870)	(0.0846)	(0.0953)
Share with 4+ years of college		0.353	-0.395
		(1.125)	(1.225)
Racial Friction		-0.177	-0.101
		(0.479)	(0.513)
Income inequality		-1.013	-1.401
		(2.232)	(2.312)
Gov employees		-1.26e-05	0.000120
		(0.000141)	(0.000248)
Ln (Population)		0.710***	0.706***
		(0.168)	(0.176)
Urban 1990		0.0121***	0.0101**
		(0.00334)	(0.00381)
Federal employee			-0.000840
			(0.00132)
Unemployment			-0.0828**
			(0.0327)
Racial fraction			0.234
			(0.386)
Constant	1.588***	3.438**	3.807***
	(0.116)	(1.265)	(1.285)
Observations	46	46	46
R-squared	0.247	0.607	0.683

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Like the result in Table 3, when more controlled variables are included, the relationship between the corruption index and real GDP growth rate is still significantly negative, though the absolute values of the coefficients become smaller compared to those in Table 3.

Then we run the 2SLS regression. Except for the endogenous variable is replaced by the Boylan–Long corruption perception index, all other variables are the same as in Table 5.

Table 8
Relation between the corruption index and real GDP growth 1990-2020 (IV regression)

	(1)	(2)
	second stage frequency	second stage frequency and distance
VARIABLES	log_average9020	log_average9020

Perception index	-0.6847	-0.1398
	(2.931)	(0.136)
Ln (Gross state product 1990)	0.1478	-0.6137***
	(4.049)	(0.205)
Northeast	0.3759	0.0149
	(1.961)	(0.112)
Midwest	-0.2252	-0.1093
	(0.724)	(0.070)
South	0.3817	0.0481
	(1.794)	(0.098)
Racial fraction	0.6627	0.1439
	(2.867)	(0.235)
Unemployment	0.1121	-0.0517
	(0.889)	(0.046)
Urban 2010	0.0367	0.0141**
	(0.123)	(0.006)
Unemployment	-0.0008	0.0000
	(0.004)	(0.000)
Federal employees	0.0009	-0.0004
	(0.008)	(0.001)
Income equality	3.1530	0.2820
	(15.870)	(1.466)
Share with 4+ years of degree	-7.2926	-0.8722
	(34.759)	(1.726)
Ln (Population)	-0.0990	0.5821***
	(3.616)	(0.183)
Constant	0.5435	-0.9300
	(8.996)	(0.889)
Observations	41	40
R-squared		0.826

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Again, when “frequency” and “distance” are introduced as the instrumental variables, the relationship between the corruption index and real GDP growth rate becomes insignificant. However, this IV regression has a main drawback that the instruments (especially for “frequency”) are insignificantly connected with the endogenous variable (as shown in Table 11). This leads to a worry about weak instruments and biased regression results (Angrist et al., 2008).

7. Limitations

The primary worry of our study is whether the frequency of the keyword “corruption” and the state’s name appearing in the historic US newspapers can evaluate the historic level of corruption by state. The value of this instrument we reported is from “Newspaper.com”, which shows a strong positive relationship with the number of US current corruption situations. However, we also utilize other newspaper search engines like Proquest and the Library of Congress to get the historical corruption data. We found the data from these two search engines only shows an insignificant relationship with the endogenous variable. Thus, doubt may be cast out on whether this instrument is a valid proxy for the US historic corruption level by state. Unfortunately, we cannot respond to this query as we are not able to dig out the internal mechanism of these newspaper search engines. Further research on search engines is highly valuable as it can contribute to a wide range of studies based on keyword searching. Nevertheless, this needs an inter-disciplinary effort especially assistance from computer science.

Another potential flaw is in section 6.3.2. The two instrumental variables we use are insignificantly connected with the Boylan–Long corruption perception index, leading to weak instruments concern. As a result, the coefficient of the explanatory variable is likely to be biased. Finally, when we collect the panel data and attempt to apply the factors fixed effect model to solve the endogeneity, only five controlled variables are included for the constraint of data availability. Consequently, this model can only explain part of the GDP's real growth rate. This regression can be optimized if variables like the annual Gini coefficient, urbanization rate, racial fraction and share of 4+ years of college by state are taken into account. But till now, only data on the share of 4+ years of college by state have been published since 2006, and all other categories' annual data remain inaccessible.

8. Conclusions

In this paper, we examined corruption’s impact on economic growth across U.S. states based on Glaser and Saks (2006). Our study verifies Glaser and Saks’ (2006) conviction that in the US corruption’s modest negative impact on economic growth disappears once a rich enough set of covariates is controlled. We discovered that when econometric methods like instrumental variables and fixed effect regression are introduced, the correlation between corruption and economic growth becomes insignificant.

Glaser and Saks (2006) use the number of corrupt officials convicted by the federal court as the proxy for corruption and run an OLS regression model. As the baseline study, we adopted this model and introduced data from a wider time period. Unlike

Glaser and Saks (2006), we found that even though more controlled variables are included, the significant negative relationship stands. However, when we use more average data instead of the starting year data, this relationship becomes less insignificant. Then, we ran a two-stage least squares (2SLS) regression by introducing two instrumental variables. In this regression, corruption has an insignificant but positive impact on economic growth.

Referring to Geol and Nelson's (2010) comments on the cross-state corruption study, we adopt the Boylan–Long corruption perception index as the explanatory variable. We ran a similar OLS and IV regression and found similar insignificant results. Finally, we incorporated panel data into the two-factor fixed effect model. Once again, the correlation between corruption and economic growth is insignificant.

Our paper contributes to the literature by introducing two reasonable instrumental variables of corruption. Though the corruption level in the past few years and some International Corruption Indexes have been used as the instrumental variables of corruption, they are likely to be endogenous to the omitted variables and fail to meet the exclusion restriction. On the other hand, we use information from historical newspapers and each state capital's distance to Boise as the instruments. They are more irrelevant to omitted variables by theory and the exclusion restriction is also supported by empirical tests.

Our project also indicates that more meticulous research on newspaper search engines is highly valuable. Also, theoretical explanations on why the US shows a different relationship between corruption and economic growth compared to many other developed countries are still worth developing.

9. Appendix

Table 9

Relationship between corruption 1976–80 and economic growth 1980–2000

	Gross state product		Median household income		Employment		Median house value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Corruption 1976–80	−.641*	.032	−.221**	−.108	−.124	.249	−.126	−.256
	(.360)	(.332)	(.106)	(.102)	(.196)	(.234)	(.199)	(.216)
Ln(Gross state product)	.048	−.390						
	(.039)	(.304)						
Ln(Income)		−.974	−.074	−.646**		−.937**		−.373
		(.826)	(.089)	(.124)		(.300)		(.291)
Ln(Employment)					.006	.397		
					(.022)	(.325)		
Ln(House Value)							−.021	−.326**
							(.108)	(.159)
South	.058	.018	.062*	.055	−.115*	−.076	.073	−.074
	(.105)	(.122)	(.035)	(.037)	(.059)	(.082)	(.076)	(.083)
Northeast	.074	−.176	.134**	.102**	−.224**	−.304**	.221**	.126**
	(.115)	(.105)	(.034)	(.031)	(.064)	(.071)	(.069)	(.069)
Midwest	−.125	.107	.022	.070**	−.212**	−.189**	.003	−.062
	(.107)	(.104)	(.032)	(.030)	(.060)	(.069)	(.072)	(.072)
Share with 4+ years of college		3.40*		1.97**		1.03		1.57
		(1.73)		(.527)		(1.24)		(1.08)
Share black		.796		.355**		.172		.656**
		(.560)		(.166)		(.374)		(.340)
Income inequality		−7.83		−2.67**		−6.05**		−3.66*
		(4.80)		(.958)		(2.23)		(1.98)
Share gov. employment		−1.69		.112		−.394		.854
		(1.24)		(.377)		(.865)		(.778)
Ln(Population)		.347		.001		−.411		.038
		(.292)		(.013)		(.314)		(.027)
Share urban		.828**		.200**		.425**		.352**
		(.268)		(.071)		(.161)		(.156)
Constant	.816	12.6	.813	7.24**	.444	12.6**	.331	7.93**
	(.394)	(10.9)	(.920)	(1.47)	(.300)	(3.60)	(1.24)	(3.02)
Adj. R^2	.03	.49	.25	.58	.22	.35	.17	.44
# Obs.	50	50	50	50	50	50	50	50

Note. Except for the corruption rate, all independent variables are measured in 1980. Gross state product and total employment are from the Bureau of Economic Analysis. Income is median household income from the 1980 Census, and housing values are the median value in the 1980 Census.

Table 10. First-stage 2SLS regression (The endogenous variable is Corruption 1990-2020)

VARIABLES	(1) Corruption 1990-2020	(2) Corruption 1990-2020
Ln (Gross State Product 1990)	-0.0802	0.0570
	(0.178)	(0.184)
Northeast	-0.0227	0.292
	(0.107)	(0.174)
Midwest	0.0991	0.288**
	(0.0933)	(0.124)
South	0.00782	0.293*
	(0.0944)	(0.157)
Racial fraction	-0.108	-0.0244
	(0.314)	(0.305)
Unemployment	0.0225	0.0446
	(0.0310)	(0.0312)
Urban2010	-0.00219	-0.00394
	(0.00324)	(0.00320)
Gov employees	-0.000320*	-0.000411**
	(0.000181)	(0.000177)
Federal employees	0.00332**	0.00332**
	(0.00140)	(0.00134)
Income inequality	6.117***	5.117**
	(2.023)	(1.992)
Share with 4+ years of college	-0.377	-0.0864
	(0.899)	(0.868)
Ln (Population)	0.0120	-0.0750
	(0.169)	(0.169)
Newspaper frequency	0.0254**	0.0246**
	(0.0111)	(0.0107)
Ln (Distance to Boise)		-0.252**
		(0.113)
Constant	-1.776	0.133
	(1.282)	(1.494)
Observations	45	44
R-squared	0.495	0.565

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 11. First-stage 2SLS regression (The endogenous variable is the Perception index)

VARIABLES	(1) Perception index	(2) Perception index
Ln (Gross state product 1990)	1.429	0.943
	(1.077)	(1.216)
Northeast	0.652	-0.0686
	(0.658)	(1.153)
Midwest	-0.231	-0.696
	(0.556)	(0.795)
South	0.628	-0.0384
	(0.580)	(1.019)
Racial fraction	0.957	0.534
	(1.983)	(2.071)
Unemployment	0.302	0.257
	(0.187)	(0.201)
Urban2010	0.0422**	0.0469**
	(0.0197)	(0.0208)
templayment9020	-0.00148	-0.00123
	(0.00110)	(0.00116)
Gov employees	0.00187	0.00199
	(0.00879)	(0.00897)
Income inequality	4.593	5.799
	(13.00)	(13.70)
Share with 4+ years of college	-11.88**	-12.58**
	(5.615)	(5.789)
Ln (Population)	-1.294	-0.927
	(1.015)	(1.104)
Newspaper frequency	-0.0119	-0.00285
	(0.0674)	(0.0696)
Ln (Distance to Boise)		0.563
		(0.742)
Constant	3.327	-1.028
	(8.075)	(10.03)
Observations	41	40
R-squared	0.657	0.663

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

10. References

Alfada, A. (2019) ‘The destructive effect of corruption on economic growth in Indonesia: A threshold model’, Heliyon, 5(10).

Andersson, S. and Heywood, P.M. (2009) ‘The politics of perception: Use and abuse of Transparency International’s approach to measuring corruption’, Political Studies, 57(4), pp. 746–767.

Ang, Y. Y. (2020). China's gilded age: The paradox of economic boom and vast corruption. 1st ed. Cambridge: Cambridge University Press.

Angrist, J. D. et al. (2008) Mostly harmless econometrics: An empiricist’s companion. 1st edition. New Jersey: Princeton University Press.

Becker, Gary S., and George J. Stigler. 1974. “Law Enforcement, Malfeasance, and Compensation of Enforcers.” Journal of Legal Studies 3 (1), pp. 1–18.

Besley, T. (2004) ‘Paying politicians: Theory and evidence’, Journal of the European Economic Association, 2(2–3), pp. 193–215.

Bose, N. et al. (2008) ‘Threshold effects of corruption: Theory and evidence’, World Development, 36(7), pp. 1173–1191.

Boylan RT, Long CX (2003) Measuring public corruption in the American States: a survey of state house reporters. State Politics Policy 3:420–438

Chen, Z. et al. (2021) Is a Corruption Crackdown Really Good for the Economy? Firm-Level Evidence from China. Journal of law, economics, & organization. [Online] 37 (2), 314–357.

Del Monte, A. and Papagni, E. (2001) ‘Public expenditure, corruption, and economic growth: The case of Italy’, European Journal of Political Economy, 17(1), pp. 1–16.

Donchev, D.D. and Ujhelyi, G. (2009) ‘What do corruption indices measure?’, SSRN Electronic Journal [Preprint].

Farrales, M.J. (2005) ‘What is corruption? A history of corruption studies and the great definitions debate’, SSRN Electronic

Friedrich, C.J. (2017) ‘Corruption concepts in historical perspective’, Political Corruption, pp. 15–24.

Gilman, S.C. (2018) ‘To understand and to misunderstand how corruption is measured: Academic research and the corruption perception index’, Public Integrity, 20(sup1).

Glaeser, Edward, Saiz, Albert. (2004) The rise of the skilled city. Brookings-Wharton Papers on Urban Affairs 5.

Glaeser, E. and Saks, R. (2006) ‘Corruption in the United States of America’, Global Corruption Report 2006, pp. 308–311.

Goel, R.K. and Nelson, M.A. (2010) ‘Measures of corruption and determinants of US corruption’, Economics of Governance, 12(2), pp. 155–176.

Gründler, K. and Potrafke, N. (2019) ‘Corruption and economic growth: New empirical evidence’, SSRN Electronic Journal [Preprint].

Heckelman, J.C. and Powell, B. (2010) ‘Corruption and the institutional environment for growth’, Comparative Economic Studies, 52, pp.351-378.

Heidenheimer, A. J. (ed) (1970) ‘Political Corruption: Readings in Comparative Analysis.’ New York: Holt, Rinehart and Winston, Inc.

Hoinaru, R., Buda, D., Borlea, S.N., Văidean, V.L. and Achim, M.V. (2020) ‘The impact of corruption and shadow economy on the economic and sustainable development. Do they “sand the wheels” or “grease the wheels?”’. Sustainability, 12(2), p.481

Klašnja, M. (2015). Corruption and the incumbency disadvantage: Theory and evidence. The Journal of Politics, 77(4), pp. 928-942.

Klaveren, V. (1970) ‘The Concept of Corruption’, in Arnold J. Heidenheimer (ed), Political Corruption: Readings in Comparative Analysis. New York: Holt, Rinehart and Winston, Inc.

Machoski, E. and de Araujo, J.M. (2020) ‘Corruption in public health and its effects on the economic growth of Brazilian municipalities’, The European Journal of Health Economics, 21(5), pp. 669–687.

Pluskota, A. (2020) ‘The impact of corruption on economic growth and innovation in an economy in developed European countries’, Annales Universitatis Mariae Curie-Skłodowska, sectio H – Oeconomia, 54(2), p. 77.

Nam, V.H et al. (2020) The impact of corruption on the performance of newly established enterprises: Empirical evidence from a transition economy. Borsa Istanbul review. [Online] 20 (4), 383–395.

Ramirez, C.D. (2014) ‘Is corruption in China “out of control”? A comparison with the US in historical perspective’, *Journal of Comparative Economics*, 42(1), pp. 76–91.

Tao, M., Goh, L. T., Zheng, Y., & Le, W. (2023) Do China's anti-corruption efforts improve corporate productivity? A difference-in-difference exploration of Chinese listed enterprises. Socio-Economic Planning Sciences, 87, 101594.