Statistical Manipulation: An Analysis of Potential Economic Data Tampering

Ryan Coutu
Stonehill College, United States

Oltiana Muharremi
Stonehill College, United States

www.ijonest.net

To cite this article:

Coutu, R. & Muharremi, O. (2024). Statistical manipulation: An analysis of potential economic data tampering. *International Journal on Engineering, Science, and Technology (IJonEST), 6*(3), 286-303. https://doi.org/10.46328/ijonest.221

International Journal on Engineering, Science, and Technology
2024, Vol. 6, No. 3, 286-303 https://doi.org/10.46328/ijonest.221

Statistical Manipulation: An Analysis of Potential Economic Data Tampering

Ryan Coutu, Oltiana Muharremi

Article Info

Article History

Received:
10 February 2024
Accepted:
24 May 2024

Keywords

Corruption
Autocracy
Statistics
Manipulation

Abstract

Information on a country's economy is important for many parties, including domestic politicians, potential investors, and international financial institutions. However, an individual country's official statistics are generated and published by the governing body, creating the opportunity for manipulation of these measures. Our research is based on the assumptions that all governments are incentivized to exaggerate economic metrics and that stricter measures on civil liberties, like speech and media, allow for this manipulation to remain. We identified two variables suitable for proxying economic growth, per capita Health Care Expenditure (HCE) and per person electricity consumption to compare to the stated economic data. We utilized a Fixed Effect-OLS regression to find the relationship between these proxy variables and Gross Domestic Product dependent on government type for 151 countries from 2000-2020. Our research revealed that, despite two government-type indicators, the relationship between electricity consumption and GDP was greater with more autocratic governments. This provides evidence to support that autocratic governments may manipulate their economic data. This research expands upon the literature on the role both corruption and government style have on GDP, as well as providing additional evidence that governments may have manipulated economic statistics.

Introduction

The calculation of economic statistics is not a perfect science. Individuals lack complete information about the current state of a country's economy and therefore these measures must be, to some extent, imperfect. Even metrics that are widely known and utilized, such as inflation and Gross Domestic Product (GDP), are flawed either theoretically, in calculation, or both (Silver, 2007; Ivković, 2016). However, governments themselves provide these estimates on economic performance, with an overall lack of alternatives due to limited incentives and the inability to produce such metrics by third parties. Therefore, we are reliant on government-produced statistics to measure the effectiveness of said government.

Furthermore, economic performance and political survival are inextricably linked. Several studies have shown that tampering with GDP figures can result in economic and political inefficiencies. Economically, a country's growth estimate influences the operations of other agents (trade partners, multilateral organizations, foreign

investors) who may be ineffective at extracting the information content of official statistics. Formal accountability models demonstrate that incumbents have a strong political motive to alter the information available to citizens (Gehlbach et al., 2016). As a result, the overestimation of GDP growth is likely to stymie political responsibility and worsen overall governance. GDP estimates are frequently produced by government-affiliated statistical agencies as preliminary estimates are developed quarterly and updated as new information becomes available. As any fundamental macroeconomics student is aware, GDP can be estimated in a variety of ways, including the sum of expenditures, the sum of incomes, or the amount of value-added across the economy. National statistics bodies collect data from various sources to develop such estimations. Sectoral surveys, manufacturing and agricultural censuses, and household surveys are among these sources. They also include data from banks, utilities, transportation companies, and other levels of government. While comprehensive, these measures are by no means perfect.

The problem arises in the fact that these statistics play large roles in the global economic and political sphere. Countries are ranked on their economic performance, these statistics can influence whether politicians stay or are removed from office, and these metrics help dictate foreign investments. Government entities are incentivized to improve economic metrics through both legitimate (implementing productive policies) and illegitimate (corruption and manipulation) ways.

Given the assumption that an incentive to exaggerate economic performance is ever-present, we expect that a well-functioning democracy can rein in, to some extent, the impulse to manipulate official statistics. A strong democracy, with a more open and examinable government structure, allows political opponents and the media the opportunity to scrutinize government figures. Furthermore, an independent judiciary can investigate and prosecute those who attempt to fiddle with the numbers. There is, in essence, a built-in system of checks and balances that works to keep data manipulation in line. More autocratic governments, with stricter government hierarchies and limitations of outside investigations, lack the system of checks and balances that work to reign in potential manipulation. We expect that data exaggeration to be more common in autocratic governments as the ability of third parties to further investigate these statistics is more limited than in democratically run governments.

In this study, we look to find evidence of increased manipulation of economic statistics by autocratic governments through the analysis of "soft" economic metrics (GDP) and proxy metrics (per capita electricity consumption and per capita health care expenditure). "Soft" metrics can be manipulated, and governments have an incentive to do so. Proxy metrics are either unable to be manipulated or manipulation does not directly translate to improved economic status. In this case, we assume countries lack the incentive to manipulate these proxy variables as it would have minimal impact on the overall economy, especially compared to GDP exaggeration. In this study, we compare the two proxy variables to reported GDP figures, analyzing the impact across regime types while considering other various factors.

This article contributes to several strands of the academic literature. Its most immediate contribution is to the literature on the manipulation of official statistics in authoritarian regimes. It primarily builds off the findings of Martinez (2018) who found that yearly GDP growth rates are inflated by a factor between 1.15 and 1.3 when

analyzing visible night light, a proxy variable for economic growth, on GDP. We take this idea of proxy metrics in an attempt to prove statistical manipulation by comparing GDP to per capita Health Care Expenditure.

Furthermore, this study includes a corruption metric, examining the role it plays in regime types, GDP exaggeration, and healthcare expenditure. Multiple studies have looked at the impact of corruption on levels of public health, mainly through the lens of child mortality (Hanf et al. 2011) and COVID-19 (Khan et al., 2022; McMann and Tisch, 2023). Understanding the unique role health plays in economic growth, especially considering corruption and autocracy measures, adds to multiple strands of economic research. Lastly, the connection between corruption and authoritarian regimes has been well documented (Nur-tegin and Czap, 2012; Zaloznaya, 2015; Alon et al., 2016). Its inclusion and relevance in this analysis are further documented later.

This paper tests whether autocratic governments can get away with greater data manipulation than their democratic peers. This includes an underlying assumption that democracies are built upon a system of checks and balances that limits potential data manipulation. Furthermore, democracies uphold pillars of civil liberties, including free speech, that grant individuals the ability to scrutinize published statistics, holding the government accountable and reining in potential manipulation. A combination of both the lack of liberties and increased corruption found in more autocratic regimes may allow and facilitate purposeful data tampering.

Literature Review

Manipulation

Administration and institutional competence that leads to successful resource allocation and management are referred to as good governance. The importance of good governance in modern economic progress has been emphasized. It is revealed to increase the country's competitiveness and keep the economy functioning properly. Poor administration, on the other hand, might impede the economy from realizing its full potential. Effective governance, according to Kraipornsak (2018), is a vital factor that may significantly contribute to the development of Asian rising economies.

Corruption as a component of weak governance implies that only a percentage of available resources is used for economic activity. As a result, the general population is unable to fully benefit from the economic events. Furthermore, transitioning economies are widely regarded as among the most corrupt. Transparency International (TI) creates and calculates the Corruption Perceptions Index (CPI) on an annual basis, and the 2022 (CPI) demonstrates that various countries are failing to tackle corruption. Western Europe and the European Union, the top-scoring regions, have seen scores stagnating or dropping for more than a decade. Denmark, Finland, and New Zealand are all exceptionally clean countries. Countries with low ratings continue to lag, unable to make significant progress. According to TI, restrictions and attacks on civic space and basic freedoms continue across the Americas, Eastern Europe, Central Asia, and Sub-Saharan Africa. Several crises threaten security, stability, democracy, and human rights. Similarly, rising authoritarianism weakens civil society's position as a watchdog in many Asia Pacific countries, while many governments prioritize economic recovery above anti-corruption efforts. Corruption is undermining democratic processes, causing widespread social unrest, and stoking carnage across

the Middle East and North Africa, where unequal political and economic power is inexorably linked to war.

Various authors have studied the manipulation of government statistics using different indexes and factors, while few have made use of the emphasis on political institutions. Andreis et al. (2022) assessed the contribution of Earth Observation (EO) data from various organizations' estimations, as well as the classification of Sustainable Development Goals (SDG) indicators based on the three pillars of sustainability. Traditional data can be expensive and have other drawbacks, whereas new technologies, such as satellite EO data and techniques, can play an essential role in compiling data for numerous SDG indicators.

Sandefur and Glassman (2015) investigated how governments in poor countries are misled by public personnel in charge of service delivery. They contended that serious problems exist in the statistics given by national and international agencies in some African countries, owing to anomalies between administrative data and household survey-based estimates in education and health. Previous evaluations of African statistics, according to Jerven (2011), have focused on the dynamic in which central governments are the providers of untrustworthy statistics. In other cases, national governments are deliberately misled, creating a huge impediment to domestic evidence-based decision-making. National statistics production in Sub-Saharan Africa is significantly reliant on foreign aid compared to other government responsibilities. Donors require this information to allocate aid amongst nations, evaluate specific initiatives, and analyze the overall economic management of recipient governments. Bruzelius et al. (2019) used advanced image recognition and machine learning algorithms to conduct a cost-effectiveness analysis for their technique for recognizing healthcare service buildings in a rural area of Liberia. Connecting such data to community health systems can be a powerful tool for service delivery, health policy planning, and advocacy, especially in poor and middle-income countries.

Communist regimes have a long history of manipulating information. Censorship and fabrication of census data were well documented under Stalin (King, 1997). Many authoritative regimes actively limit and undercut media inquiries, limiting their ability to investigate economic disinformation (Lucas, 2022). The most extreme extent of media suppression was seen following the USSR's population census in 1937. While the data was "exceptionally thorough and complete," Stalin had it entirely suppressed following the reveal of data that highlighted the government's inability to control the 1932-33 famine, and the officials were arrested and executed, (Merridale, 1996). King et al. (2013, 2017) found proof of restriction and falsification of social media information in communist China.

Furthermore, according to Martinez (2018), the autocracy gradient is more than twice as wide in countries with communist backgrounds. China is classified as a civil and communist dictatorship with a strong dictatorial regime. Several studies have found that China's GDP growth has been consistently inflated in national accounts (Young, 2003; Madisson, 2006). Others contend that there is no evidence of manipulation or that growth has been exaggerated (Mehrotra and Paakkonen, 2011; Nakamura et al., 2016; Clark et al., 2017). Martinez (2018) discovered substantial evidence that Chinese GDP estimates were significantly overstated, and China was one of the countries with the largest reported excess in GDP elasticity when compared to visible night light.

Electricity

Common sense as well as several studies support the idea that electricity access allows for increased economic development and progress. Multiple studies have shown that is a superior form of energy source, enabling superior information and communication technologies and more productive manufacturing and organization capabilities (Kander, Malanima, and Warde, 2014). However, nothing beyond the simple correlation between the two variables (electricity Consumption and GDP Growth) has been proven. Stern, Burke, and Burns (2019) described the relationship between Electricity and Consumption hand-in-hand but stipulated that simple access to electricity is not likely sufficient for economic growth.

However, in our case, Electricity Consumption, while not a perfect measure of the economy, provides a good enough proxy variable that a government would be highly unlikely to manipulate. Furthermore, the results are validated and summated by a third-party source, providing some additional support that these numbers are valid. By comparing the relationships between these variables and GDP across government type we can look to identify potential evidence of manipulation in the stated GDP data.

Healthcare

Researchers are becoming increasingly interested in the relationship between income growth and healthcare costs. Even as the economic structures and health expenditures of the nations included variety, there is a tendency for an increase in both per capita GDP and health expenditures for all. Several studies have been carried out to analyze the relationship between Health Care Expenditure and GDP, which are the determinants of health spending. Mushkin (1962) outlined the importance of healthcare spending in driving a country's economic growth through the Health-led Growth Theory. Investing in health (capital) will boost income and economic growth by influencing human and capital accumulation.

According to Elmi and Sadeghi (2012), there is bidirectional, long-term causation between GDP and HC, and economic growth plays a vital role in increasing healthcare spending, which serves as a long-term growth engine for developing countries. Furthermore, the joint test shows that variables have bilaterally strong causation, meaning that when a shock occurs in the system, the variables will make short-run adjustments to restore long-run equilibrium.

Bhargava et al. (2001) evaluated the influence of health indicators on economic growth rates in industrialized and developing nations from 1965 to 1990 and discovered a positive but modest link between health and economic growth. Furthermore, Bloom et al. (2001) observed comparable good outcomes. According to their results, health has a positive and statistically significant influence on economic growth. Mayer (2001) investigated if there is a Granger causality of healthcare expenditures on income for 18 Latin American nations and discovered a high correlation between income on healthcare.

Based on data from 1978 to 2005, Li Huang (2009) investigated the link between per capita real GDP growth and

physical capital, human capital, and health investment in the production function. Both health and education have a major beneficial impact on economic growth. Ak (2012) investigated the presence of a long-term causation link between health expenditures, economic growth, and life expectancy at birth in the Turkish economy. The findings revealed a long-term association between health expenditures and economic growth, but no short-term relationship between the two datasets. Bedir (2016) found that income is a major factor in explaining differences in healthcare costs between nations. As a result, it appears that rises in income levels stimulate healthcare spending in several developing market economies.

Corruption

The links between autocracy and corruption have been documented thoroughly. Autocratic countries often lack the system of checks and balances often found in democracies which limits political corruption. Nur-tegin and Czap (2012) found that the level of corruption is greater in stable dictatorships than in unstable democracies, which supports the established idea that corruption is rife in autocratic governments. This, in turn, influences reported economic growth, potentially pushing it artificially higher. Saha and Sen (2020) saw that in strongly autocratic countries, higher corruption may lead to higher growth. Furthermore, autocracies outperform anocracies as they can utilize corruption more effectively (Alon et al. 2016). However, due to the complex nature of the links between governance and corruption, current literature may be limited due to methodological restraints.

Asymmetries of beneficiaries of corruption under certain regimes may make it more difficult to fully understand the connection between corruption, governance type, and economic growth (Zaloznaya 2014). The findings in this paper shed light on a source of non-classical measurement error in reported GDP that could lead to bias in a wide array of empirical studies that make use of this data. Instead, government officials may be purposefully altering their economic statistics, bolstering their country's economy while making it seem that their policies have been successful, keeping them in office. We cannot understand the full scope of a country's economic growth until we analyze if corruption and manipulation played a role.

Data

The data utilized in this study is a compilation of different sources from various institutions such as the World Bank, Global Change Data Lab, and Freedom House. This combination provides a time series, cross-section dataset of 151 countries from 2000-2020 across multiple variables including GDP, Population, per Capita Health Care Expenditure (HCSpC), per Capita electricity consumption, Control of Corruption (CCI), Regime Type (RT) and Freedom in the World Index (FWI). Combining these various data sources allows us to fully understand the relationships between the dependent variables and included independent variables, especially the interaction between the proxy variables and stated government type.

Our GDP data is from the World Bank's World Development Indicators database. This data is a culmination of national statistic accounts and Organization for Economic Co-operation and Development (OECD) National Account files when applicable. Other data, including population and HCSpC, also come from the World Bank

database to ensure legitimacy and avoid discrepancies between data sources. Additionally, the Electricity consumption measures came from Our World in Data.

The main measure of Autocracy in this data set is FWI data from Freedom House. Using this measure gives us greater insight into the actual limitations imposed on citizens by the government. Especially considering how important media control is in allowing governments to manipulate data, this variable is useful in describing autocracy scores for our study as a measure of de jure government types (Freedom House, 2017). Greater media restrictions, and greater ability to manipulate data, would come from a higher FWI score.

Freedom House publishes FWI data annually, with the inputs being questionnaire answers from analysts and country experts. Two sub-indices, 'political rights' and 'civil liberties' are scored on a scale from 1 to 7, with lower numbers being associated with higher enjoyment of rights and liberties. The 'political rights' index is based on 10 questions regarding the Electoral Process, Political Pluralism and Participation, and the Functioning of Government. The 'civil liberties' index is based on 15 questions regarding Freedom of Expression and Belief, Associational and Organizational Rights, Rule of Law, Personal Autonomy, and Individual Rights. These scores are averaged, and a final freedom ranking is determined. A score between 1 and 2.5 is considered free, 3 and 5.5 partially free, and 6 and 7 not free.

However, democracy, and its measurement, is not a simple task. While FWI data does provide better insight into government-imposed restrictions, it does so without considering the exact government characteristics. Therefore, a supplement data point, RT from the Regimes of the World (RoW) database in conjunction with the Varieties of Democracy Institute can be used. RoW defines four types of democracies, Liberal Democracy, Electoral Democracy, Electoral Autocracy, and Closed Autocracies. These scores are created from survey questions from 3,500 different country experts every year which assess the characteristics of democracy within a given country.

We utilize two models in this analysis, one which utilizes FWI data as the main government autocracy measure, and another using RT data. These are then compared. Ideally, both measures would show similar results, those being that as governments become more autocratic (less free), the relationship between the proxy variable and GDP increases. This is done to test the robustness of the model while providing further insight into the impacts of both de facto and de jure measures of government on potential manipulation. Utilizing both allows us to cover two major measures of government types, as no one measure is all-encompassing.

Lastly, the Control of Corruption metric from the Millennium Challenge Corporation is used to understand the role corruption has on GDP manipulation. This indicator combines 22 different assessments and surveys. Evaluated factors such as the prevalence of grand corruption and petty corruption at all levels of government and the estimated cost of bribery as a share of a company's annual sales. A higher CCI score shows more control of corruption and thus less corruption within the country. Understanding this metric and its impact on GDP is important, as corruption has been shown to hinder economic growth, as seen in Cieślik and Goczek (2017) and Thach et al. (2017). By controlling for corruption, we can analyze the impact solely of government type on GDP growth through a proxy metric.

The following tables outline the variables and summary statistics utilized in the dataset. Table 1 provides further insight into the data utilized, while both Tables 2 and 3 cover summary statistics based on government measures. These are summarized across the two measures, regime type and freedom in world rankings to better understand the scope of the variables utilized in the dataset.

Table 1. Variable Descriptions

Variable	Summary
lnGDPMill	log-transformed Country GDP ($000,000's)
Pop_Mil	Country Population (000,000's)
CCI	Control of Corruption Rating
Regime Type Categorical Variables	FWI Categorical Breakdown Electoral Democracy (ED) Electoral Autocracy (EA) Closed Autocracy (CA)
Freedom in World Index Categorical Variables	FWI Categorical Breakdown Partially Free (PF) Not Free (NF)
lnHCSpC	log-transformed Country Health Care Expenditure per Capita
lnEnergyUsePerPerson	log-transformed Energy Consumption per Capita

Table 2. Freedom in World Index Summary Statistics

Summary All Countries

VarName	Obs	Mean	SD	Min	Median	Max
lnGDPMil	2957	3.80	2.113	-1.822796	3.669336	9.89993
lnEnergyUsePerPerson	2957	9.21	1.634	4.65501	9.49529	12.4783
lnHCSpC	2957	5.64	1.661	1.744045	5.546693	9.305514
CCI	2957	-0.03	1.018	-1.672809	-.3000793	2.459118
PopMil	2957	43.19	154.142	.081131	9.446836	1411.1

Summary Free

VarName	Obs	Mean	SD	Min	Median	Max
lnGDPMil	1275	4.48	2.291	-1.822796	4.615171	9.89993
lnEnergyUsePerPerson	1275	10.02	1.114	5.85997	10.2118	12.1458
lnHCSpC	1275	6.77	1.495	3.25108	6.830652	9.305514
CCI	1275	0.70	0.941	-1.071737	.6292796	2.459118
PopMil	1275	43.72	156.066	.098462	7.223938	1383.112

Summary Partially Free

VarName	Obs	Mean	SD	Min	Median	Max
lnGDPMil	977	3.03	1.746	-.4381119	2.816736	7.827477
lnEnergyUsePerPerson	977	8.37	1.561	4.98352	8.29474	12.0216
lnHCSpC	977	4.70	1.192	1.744045	4.644979	8.180525
CCI	977	-0.50	0.643	-1.613251	-.6571345	2.301146
PopMil	977	29.87	65.385	.081131	10.5697	1396.387

Summary Not Free

VarName	Obs	Mean	SD	Min	Median	Max
lnGDPMil	705	3.62	1.819	-.4229621	3.626459	9.589909
lnEnergyUsePerPerson	705	8.89	1.822	4.65501	9.16993	12.4783
lnHCSpC	705	4.91	1.236	1.847796	4.852332	7.812172
CCI	705	-0.67	0.703	-1.672809	-.8737713	1.558678
PopMil	705	60.69	221.827	.282507	11.27661	1411.1

Table 3. Regime Summary Statistics

Summary All Countries

VarName	Obs	Mean	SD	Min	Median	Max
lnGDPMil	2957	3.80	2.113	-1.822796	3.669336	9.89993
lnEnergyUsePerPerson	2957	9.21	1.634	4.65501	9.49529	12.4783
lnHCSpC	2957	5.64	1.661	1.744045	5.546693	9.305514
CCI	2957	-0.03	1.018	-1.672809	-.3000793	2.459118
PopMil	2957	43.19	154.142	.081131	9.446836	1411.1

Summary Liberal Democracy

VarName	Obs	Mean	SD	Min	Median	Max
lnGDPMil	356	3.84	2.083	-.4381119	3.906929	9.589909
lnEnergyUsePerPerson	356	9.45	1.849	5.0577	9.365	12.4783
lnHCSpC	356	5.35	1.373	1.847796	5.495773	7.812172
CCI	356	-0.24	0.725	-1.672809	-.2412948	1.558678
PopMil	356	92.76	306.799	.536758	8.427039	1411.1

Summary Electoral Democracy

VarName	Obs	Mean	SD	Min	Median	Max
lnGDPMil	924	3.19	1.665	-.3729076	3.080073	7.895714
lnEnergyUsePerPerson	924	8.38	1.659	4.65501	8.19683	12.0125
lnHCSpC	924	4.57	1.160	1.744045	4.357431	8.180525
CCI	924	-0.74	0.659	-1.672096	-.8848099	2.301146
PopMil	924	34.12	98.148	.081131	11.93358	1396.387

Summary Electoral Autocracy

VarName	Obs	Mean	SD	Min	Median	Max
lnGDPMil	936	3.13	2.021	-1.822796	2.795363	7.730731
lnEnergyUsePerPerson	936	8.85	1.273	5.61718	9.11175	11.7443
lnHCSpC	936	5.24	1.151	2.540154	5.372344	8.606693
CCI	936	-0.27	0.549	-1.581135	-.3605599	1.717748
PopMil	936	45.36	162.224	.089949	7.297136	1338.636

Summary Closed Autocracy

VarName	Obs	Mean	SD	Min	Median	Max
lnGDPMil	741	5.38	1.899	.3478369	5.508367	9.89993
lnEnergyUsePerPerson	741	10.58	0.774	7.40567	10.6484	12.1458
lnHCSpC	741	7.62	1.077	3.336399	7.947284	9.305514
CCI	741	1.27	0.743	-.7454428	1.347415	2.459118
PopMil	741	27.95	54.460	.093419	8.363404	331.5011

Empirical Strategy

To understand the relationships between the proxy variables and GDP growth, especially in consideration of government type, we utilized a fixed-effect ordinary least-squared regression model. This model will allow us to understand the relationship between either health care expenditure or energy consumption on GDP while controlling for other various factors. Utilizing a fixed-effect model will allow us to control for country-specific characteristics that would not be considered in any additional models. Furthermore, the interaction of these variables grants insight into the relationships that are useful to understand the role government type has on GDP growth. Lastly, we utilized a log-log regression due to the distribution of our dependent and primary independent variables.

To understand the impact of our proxy variable on the “soft” GDP metric given the given government type, we

must interact with the two variables. A basic model showcasing the interaction is as follows:

$$y_{i,t} = \beta_0 + \beta_1 * Proxy_{i,t} + \beta_2 * PF_{i,t} + \beta_3 * NF_{i,t} + \beta_4 * PF_{i,t} * Proxy_{i,t} + \beta_5 * NF_{i,t} * Proxy_{i,t} + \varepsilon_{i,t}$$

Equation 1. Basic Model Equation

where $y_{i,t}$ is the GDP of country i in year t and ε is the stochastic error term. This interaction allows us to understand the relationship between the proxy metric and GDP given the government type, either on a strict definition of government type or by the civil liberties individuals are granted. This will grant us the ability to determine if these governments are exaggerating the data. Since we can expect the relationship between the proxy variables and GDP growth to be constant amongst government types, as no literature has found evidence suggesting otherwise, any deviation from this could suggest manipulation. Under this paper's hypothesis, a larger coefficient, especially amongst more autocratic and stricter government types would be in line with previous literature to show evidence of manipulation.

Next, we incorporate several controls into the model. Primarily, we utilize a fixed-effect regression model to keep constant the assumed country-specific characteristics over time. Therefore, we can get a more accurate estimation of the relationship between the independent variables, especially the interactions variables, have with the dependent variable. Additionally, we add the independent variables of population and control of corruption to limit the impact on the relationship these variables have on GDP. The full explanatory model is highlighted below.

$$y_{i,t} = \beta_0 + \beta_1 * Proxy_{i,t} + \beta_2 * PF_{i,t} + \beta_3 * NF_{i,t} + \beta_4 * PF_{i,t} * Proxy_{i,t} + \beta_5 * NF_{i,t} * Proxy_{i,t} + \beta_6 * z_i + \varepsilon_{i,t}$$

Equation 2. Full Model Equation

where z represents the control variables (population and control of corruption) included in the model. Utilizing this model will allow us to understand the relationship between the proxy variable and GDP across the government types, both stated regime types, as well as the civil liberties individuals, are granted in that country.

Results

The following section highlights the results associated with each model. Four fixed-effect regressions were run, one for each combination of government type and proxy variable (Energy and FWI, Energy and Regime Type, Health Care Expenditure and FWI, and Health Care Expenditure and Regime Type). Doing so will allow us to fully understand the relationships of each variable and GDP regardless of the government measure used. This provides greater robustness as we are testing two significantly different variables that a government is highly unlikely to manipulate. Since we utilized a log-log regression model, the coefficient interpretations are as follows: a 1% increase in the independent variable (in this case the proxy variable) will result in a coefficient increase in the dependent variable (for example, a 1% increase in electricity consumption will result in a .5% increase in GDP). Utilizing the interaction variables, we can find the coefficients across government types.

Electricity Consumption

We see statistically significant results for both regime types when utilizing electricity consumption as the proxy variable that points to evidence of potential manipulation. As countries become increasingly autocratic and increasingly strict in their allowance of civil liberties, the coefficients on the interaction term get greater. The further sections provide these breakdowns based on government measures for electricity data.

De Facto Autocratic Measure

For the *de facto* government measure, that is what civil liberties the individuals have within the country, as these rights become more restricted, the relationship between electricity consumption and GDP increases. This is in line with previous literature in identifying potential data exaggeration. Figure 1 provides the regression analysis results, which show both the coefficients as well as the significance. Additionally, Table 4 highlights solely the relationship between GDP and electricity consumption based on government type. This represents the line of best fit for the regression which is calculated both with the coefficients of the interaction variable as well as the relationship between the electricity proxy variable and GDP for all countries. We can see that as countries become more restrictive, this relationship between the variables increases.

VARIABLES	lnGDPMil
lnEnergyUsePerPerson	0.672***
	(0.0238)
PF	-0.314**
	(0.157)
NF	-0.549***
	(0.177)
Energy_PF	0.0390**
	(0.0176)
Energy_NF	0.0681***
	(0.0203)
Pop_Mil	0.00585***
	(0.000390)
CCI	0.0768***
	(0.0202)
Constant	-2.659***
	(0.217)
Observations	2,957
Number of id	151
R-squared	0.402

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

Figure 1. Electricity & FWI Results

Table 4. Summary of Relationships

Elecricity Consumption	
Freedom Ranking	Coefficient
Free	0.672
Partially Free	0.711
Not Free	0.740

De Jure Autocratic Measure

For the *de jure* government measure, that is the stated government type of the country, as governments become increasingly autocratic, the relationship between electricity consumption and GDP increases. This is in line with previous literature in identifying potential data exaggeration. Figure 2 provides the regression analysis results, which show both the coefficients as well as the significance. Additionally, Table 5 highlights solely the relationship between GDP and electricity consumption based on government type.

VARIABLES	lnGDPMil
lnEnergyUsePerPerson	0.619***
	(0.0330)
ED	-0.649**
	(0.284)
EA	-0.756**
	(0.296)
CA	-1.135***
	(0.311)
lnEnergy_ED	0.0682**
	(0.0281)
lnEnergy_EA	0.0741**
	(0.0299)
lnEnergy_CA	0.112***
	(0.0324)
Pop_Mil	0.00590***
	(0.000390)
CCI	0.0665***
	(0.0203)
Constant	-2.093***
	(0.325)
Observations	2,957
Number of id	151
R-squared	0.408

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

Figure 2. Electricity & Regime Type Results

Table 5. Summary of Relationships

Elecricity Consumption	
Regime Type	Coefficient
Liberal Democracy	0.619
Electoral Democaracy	0.687
Electoral Autocracy	0.693
Closed Autocracy	0.731

Per Capita Health Care Expenditures

However, when we look at the results provided when utilizing the health care measure as a proxy, we see limited evidence supporting widespread potential data manipulation. We see either limited statistical significance (regime type) or a decreasing relationship as countries become less free. Ultimately, this may be due to our use of healthcare expenditure as the main independent variable in these relationships. Further work may be done to identify a widespread variable that could be used in further work on the subject,

De Facto Autocratic Measures

For the *de facto* government measure, that is what civil liberties the individuals have within the country, as these rights become more restricted, the relationship between electricity consumption and GDP decreases This does not support our hypothesis or previous literature in identifying potential data exaggeration. Figure 3 provides the regression analysis results, which show both the coefficients as well as the significance. Additionally, Table 6 highlights solely the relationship between GDP and electricity consumption based on government type.

VARIABLES	(1) lnGDPMil
lnHCSpC	0.722***
	(0.0166)
PF	0.141*
	(0.0793)
NF	0.245***
	(0.0886)
lnHCSpC_PF	-0.0215
	(0.0154)
lnHCSpC_NF	-0.0386**
	(0.0178)
Pop_Mil	0.00451***
	(0.000311)
CCI	0.0117
	(0.0162)
Constant	-0.498***
	(0.0928)
Observations	2,957
Number of id	151
R-squared	0.622

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

Figure 3. Health Care & FWI Results

Table 6. Summary of Relationships

Health Care Expenditure	
Freedom Ranking	Coefficient
Free	0.722
Partially Free	0.701
Not Free	0.683

De Jure Autocratic Measures

For the *de jure* government measure, that is what the stated government type of the country is, as governments become increasingly autocratic, the relationship between electricity consumption and GDP is mixed. This does not support our hypothesis or previous literature in identifying potential data exaggeration. Furthermore, we see limited statistical significance in these variables, which makes it hard to support evidence of manipulation. Figure 4 provides the regression analysis results, which show both the coefficients as well as the significance. Additionally, Table 7 highlights solely the relationship between GDP and electricity consumption based on government type.

VARIABLES	(1) lnGDPMil
lnHCSpC	0.658***
	(0.0188)
Pop_Mil	0.00464***
	(0.000311)
CCI	-0.00591
	(0.0162)
ED	-0.0754
	(0.0780)
EA	-0.130
	(0.0867)
CA	0.0283
	(0.144)
HCSpC_ED	0.0379**
	(0.0192)
HCSpC_EA	0.0558***
	(0.0206)
HCSpC_CA	0.0338
	(0.0267)
Constant	-0.271***
	(0.0832)
Observations	2,957
Number of id	151
R-squared	0.624

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

Figure 4. Health Care & Regime Type Results

Table 7. Summary of Relationships

Health Care Expenditure	
Regime Type	Coefficient
Liberal Democracy	0.658
Electoral Democaracy	0.696
Electoral Autocracy	0.714
Closed Autocracy	0.692

Conclusions

To conclude, we see mixed results across the four models analyzed. Utilizing electricity consumption provides the most support in identifying potential manipulation, as the values are both in line with our hypothesis and previous literature for both government types as well as being statistically significant. However, those results are not matched when analyzing the Health Care expenditure values. With these values, we see limited statistical significance as well as mixed results in the coefficients of each value. Further considerations of the health expenditure value, especially that of the health care system within the country, could provide further insight into the validity of the results.

There is immense future work that could be considered for this topic, specifically in two areas, validating results and identifying other potential proxy variables. Ultimately, we looked to identify several additional variables to be used as proxies, however, limited widespread availability, either with restrictions of years the data was provided (literature or education measures and financial inclusion) as well as limited countries the data was available for made it difficult to identify solid variables to be used as proxies. Finally, the decision to utilize both health care expenditures and electricity consumption came down to a "best available" verdict, not one of the best possible. Regardless, the results received from the electricity consumption that point to the potential of data manipulation call into question the historic economic growth some of the more autocratic countries have reported in the previous decades. This builds upon previous literature by providing further examples of proxy metrics that point to data manipulation in more autocratic countries, while potentially spurring the opportunity for other variables to be utilized in the future.

This paper should further act as a cautioning sign for all agents that heavily utilize economic statistics from partially or fully autocratic governments. It is becoming harder and harder to take these values at face value, especially for autocratic governments which may produce less information to begin with. This alters our perception of economies across the globe. It also provides reasoning for further improvement of economic data metrics, whether creating and utilizing proxy metrics or through the implementation of policies that push for transparency within these autocratic countries.

References

Ak, R. (2012). The relationship between health expenditure and economic growth: Turkish case. International Journal Business Management Economic Research (IJBMER), 3(1), 404-409.

Alon, I., Li, S., & Wu, J. (2016). Corruption, regime type, and economic growth. Public Finance and Management, 16(4), 332-361.

Andries, A., Morse, S., Murphy, R. J., Lynch, J., & Woolliams, E. R. (2022). Using Data from Earth Observation to Support Sustainable Development Indicators: An Analysis of the Literature and Challenges for the Future. Sustainability, 14(3), 1191. https://doi.org/10.3390/su14031191

Andrzej Cieślik, & Łukasz Goczek. (2018). Control of corruption, international investment, and economic growth – Evidence from panel data. World Development, 103, 323-335.

https://doi.org/10.1016/j.worlddev.2017.10.028

Bedir, S. (2016). Healthcare Expenditure and Economic Growth in Developing Countries. Advances in Economics and Business, 4(2), 76-86. http://www.hrpub.org DOI: 10.13189/aeb.2016.040202

Bhargava, A., Jamison, D. T., Lau, L., & Murray, C. J. L. (2001). Modeling the effects of health on economic growth. GPE Discussion Paper Series: No. 33.

Bloom, D. E., Canning, D., & Sevilla, J. (2001). The effects of health on economic growth: theory and evidence. Working Paper 8587, National Bureau of Economic Research.

Bruzelius, E., Le, M., Kenny, A., Downey, J., Danieletto, M., Baum, A., Doupe, P., Silva, B., Landrigan, P. J., & Singh, P. (2019). Satellite images and machine learning can identify remote communities to facilitate access to health services. Journal of the American Medical Informatics Association, 26(8-9), 806-812. https://doi.org/10.1093/jamia/ocz111.

Clark, H., Pinkovskiy, M., & Sala-i Martin, X. (2017). China's GDP Growth May be Understated. NBER Working Paper 23323.

Elmi, Z. M., & Sadeghi, S. (2012). Health Care Expenditures and Economic Growth in Developing Countries: Panel Co-Integration and Causality. Middle East Journal of Scientific Research, 12(1), 88-91. DOI: 10.5829/idosi.mejsr.2012.12.1.64196

Frajman Ivković, A. (2016). LIMITATIONS OF THE GDP AS A MEASURE OF PROGRESS AND WELL-BEING. Ekonomski vjesnik/Econviews - Review of Contemporary Business, Entrepreneurship and Economic Issues, 29(1), 257–272.

Freedom House. (2017). Populists and Autocrats: The Dual Threat to Global Democracy. Freedom in the World Report 2017.

Gehlbach, S., Sonin, K., & Svolik, M. W. (2016). Formal Models of Nondemocratic Politics. Annual Review of Political Science, 19(1), 565–584.

Hanf M, Van-Melle A, Fraisse F, Roger A, Carme B, et al. (2011). Corruption Kills: Estimating the Global Impact of Corruption on Children Deaths. PLOS ONE, 6(11), e26990.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. American Economic Review, 102(2), 994-1028.

Jerven, M. (2011). Users and producers of African income: Measuring the progress of African economies. African Affairs, 110(439), 169–190.

Kander, A., Malanima, P., & Warde, P. (2013). Power to the People: Energy in Europe over the Last Five Centuries (Vol. 46). Princeton University Press. https://doi.org/10.2307/j.ctt4cgb93

Kanybek Nur-tegin & Hans J. Czap. (2012). Corruption: Democracy, Autocracy, and Political Stability. Economic Analysis and Policy, 42(1), 51-66.

Khan AR, Abedin S, Rahman MM, Khan S (2022) Effects of corruption and income inequality on the reported number of COVID-19 cases and deaths: Evidence from a time series cross-sectional data analysis. PLOS Glob Public Health, 2(11), e0001157.

King, D. (1997). The Commissar Vanishes: The Falsification of Photographs and Art in Stalin's Russia. Canongate Books.

King, G., Pan, J., & Roberts, M. E. (2013). How Censorship in China Allows Government Criticism but Silences Collective Expression. American Political Science Review, 107(2), 1–18.

King, G., Pan, J., & Roberts, M. E. (2017). How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument. American Political Science Review, 111(3), 484–501.

Kraipornsak, P. (2018). Good Governance and Economic Growth: An Investigation Of Thailand And Selected Asian Countries. Eurasian Journal of Economics and Finance, 6(1), 93-106. DOI: 10.15604/ejef.2018.06.01.009

Li, H., & Huang, L. (2009). Health, education, and economic growth in China: empirical findings and implications. China Economic Review, 20, 374-387.

Lucas, E. (2022). How Autocrats Undermine Media Freedom. Journal of Democracy, 33(1), 131-146. doi:10.1353/jod.2022.0009.

Madisson, A. (2006). Do Official Statistics Exaggerate China's GDP Growth? A Reply to Carsten Holz. Review of Income and Wealth, 52(1), 121–126.

Martinez, L. R. (2018). How Much Should We Trust the Dictator's GDP Growth Estimates? Journal of Political Economy, 130(10), 2731-2769. doi:10.1086/720458

Mayer, D. (2001). The long-term impact of health on economic growth in Latin America. World Development, 29(6), 1025-1033.

McMann, Kelly M., & Tisch, Daniel J. (2023). How Lower Levels of Corruption in Democracies Prevented COVID Deaths. V-Dem Working Paper 2023:139.

Mehrotra, A., & P¨a¨akk¨onen, J. (2011). Comparing China's GDP Statistics with Coincident Indicators. Journal of Comparative Economics, 39(3), 406–411.

Merridale, C. (1996). The 1937 Census and the Limits of Stalinist Rule. The Historical Journal, 39(1), 225–240.

Mushkin, S. J. (1962). Health as an Investment. Journal of Political Economy, 70(5), 129-157.

Nakamura, E., Steinsson, J., & Liu, M. (2016). Are Chinese Growth and Inflation Too Smooth? Evidence from Engel Curves. American Economic Journal: Macroeconomics, 8(3), 113–144.

Saha, S., & Sen, K. (2021). The corruption–growth relationship: Does the political regime matter? Journal of Institutional Economics, 17(2), 243-266. doi:10.1017/S1744137420000375

Sandefur, J., & Glassman, A. (2015). The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics. Journal of Development Studies, 51(2), 116–132. DOI: 10.1080/00220388.2014.968138

Silver, M. (2007). Core Inflation: Measurement and Statistical Issues in Choosing among Alternative Measures. IMF Staff Papers, 54(1), 163–190.

Stern, D. I, Burke, P. J, & Bruns, S. B. (2019). The Impact of Electricity on Economic Development: A Macroeconomic Perspective. UC Berkeley: Center for Effective Global Action. Retrieved from https://escholarship.org/uc/item/7jb0015q

Thach, N., Duong, M., & Tran, O. (2017). Effects of Corruption on Economic Growth - Empirical Study of Asia Countries.

Young, A. (2003). Gold into Base Metals: Productivity Growth in the People's Republic of China during the Reform Period. Journal of Political Economy, 111(6), 1220–1261.

Zaloznaya, M. (2015). Does Authoritarianism Breed Corruption? Reconsidering the Relationship Between Authoritarian Governance and Corrupt Exchanges in Bureaucracies. Law & Social Inquiry, 40(2), 345–376. JSTOR. Accessed 15 Sept. 2023.

Author Information

Ryan Coutu

https://orcid.org/0009-0005-4876-0273

Stonehill College

320 Washington Street, Easton MA

United States

Contact e-mail: *ryancoutu1@outlook.com*

Oltiana Muharremi

https://orcid.org/0000-0002-3931-2656

Stonehill College

320 Washington Street, Easton MA

United States