



Fatima Zain¹, Zafar Manzoor², Dr. Azma Batool³

¹Research Scholar, Department of Economics, Forman Christian College,

E mail: 61933625@formanite.fccollege.edu.pk

²Lecturer, Department of Economics, Forman Christian College,

E mail: zafarmanzoor@fccollege.edu.pk

³Assistant Professor, Department of Economics, Forman Christian College,

E mail: azmabatool@fccollege.edu.pk, Oric ID: <https://orcid.org/0009-0005-0163-4069>

KEYWORDS	ABSTRACT
Shadow Economy; Corruption; Fixed Effects; Panel Data; Low-Income Countries; High- Income Countries, Principal Component Analysis, Driscoll- Kraay	The paper explores the relationship between corruption and the shadow economy in high-income and low-income countries. Shadow economy refers to legal economic activity that is deliberately concealed from the government to evade taxes and regulations. The study examines whether shadow economy and corruption are complements or substitutes. To overcome the limitations of previous cross-sectional studies, a balanced panel of 98 countries (2010-2020) is employed, country Fixed Effects estimation is used, and standard errors are derived from the Driscoll-Kraay method. A composite governance index is constructed using Principal Component Analysis to deal with multicollinearity. The results show that corruption and shadow economy are complements in low-income countries, meaning that enhanced corruption control lowers activity in the shadow economy in these countries. In high-income countries, this relationship is not statistically significant: in these nations, control of corruption becomes part of the general institutional quality, without having a unique and independent impact on the informal sector.
ARTICLE HISTORY Date of Submission: 04-07-2026 Date of Acceptance: 16-08-2026 Date of Publication: 21-08-2026	 2026 Journal of Social Research Development
Correspondence	Dr. Azma Batool
Email:	azmabatool@fccollege.edu.pk
DOI	https://doi.org/10.59075/jsrd.v7i6.566
Volume/Issue	7(6), 2026, 364-388

Introduction

The shadow economy and corruption are the two prevalent and intertwined phenomena in the economic world. The term “shadow economy” can be defined as all legal and illegal production of goods and services deliberately concealed from public authorities to avoid taxation, contributions to social security, and compliance with regulatory bodies (Dreher & Schneider, 2010). In both

developed and developing nations, it represents a substantial share of economic activity (Schneider & Medina, 2018). A common definition of corruption is the misuse of official authority for personal benefit. It alters the allocation of resources, further weakens institutions, and undermines the rule of law. The relationship between these two may be fundamentally shaped by a country's level of income. The level of wealth and the quality of its governing institutions plays a crucial role in determining whether corruption and the shadow economy promote one another or function as substitutive forces. The nature of their relationship remains a subject of theoretical and empirical debate.

There are two coherent prediction mechanisms that predict different outcomes for the relationship between shadow economy and corruption, and both are logically valid. The complementarity hypothesis suggests that corruption is an operational tool for informality: corrupt officials provide informality with immunity from detection in exchange for bribes, informality, in turn, creates a demand for corrupt protection, creating a positive feedback cycle between informality and corruption (Hindriks et al., 1999). The argument is formally stated by Hibbs and Piculescu (2005): Corrupt bureaucrats tolerate the informal production of goods and services, and thus formal and informal production are mutually reinforcing within the same institutional structure. The substitution hypothesis, however, suggests that the underground option is a way of dealing with corruption by providing an exit opportunity for the entrepreneur who is faced with extortionate demands from a corrupt official (Dreher et al., 2008), as the credible threat of economic exit reduces the ability of the corrupt official to extract rents (Choi & Thum, 2005). The empirical issue is therefore not which theory is abstractly correct, but which mechanism predominates under which institutional conditions, and whether this predominance is systematically linked to a country's level of economic and institutional development.

Dreher and Schneider tested these hypotheses on a cross-section of 98 countries over 1999-2002. They concisely presented their findings, concluding that the hypotheses built on perception-based corruption indices are, at best, weakly supported. The complement result only holds under a structural corruption measure and only for instrumental variable specifications that they were aware were open to weak-instrument concerns. The cross-sectional design was a limiting constraint on their inference: it allows them to observe differences in corruption between countries at a single point in time, but cannot explain unobserved, time-invariant country-specific determinants of corruption. These are characteristics that also affect informal economic activity and drive differences in corruption. Dreher and Schneider noted that they had previously estimated panel regressions in a working paper version of their research, but they were unable to find sufficient volume of data to make the country sample large enough to be reliable resulting in the panel results being dropped. The methodological lesson could have been definitive testing of the complement and substitute hypotheses calls for a panel design that features the coverage and econometric structure to make use of variation within countries over time, not variation in structure across countries at a given point in time. This study tackles that requirement head on.

Problem Statement

The empirical research on corruption and shadow economy offers some conflicting findings about how the subject has been investigated rather than the question itself. Recent work depends on cross-sectional data, which captures variation across countries at a single point in time but cannot account for how these relationships evolve within countries over time, nor can it adequately control for unobserved country-specific characteristics that simultaneously drive both corruption and informal economic activity. The consequence is that estimated effects in these studies may partly reflect structural differences between countries rather than a true relationship between the two variables. This study addresses that limitation directly. By working with a balanced panel of 98 countries over the period 2010–2020, it becomes possible to exploit within-country variation over time, hold fixed time-invariant country characteristics through Fixed Effects estimation, and draw more credible inferences about the direction and magnitude of the corruption-shadow economy relationship. Beyond the panel-versus-cross-section distinction, one further problem in existing literature remains under addressed. When studies include governance indicators alongside corruption, they rarely correct for the high multicollinearity among these variables, which inflates standard errors and makes it difficult to isolate the independent contribution of any single governance dimension. This study will apply Principal Component Analysis to combine the correlated governance indicators before estimation to ensure that the results reflect unaffected institutional effects rather than statistical noise produced by collinear regressors.

Considering these gaps, this study addresses the following core research problem: do high-income and low-income nations have different relationships between corruption and the shadow economy in a panel data setting, and if so, in what direction and with what magnitude? By employing panel data over the years 2010 to 2020, applying the Fixed Effects estimator confirmed by the Hausman test, correction for some violations of OLS through Driscoll-Kraay standard errors, and resolving multicollinearity through PCA, this study offers a more methodologically rigorous answer to this question.

Significance of the Study

Cross-sectional data is used in most of the current research on the link between corruption and the shadow economy. It captures variation across countries at a single point in time but cannot account for how these relationships change within the countries over time. Dreher and Schneider (2010) being the foundational literature in this study, themselves acknowledged this limitation directly, stating that panel regressions were discarded due to insufficient data coverage. Their cross-sectional analysis was of 98 countries over the time of 1999–2002. They did distinguish between high-income and low-income countries but produced weak and inconclusive evidence. The authors mentioned that consistent panel data on corruption based on a structural model would be needed to test the hypotheses properly, but there was unavailability of such data at the time. This study addresses that gap directly. It uses a balanced panel of 98 countries over 2010–2020, exploring within-country variation over time and holding time-invariant country characteristics constant through Fixed Effects estimation confirmed by the Hausman test. Models estimate separately for

high-income and low-income samples. This produces results that are both statistically more credible and directly comparable across the two income groups.

This study makes a methodological contribution by explicitly addressing multicollinearity among governance indicators, a problem that is pervasive in previous literature but rarely corrected for. The Worldwide Governance Indicators that includes Rule of Law, Government Effectiveness, Regulatory Quality, Voice and Accountability, and Political Stability can measure related dimensions of a common institutional construct (Kaufmann et al., 2010) and are therefore highly correlated. When included simultaneously as regressors, their Variance Inflation Factors exceed 10, which inflates standard errors and distorts the estimated contribution of the primary variable of interest (O'Brien, 2007). This study applies Principal Component Analysis, following Jolliffe (2002), to collapse the five indicators into a single composite governance index prior to estimation. The first principal component captures 83 percent of the variance in governance indicators for high-income countries and 59 percent for low-income countries and reduces all VIF values below the acceptable threshold of 10. Finally, the results of this study have special significance for developing economies like Pakistan where corruption and shadow economy are longstanding structural issues. Pakistan's shadow economy has been estimated to contribute 30-40 percent of GDP in Pakistan (Schneider, 2012) and its Control of Corruption scores have been consistently low as per Transparency International (2023). The uncovering that corruption and the shadow economy are complementary in low-income nations has direct policy implications for the sequencing of anti-corruption policy in these contexts. Reforms that weaken corruption, but do not provide an alternative to the informal sector in the formal sector, can threaten the livelihoods of informal workers who are relying on the corrupt facilitation of the informal sector, particularly in South Asian developing economies (Huynh & Nguyen, 2020).

Literature Review

Defining the Shadow Economy and Corruption

One of the most studied yet least measurable phenomena in Economics is the shadow economy. Part of this challenge starts at the definitional level because there are no universally accepted definitions. The concept is taken from a broader political economic perspective by Gyls (2005), who defines it as 'economic activity that is not a part of the official economic structures and escapes the regulatory and fiscal control of the state'. However, Schneider and Enste (2000) have provided a more standard definition for empirical purposes, namely the shadow economy can be described as the market for the production of lawful goods and services but are intentionally concealed from public authorities so as to evade tax, social security contributions, or legal requirements regarding the labour market and administrative compliance. This definition specifically excludes activities which are solely illegal, for example drug trafficking, robbery, etc. This definition is similar to the one used by Buehn (2012) and Schneider et al. (2010). Wong (2023) also discussed that the informal economy was best described as the lawful production of goods and services, purposefully hidden from authorities. This is consistent with the most used description on shadow economy.

This study refers to corruption as the intentional misuse of a public office for private benefits, as formulated by Transparency International, and often used in the literature. Shleifer and Vishny (1993) adopt a similar definition in their inspiring industrial organization study of corruption, and so does Tanzi (1998). However, Buehn and Schneider (2012) point out that the World Bank provides a more precise definition: Corruption undermines the rule of law and the institutional framework of a country, particularly affecting those who are the poorest and most disadvantaged. The empirical measure used in this study is the Control of Corruption indicator from the World Bank's Worldwide Governance Indicators (WGI) framework, intending to show the perceptions of the extent to which public power is employed for private gain, whether in major or minor ways. The problem of measurement is common for both the shadow economy and corruption: both being hidden by their very nature. The people who are involved in it actively hide their activities from the authorities and it is impossible to see them directly. Researchers have devised three general methods of estimation for the shadow economy. The first is the currency demand approach which was developed by Tanzi (1983). It uses excess demand for cash to estimate informal economic activity, because it cannot be explained by official transactions. The second is the Multiple Indicators Multiple Causes (MIMIC) approach which is widely used by Schneider and colleagues. It considers shadow economy as an unobserved variable that is caused by other observable variables (tax burden, regulatory intensity, etc.). They are made visible by indicators like circulation of money, participation in the labor market, etc. The third method is the physical input method: based on electricity consumption, which becomes a proxy for total economic activity, and subtracts official GDP to estimate the underground economy. Estimates for 162 countries from 1999 to 2006/2007 are produced by Schneider et al. (2010). These are one of the most widely cited benchmarks in cross-country literature. In the area of corruption, the most common measures are perception based indices, in particular the Transparency International Corruption Perceptions Index (CPI) and the World Bank's Control of Corruption score. These go together with structural model-based indices such as the Dreher et al. (2007) index which defines the likely causes and consequences of corruption, rather than expert perceptions of corruption. The selection of corruption measure is not unbiased: As Dreher and Schneider (2010) document, the results can vary substantially across corruption measures, such as between perception-based and model-based measures, which have direct methodological implications for this study.

Empirical Evidence from Developing and Low-Income Countries

Empirical research on developing and low-income countries is always found to support the complementarity prediction. In countries with poor institutions, research in various regions, eras, and methods reveals a correlation between informal economy and corrupt governance. Dang and Nguyen investigate how corruption and the quality of institutions affect the shadow economy in 29 Asian countries. Their findings are consistent with the positive and statistically significant effect of corruption on the shadow economy and the negative effect of institutional quality. Importantly, they also examine the role of interaction effects. They discovered that corruption and economic freedom held a negative relationship; that is, in countries with stronger economic institutions, corruption's expansion of informality is less. The pattern is a conditional one that aligns with the

larger income-heterogeneity thesis of Dreher and Schneider (2010) and offers more specific insights into the institutional pathways by which corruption spreads informal work in developing countries. Elgin and Oztunali explore the informal sector's development in 141 countries from 1984 to 2009, employing a two-sector dynamic general equilibrium model and panel data estimation. Their conclusions point to the importance held by institutional quality as a conditioning factor in the association between informality and economic development. The result is counterintuitive, but in context, makes perfect sense: in countries that lack institutional quality, an increase in GDP per capita is associated with an increase in the informal sector. In other countries that have well-developed institutions, it is the other way around.

Torgler and Schneider (2009) take a complementary path in analyzing the determinants of the shadow economy, concentrating on subjective aspects of compliance behavior, namely tax morale and perceived institutional quality. They discovered that there is robust empirical evidence that greater tax morale is associated with a significantly smaller shadow economy, as is greater perceived institutional quality, even after accounting for objective factors like tax rates, public expenditure levels, and regulatory intensity. Schneider et al. (2010) provide the most extensive cross-country estimates of the shadow economy. They are for 162 countries for the period of 1999 to 2006/2007. Other prime findings include the fact that the shadow economy in Sub-Saharan Africa is on average 37.6 per cent of GDP, compared with 36.4 per cent in transition economies (in Europe and Central Asia) and 13.4 per cent in high-income OECD countries. The distribution by income groups offers good descriptive evidence for the hypothesized link between development and informality. The study further discusses how the magnitude of the shadow economy is primarily caused by the combination of direct and indirect taxes, as well as labour market regulations and the quality of public goods and services, all of which are generally lower in low-income countries, and are closely related to corruption. Wong (2023) adds an inequality dimension to the corruption-informality relationship, looking at 127 countries between 1964 and 2007. In high-corruption, high-informality settings, the benefits of growth are more broadly shared as both bribery and informal work offer alternative avenues to access resources that bypass formal constraints. This is in line with the argument of greasing the wheels for developing economies and further suggests that the concept of corruption and informality in low-income countries cannot be separated from one another. Huynh (2020) analyzes the shadow economy in 22 Asian developing countries from 2002 to 2015 and investigates how it affects air pollution and how fiscal policy acts as a moderator. The study confirms the size and persistence of the shadow economy in Asian developing countries and the sensitivity of the shadow economy to fiscal policy instruments. An increase in government spending leads to a decrease in informality, an increase in taxes leads to an increase in informality.

Empirical Evidence from High-Income Countries

The developing-country evidence is not as clear-cut as the high-income country evidence. Overall, the research reviewed in this paper suggests that the shadow economy is smaller and less bribery-intensive in advanced economies, as predicted by Dreher and Schneider (2010) in their substitution or null-relationship hypothesis. But there are not complete complementarities, especially if

corruption is interpreted in a broad sense, such as large-scale institutional failures. Baklouti and Boujelbene (2020) investigate the impact of the interaction between shadow economy and corruption on economic growth in 34 OECD countries from 1995 to 2014. They conclude that the shadow economy exacerbates the detrimental impact of corruption on growth and that, therefore, there is a large positive relationship between the two variables even in advanced economies. This outcome is not in direct contradiction to the Schneider (2010) substitution hypothesis, which suggests that the coefficient of the shadow economy in regressions should be negative or null, not positive and significant, as the shadow economy grows. Instead, it implies that corruption and informality can work together, albeit in distinct ways, to reduce official economic activity. The take-home message is that anti-corruption and formalization policies are not redundant in OECD countries; both are needed to achieve the highest growth. This study is conducted during the period 2010-2022, which is also the time span covered by Yefimenko (2024), who has recently and thoroughly analyzed the trends in corruption and shadow economy in 34 European countries in the same period. The analysis shows some interesting differences among the European sample. Denmark, Norway, Finland, Sweden, and Switzerland are among the countries that have low corruption and low shadow economy across the time series, consistent with the substitution hypothesis, which suggests that high institutions will have an impact on reducing both measures. In contrast, countries such as Ukraine, Moldova, Bulgaria, and Albania still have a prominent level of corruption and a large shadow economy, despite a number of anti-corruption reforms.

The Role of Governance and Institutional Quality

One of the common themes across the literature reviewed above is that institutional quality is not just one of the many variables in the shadow economy regression but is a key variable in the magnitude and characteristics of the informal sector. In the cross-sectional analysis, Dreher and Schneider (2010) report that the shadow economy is always negatively related to the rule of law. Buehn and Schneider (2012) also get this result in their SEM framework. Government Effectiveness reflects the quality of public services, the autonomy of the civil service from political influences, and the trustworthiness of government promises of policy. Elgin and Oztunali conclude that the quality of institutions, which is one of its key components government effectiveness, is a fundamental determinant of the relationship between GDP per capita and informal sector size. Regulatory Quality reflects the government's capacity to make and execute well-designed policies that foster private sector growth. Excessive regulatory restrictions, arbitrary licensing requirements, and uncertainty about regulatory policies are two key factors that lead to informal operations (Johnson et al., 1997). Voice and Accountability is the extent to which citizens have the right to choose their government and press and civil society is free. Increased accountability lowers corruption and increases the responsiveness of officials to the social costs of informality, which is an indirect pathway through which democratic governance is able to decrease the size of the informal sector (Schneider & Enste, 2000). The tax burden is one of the most established factors in theoretical and empirical literature that dictates the extent to which the informal economy grows. Higher taxes can make underground work more attractive by making the gap between the total cost of labor in the official economy and the unofficial economy bigger, as Johnson et al. (1997)

show. They conclude that the overall burden of direct and indirect taxes, as measured by Schneider et al. (2010) in their sample of 162 countries, is one of the major determinants of the size of the shadow economy in the world. The negative correlation cross-country, between GDP per capita and size of the shadow economy, is also well documented by Schneider, Buehn and Montenegro (2010) and by Elgin and Oztunali (2010) in a dynamic panel setting. The variable is put in log form to allow for diminishing returns, and to minimize the effects of extreme values at the ends of the income distribution. Multiple governance indicators and the corruption variable lead to a well-documented multicollinearity issue. As Kaufmann et al. (2010) point out, the six WGI dimensions are conceptually and empirically linked: countries with high rule of law tend to have higher levels of effective government, regulatory quality and corruption. The standard errors of all the coefficients associated with governance indicators are inflated, and the individual coefficients are less precise, in a regression with several of these indicators entered as separate variables. This study addresses this issue by using Principal Component Analysis (PCA), which reduces five non-corruption governance indicators to a single indicator that reflects the dimension of institutional quality which is common to all indicators, while eliminating multicollinearity. The resulting composite governance index (Gov. Index PCA) is then used as one variable in the regression and is expected to have a negative sign in the regression equation: the greater the quality of institutions, the smaller the shadow economy.

Panel Data Approaches in the Shadow Economy Literature

Most of the research examining the relationship between the informal sector and corruption, such as the most prominent studies cited in this section, is based on cross-sectional data. Cross-sectional analysis is useful because it can encompass many countries and capture general trends by income and institutional context. It does, however, have one major drawback: it cannot control for time-invariant and country-specific factors that are also linked to both corruption and the scale of the informal economy. These differences in colonial history, legal foundations, cultural norms about taxation and compliance, and informal institutions norms are unobservable that remain constant over time across countries. This is an acknowledged limitation of the panel regressions themselves, as Dreher and Schneider (2010) explicitly state that they were not included in their analysis because of the "insufficient data coverage. The cross-sectional approach thus leaves open the possibility that the relationships estimated are only part of the result of omitted time-invariant heterogeneity. This issue can be addressed by panel data with country fixed effects, which is based on variation across time within countries. A fixed effects estimator determines the corruption-shadow economy relationship from the alteration in corruption and the discrepancy in the shadow economy in the sample of Bangladesh, while the other attributes of Bangladesh that vary over the sample period are held constant. This is a much more believable method of identification. It is necessary to choose the appropriate estimator between the fixed and random effects by an empirical method. The standard specification test for this choice was developed by Hausman (1978). The Hausman test is employed in order to check if the country-specific effects are correlated with the regressors. Random effects is inconsistent if they are correlated, in that case fixed effect might be used. For the low-income countries, the Hausman test gives a chi-squared value of 57.66 and a p

value of < 0.001 . On the other hand, for high-income countries, the Hausman test gives a chi-squared value of 17.25 and a p value of 0.028. In addition to the selection of the estimator, large cross-country panel datasets also have a few error misspecification issues that are often not considered in the literature but can significantly bias inferences. The first is heteroscedasticity, in which the error variances are not the same for each country, which is a natural phenomenon for panels that include both small and large economies. The second is serial correlation, in which errors in one year are correlated with errors in the following year in a country, which is to be expected in macro data, such as annual data, where shocks are persistent. The third is cross-sectional dependence, which occurs when shocks to one country ripple through the other countries due to trade, capital flows, contagion and common factors. The latter issue is exacerbated in a panel spanning 2010–2020, which encompassed the post-Global Financial Crisis recovery, the European debt crisis, commodity price cycles, and the early months of the COVID-19 pandemic, all factors that impacted multiple nations at the same time and in correlated fashion. Conventional clustered standard errors and ordinary least squares are not corrected for cross sectional dependence. Driscoll and Kraay (1998) designed standard errors that are robust to all three types of misspecifications, heteroscedasticity, serial correlation and cross-sectional dependence, and that are consistent for large panels. This study uses Driscoll-Kraay standard errors in both samples. In this respect this study is an improvement over the existing cross sectional literature on three fronts: first, the use of a panel data (country fixed effects) instead of a cross section; second, the use of Driscoll-Kraay standard errors instead of the conventional OLS standard errors; and third, the use of PCA to solve the multicollinearity problem that affects any governance regression. Each of these choices is motivated directly by the limitations identified in the foundational paper of Dreher and Schneider (2010) and by the empirical diagnostics carried out on the dataset used in this study. They constitute a more methodologically sound test of the income-heterogeneity hypothesis than has been found in the undergraduate economics literature before.

Research and Methodology

Theoretical Framework

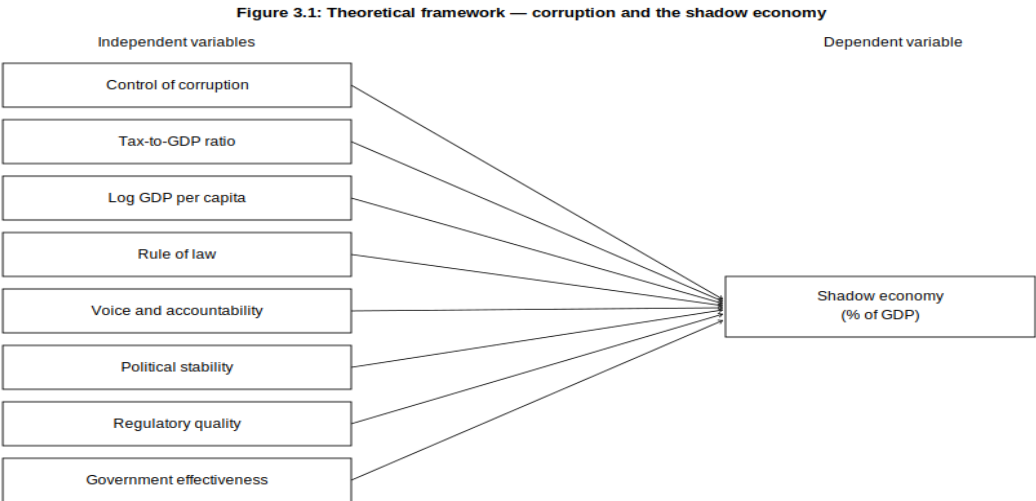
The premise of the current study is based on a framework of income heterogeneity developed by Dreher and Schneider (2010). This implies that the link between corruption and the informal economy is not the same in all countries but is systematically determined by economic development and institutional strength. Based on this approach, the relationship between corruption and informality is either positive or negative. Complementarity is assumed to be the theory for low-income countries. Insufficient protection by law enforcement and low institutional quality may result in a situation where informal producers are unable to operate without buying protection from public bodies. If bribery is the price of impunity, then corrupt officials charge rents from the firms in the shadow economy and in exchange receive the freedom of the firm to continue to operate without interference. This forms a vicious circle: a bigger shadow economy leads to a increased level of corruption, and more that corruption increases, the more the shadow economy increases. The complementary mechanism is reflected in the models of Johnson et al. (1997),

Hindriks et al. (1999) and Hibbs and Piculescu (2005). The empirical implication is that the coefficient of the Control of Corruption indicator in the regression should be negative and statistically significant in the low-income sample. The negative coefficient for the Control of Corruption suggests that a rise in corruption control is correlated with a decrease in the informal economy as predicted by complementarity.

In high-income countries, the prediction is, theoretically speaking, substitution. Entrepreneurs can rely on robust anti-corruption enforcement, independent courts and judiciaries to use effective legal tools when they encounter corrupt officials. In other words, bribery is not a usual reaction to regulatory burden because firms comply formally, litigate when required, or lobby via proper channels. Those that do operate informally typically do so for tax and compliance reasons rather than purchase official protection. In advanced economies, corruption is likely to relate to large public procurement, which is part of the formal economy. Therefore, the shadow economy and corruption are largely institutionally separated. The empirical implication is that Control of Corruption coefficient should be positive exhibiting a substitution relationship, where the shadow economy increases the more corruption increases.

The income-differentiated theoretical predictions are directly tested in the empirical model estimated in this study. The model is constructed separately for the low and high-income country samples, and the sign, size and statistical significance of the Control of Corruption coefficient is compared between the two sample groups. Figure 1 shows the theoretical framework in the form of a diagram. The Shadow economy (% of GDP) is the dependent variable and Control of Corruption is the main independent variable. There are seven other independent variables used as controls. This same model is used for high-income and low-income samples.

Figure 1: Theoretical Framework for Corruption and the Shadow Economy



Source: Based on Dreher and Schneider (2010).

Data Sources and Sample Construction

The data used in this study is generally utilized in the cross-country empirical economics literature. The size of the shadow economy is expressed as a percentage of GDP and based on the estimates of the World Bank. It uses Multiple Indicators Multiple Causes (MIMIC) methodology which assumes the scale of the shadow economy is an unobserved latent variable that is determined by observable causes and indicators. The estimates derived from MIMIC are some of the more commonly cited in the cross-country shadow economy literature. The main independent variable, Control of Corruption, is from the World Bank's Worldwide Governance Indicators (WGI) database. This indicator is indicative of how much one believes that public powers are used for private benefit, including petty and grand corruption. Higher values of the indicator indicate higher levels of corruption control. In the entire WGI data set, it ranges from about -2.5 to $+2.5$. There are also five more governance indicators used by WGI: Rule of Law, Voice and Accountability, Political Stability and Absence of Violence, Regulatory Quality, and Government Effectiveness. The indicators are correlated with one another and with the Control of Corruption variable, requiring the PCA treatment described in Section 3.7. World Bank World Development Indicators (WDI) data is used for both tax revenue as a percentage of GDP and real GDP per capita (in constant US dollars). All variables are available at annual frequency and have full coverage for the sample window 2010–2020. The final balanced panel was built by a carefully designed filtering process. Beginning with the full World Bank dataset, the data was first restricted to the years 2010–2023. Next, each country was rated on the following nine variables: the shadow economy, Control of Corruption, Tax-to-GDP, real GDP per capita, Rule of Law, Voice and Accountability, Political Stability and Absence of Violence, Regulatory Quality, and Government Effectiveness, to see if they were complete. Only countries with 11 complete observations for these nine variables were included. From these, only countries with complete data for the last ten years (2010 to 2020) were selected. The restriction results in an 11-year observations per country panel, avoiding compositional bias that can occur if countries come in and out of the sample in a systematic way. Income classification is based on the classification of the World Bank. The final balanced panel of 98 countries, observed each year between 2010 and 2020, gives 1,078 country-year observations. The sample of low-income nations consists of 24 countries and 264 observations, while the sample of high-income nations consists of 74 countries and 814 observations. In this research, these two samples are analyzed separately for all regression analysis. Table 1 gives a brief description of all the variables included in the model. It also provides information for data sources and literature sources.

Table 1: Defining the Variables

Variables	Type of Variable	Definition of the Variable	Data Sources	Literature Sources
1. Shadow Economy (% of GDP)	Dependent	Measured as a percentage of official GDP using the MIMIC method. Captures all market-based legal production deliberately concealed	World Bank (Medina & Schneider, 2018)	Schneider et al. (2010); Medina & Schneider (2018); Dreher & Schneider

		from public authorities to avoid taxation and regulatory compliance.			(2010)
2. Control of Corruption	Independent (Primary)	World Bank WGI indicator reflecting perceptions of the extent to which public power is used for private gain. Scaled so higher values indicate better corruption control. Range: -2.5 to +2.5.	World Bank WGI		Kaufmann et al. (2010); Dreher & Schneider (2010); Buehn & Schneider (2012)
3. Tax-to-GDP Ratio	Independent (Control)	Tax revenue as a percentage of GDP. Proxy for fiscal burden. Higher tax burden increases incentives to operate informally.	World Bank WDI		Johnson et al. (1997); Schneider et al. (2010); Dreher & Schneider (2010)
4. Log GDP per Capita	Independent (Control)	Real GDP per capita in constant USD, entered in log form. Controls for level of economic development. Richer countries have smaller shadow economies.	World Bank WDI		Schneider et al. (2010); Elgin & Oztunali (2010); Dreher & Schneider (2010)
5. Governance Index (PCA)	Independent (Control)	Composite index from PCA of five WGI indicators: Rule of Law, Voice and Accountability, Political Stability, Regulatory Quality, and Government Effectiveness. First principal component captures 83.1% (high-income) and 59.1% (low-income) of joint variance.	World Bank WGI		Kaufmann et al. (2010); Jolliffe (2002); O'Brien (2007)

Source: Construction based on World Bank WGI and WDI databases.

Model Specification

This study estimates the baseline model by regressing the shadow economy on the Control of Corruption indicator, in addition to a set of control variables, separately for the low-income and high-income country sample. It applies the Fixed Effects estimator. The model has the following equation:

$$SE_{it} = \alpha_{it} + \beta_1 CoC_{it} + \beta_2 LogGDP_{it} + \beta_3 TaxGDP_{it} + \beta_4 GovIdx_{it} + \varepsilon_{it} \quad (Eq1)$$

Where,

SE_{it} : shadow economy in country i in year t , as a percentage of official GDP

α_{it} : Country fixed effect. It captures all time-invariant country-specific factors such as colonial history, legal origin, and geographic characteristics

CoC_{it} : World Bank WGI indicator of Control of Corruption.

LogGDP_{it}: natural log of real GDP per capita in constant dollars from the United States.

TaxGDP_{it}: Tax revenue as a percent of GDP. Use proxy for the cost to firms and households.

GovIdx_{it}: Five components of the World Governance Index (WGI) combined into a composite governance index.

$\beta_1, \beta_2, \beta_3, \beta_4$: slope parameters to be determined for each of the independent variables.

ε_{it} : unique error term,

Subscript *i* refers to the country and subscript *t* to the year. The model is estimated separately for the low-income and high-income sample, following the stratified estimation approach. The relation between the shadow economy and corruption separately for each of the income groups. The coefficient of primary interest is β_1 , the coefficient of the Control of Corruption indicator. The complementarity hypothesis (H1) predicts that, in the low-income group, the sign of β_1 should be negative, that is, the shadow economy should be smaller with the better control of corruption. If the substitution hypothesis (H2) is accepted, then we would expect the parameter β_1 to be positive confirming the substitutive effect.

Estimation Strategy

The most common estimators for a panel dataset are Pooled Ordinary Least Squares (OLS), Random Effects (RE), and Fixed Effects (FE). Pooled OLS does not take the panel structure into account. It is inconsistent if unobserved country-specific effects are correlated with the regressors, as is often the situation in macro panels where countries have systematic differences in their unobserved institutional histories (Wooldridge, 2010). Random Effects is efficient if the country effect is not correlated with the regressors, but inconsistent if it is correlated. Fixed Effects eliminates all country-specific heterogeneity and is consistent in both the presence and absence of correlation between country effects and the regressors (Hsiao.C, 2014). The full 98-country panel was subjected to three specification tests to choose the proper estimator. First, an F-test between the Fixed Effects model and Pooled OLS is significant, with $F = 1,055$ on (98, 982) degrees of freedom, and $p\text{-value} < 2.2 \times 10^{-16}$, providing strong rejection of the pooled model and supporting the existence of significant country-specific effects. Second, the Breusch-Pagan Lagrange Multiplier test shows that there are random effects: $\chi^2 = 4,887.8$ with one degree of freedom and $p\text{-value} < 2.2 \times 10^{-16}$. Third, the Hausman test for a difference in the two estimators, Fixed Effects and Random Effects, gives a χ^2 of 29.90 and 8 degrees of freedom, with a $p\text{-value}$ of 0.0002, which suggests the rejection of the null hypothesis of no systematic difference between the two estimators and the conclusion that the country-specific effects are correlated with the regressors.

An additional Hausman test was performed for each income sample based on the unabbreviated 8-variable model before PCA. For the low-income sample, the test results in a $\chi^2 = 57.66$ with 8 degrees of freedom and $p\text{-value} = 1.34 \times 10^{-9}$, which is highly significant and rejects Random

Effects. Again, for the high-income sample, the Chi-squared value is 17.25, with eight degrees of freedom, and the p-value is 0.028, which is under the five percent significance level, so we reject Random Effects. Therefore, Fixed Effects is confirmed to be the correct estimator in both samples. The Fixed Effect is used to estimate all regressions in this study.

Diagnostic Tests and Standard Error Correction

Fixed Effects coefficient estimates are unbiased, but valid inference requires the error terms to have certain properties (Wooldridge, 2019). The three typical violations of macro panels are heteroscedasticity, serial correlation, and cross-sectional. All three were statistically tested.

Breusch-Pagan Test

The Breusch-Pagan test confirmed heteroscedasticity for both low-income and high-income samples. The results are provided in the appendix.

Breusch-Godfrey-Wooldridge Test

The results of the Breusch-Godfrey-Wooldridge test confirmed the serial correlation in both the samples which was expected given the constant nature of shadow economy. The table containing results is shown in the appendix.

Breusch-Pagan LM Test

Cross-sectional dependence was tested using the Breusch-Pagan LM test on Fixed Effects residuals. For both the samples the presence of cross-sectional dependence was confirmed. The results are shown in the appendix. The Driscoll-Kraay (1998) standard errors can be applied which accounts for all three types of misspecifications: heteroscedasticity, serial correlation up to a specified lag, and cross-sectional dependence. The same diagnostic tests were performed on the final model, and the three violations were checked, so the standard errors of Driscoll-Kraay are needed for the final model specification.

Data Analysis

However, prior to econometrics, there are theoretically important patterns that can be gleaned from a distribution of variables. There are 2010–2020 country-year observations for 814 country years of 74 high-income countries and 264 country years of 24 low-income countries. Data for the descriptive analysis is presented by income group. The data shown in descriptive analysis is sorted according to income.

Descriptive Statistics by Income Group

Table 2 presents the mean, standard deviation, median, minimum and maximum values, the standard error, skewness and kurtosis of each variable for both samples of income. Table 2 shows descriptive data for each variable measured in this study, by income group. The mean of the low-income sample is 37.3 percent of GDP. The low-income sample has less dispersion ($sd = 7.6$) and a distribution that is skewed towards the center between 17.6 and 55.1 percent, indicating that informality is not a country-specific phenomenon but a structural characteristic of the institutional weakness (Dreher & Schneider, 2010). The high-income sample, however, exhibits significantly greater variability ($sd = 12.2$) ranging from 8.1 to 67.0 percent, indicating a more diverse mix of country-specific factors that influence informality among the high-income countries. Nearly zero kurtosis is found in both distributions, allowing the use of the shadow economy in levels, not logs, in regression. The distribution of shadow economy for each income group is shown in Figure 2. The time is set from 2010 to 2020. The comparison gives a better understanding in terms of its spread over the years.

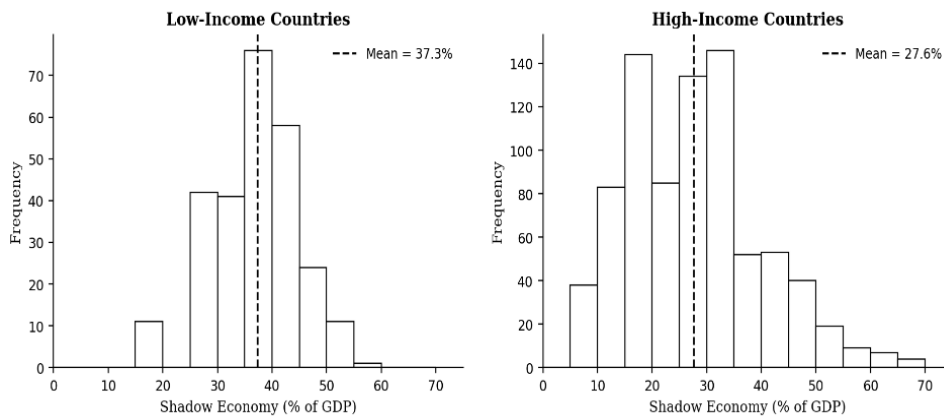
Table 2: Descriptive Statistics by Income Group (2010–2020)

Variables	Mean	Sd	Median	Min	Max	Se	Skewness	Kurtosis
Panel A: High-Income Countries (74 countries, 814 observations)								
Shadow Economy (%GDP)	27.64	12.17	27.30	8.10	67.00	0.43	0.75	0.19
Control of Corruption	0.50	1.01	0.40	-1.65	2.40	0.04	0.13	-0.88
Tax-to-GDP Ratio (%)	17.83	6.07	17.62	0.79	37.61	0.21	0.09	-0.49
Log GDP per Capita	9.61	0.96	9.51	7.65	11.59	0.03	0.08	-0.97
Rule of Law	0.53	0.93	0.49	-1.42	2.12	0.03	0.06	-0.93
Voice & Accountability	0.48	0.92	0.69	-2.00	1.74	0.03	-0.35	-0.72
Political Stability	0.30	0.74	0.42	-2.02	1.62	0.03	-0.59	0.52
Regulatory Quality	0.68	0.80	0.65	-1.71	2.25	0.03	-0.06	-0.69
Government Effectiveness	0.61	0.85	0.54	-1.36	2.28	0.03	0.13	-0.81
Panel B: Low-Income Countries (24 countries, 264 observations)								
Shadow Economy (%GDP)	37.33	7.56	38.00	17.60	55.10	0.46	-0.27	0.00

Control of Corruption	of	-0.52	0.55	-0.61	-1.46	1.62	0.03	0.72	1.48
Tax-to-GDP Ratio (%)		14.59	6.33	13.02	6.83	38.08	0.39	1.56	2.93
Log GDP per Capita		7.36	0.76	7.37	6.05	9.10	0.05	0.06	-0.63
Rule of Law		-0.49	0.44	-0.57	-1.28	0.63	0.03	0.28	0.19
Voice & Accountability		-0.42	0.51	-0.38	-1.82	0.60	0.03	-0.33	0.32
Political Stability		-0.42	0.70	-0.36	-2.21	1.11	0.04	-0.38	0.76
Regulatory Quality		-0.50	0.37	-0.52	-1.45	0.31	0.02	0.12	0.27
Government Effectiveness		-0.57	0.46	-0.69	-1.38	0.69	0.03	0.55	0.09

Note: Calculations based on World Bank WGI and WDI databases. Sd represents Standard Deviation and Se represents Standard Error

Figure 2: Distribution of Shadow Economy (% of GDP) by Income Group (2010–2020)



Source: Own construction based on World Bank data Dashed line denotes the group mean.

Selection of Estimation Method

The panel of all 98 countries was subjected to three successive specification tests to determine the best estimator. The results are presented in Table 3. The F test for individual effects is highly

significant ($F = 1055$, $p < 0.001$), indicating that there are country-specific effects, and that pooled OLS would lead to highly biased estimates. The Breusch–Pagan Lagrange Multiplier test corroborates the existence of panel effects ($\chi^2 = 4887.8$, $p < 0.001$). The Hausman (1978) specification test fails to accept the null hypothesis of uncorrelated individual effects with the regressors ($\chi^2 = 29.90$, $p = 0.0002$), which shows that Random Effects is inconsistent. Therefore, Fixed Effects is the appropriate and required estimator, and all subsequent regressions use the Fixed Effects within estimator. Three tests were carried out on the full panel data. The results are shown in table 3. It confirms that Fixed Effect is the most suitable decision.

Table 3: Model Selection Tests on the Full Panel

Dependent variable: Shadow Economy (% of GDP)

Test	Statistic	p-value / Decision
F-test: Fixed Effects vs. Pooled OLS	$F = 1055.0$	< 0.001 suggests Fixed Effects strongly preferred
Breusch-Pagan LM: Random Effects vs. Pooled OLS	$\chi^2 = 4887.8$	< 0.001 suggests Random Effects exist; pooling invalid
Hausman Test: Fixed vs. Random Effects	$\chi^2 = 29.90$	0.0002 suggests Fixed Effects confirmed

Residual Diagnostics

A battery of diagnostic tests was run on the Fixed Effects residuals for both income samples. The tests for heteroscedasticity (Breusch–Pagan), serial correlation (Breusch–Godfrey–Wooldridge), and cross sectional dependence (Breusch–Pagan LM) all fail to accept their respective null hypotheses at the one percent level of significance in both samples (full results are presented in the Appendix). The pattern is typical of macroeconomic panels across countries of differing institutional development, in which the informality and governance dimensions of the panel evolve over time at the country level, but vary across countries and are subject to global shocks. Cross-sectional dependence requires clustered standard errors. Throughout, the Driscoll–Kraay (1998) estimator is used, which corrects simultaneously for heteroscedasticity, serial correlation and cross sectional dependence. All regression tables in this section are corrected. The five WGI governance indicators were highly collinear prior to PCA. In the high-income sample, these were reduced to a single composite governance index with a maximum VIF of 9.40 and in the low-income sample, a maximum VIF of 3.83. The coefficient of the VIF of Control of Corruption was reduced from 10.75 to 8.99, thus allowing confidence in the interpretation of the coefficient. The results of the post-PCA VIF diagnostics are summarized in Table 4. It shows the value of VIF for both the income groups. All of them lie below threshold.

Table 4: Variance Inflation Factor Diagnostics for Preferred PCA Model

Variable	VIF for Low-Income	VIF for High-Income	Status
Control of Corruption	3.32	8.99	Below threshold
Tax-to-GDP Ratio	1.42	1.87	Below threshold
Log GDP per Capita	2.18	3.10	Below threshold
Governance Index (PCA)	3.83	9.40	Below threshold

Note: A Threshold VIF > 10 is considered problematic multicollinearity (O'Brien, 2007). All values are below threshold in both samples following PCA.

Regression Results

Table 5 displays the regression coefficients and t-values for the low-income countries. The preferred model used PCA approach to get unbiased results.

Table 5: PCA Model for Low-Income Countries

Dependent variable: Shadow Economy (% of GDP)

Variable	Coefficient	t-value
Control of Corruption	-0.7136	-3.080**
Tax-to-GDP Ratio	-0.0261	-1.021
Log GDP per Capita	-7.0624	-15.895***
Governance Index (PCA)	-0.1566	-2.254**

Note: (***) $p < 0.01$, (**) $p < 0.05$. Fixed Effects, Driscoll-Kraay standard errors. PCA captures 59.1% of governance variance. Max VIF = 3.83. N = 264, 24 countries, 2010–2020.

Table 6 displays the regression coefficients and t-values for the high-income countries. Using Driscoll-Kraay standard errors and Fixed Effects gave accurate results.

Table 6: PCA Model for High-Income Countries

Dependent variable: Shadow Economy (% of GDP)

Variable	Coefficient	t-value
Control of Corruption	0.0853	1.278
Tax-to-GDP Ratio	-0.0464	-4.926***
Log GDP per Capita	-6.0210	-23.487***
Governance Index (PCA)	-0.2789	-6.143***

Note: (***) $p < 0.01$, (**) $p < 0.05$. Fixed Effects, Driscoll-Kraay standard errors. PCA captures 83.1% of governance variance. Max VIF = 9.40. N = 814, 74 countries, 2010–2020.

The robust standard errors corresponding to both high-income and low-income countries are reported separately in Table 7. Driscoll-Kraay standard errors were specifically chosen and were used in both preferred models for high-income and low-income countries.

Table 7: Driscoll-Kraay Robust Standard Errors for Preferred PCA Model

Variable	Driscoll-Kraay SE (Low-Income)	Driscoll-Kraay SE (High-Income)
Control of Corruption	0.2316	0.0668
Tax-to-GDP Ratio	0.0256	0.0094
Log GDP per Capita	0.4443	0.2564
Governance Index (PCA)	0.0694	0.0454

Note: Standard errors produced by the Driscoll-Kraay (1998).

Discussion of Results

Low-Income Countries

The most consistent and statistically significant result from this study is the negative and significant coefficient for Control of Corruption in all the low-income specifications. The preferred PCA model indicates that a 1-point increase in the WGI Control of Corruption score is correlated with a 0.714 percentage-point decrease in the shadow economy share of GDP ($\beta = -0.714$, Driscoll-Kraay SE = 0.232, $p = 0.002$). The result is robustly stable in sign and significance for all specifications estimated. In institutionally weak settings, informal enterprises pay bribes to tax inspectors, labour regulators, and enforcement authorities to get the immunity that formal legal protection otherwise affords. Corruption and informality are co-produced equilibria, reinforcing each other. The more anti-corruption moves are taken and the higher the Control of Corruption score is, the weaker the network of protection based on bribery, and the smaller the shadow economy becomes (Schneider, 2010). This is reflected in the within-country variation between 2010 and 2020, which signals the weakening of the corrupt facilitation network. The log GDP per capita coefficient is large and highly significant ($\beta = -7.062$, $p < 0.001$) and thus confirms that income-driven formalization is the most important structural cause of informality, even in low-income contexts. The composite Governance Index is negative and significant ($\beta = -0.157$, $p = 0.025$), suggesting that the quality of governance does have an independent effect on informality, beyond the control of corruption. As in low-income countries, where tax evasion is widespread, the coefficient of the Tax-to-GDP ratio is negative, but statistically not significant ($\beta = -0.026$, $p > 0.10$), and it is in line with the evidence that in such countries there is no meaningful link between tax rates and enforcement capacity (Dreher & Schneider, 2010).

High-Income Countries

The higher income results are more complex. The Control of Corruption coefficient is positive in direction ($\beta = +0.085$), but statistically insignificant ($p = 0.202$) in the preferred PCA model. This is in line with the concept of institutional absorption, which suggests that corruption control is not a distinct institutional dimension in a high-income economy with functional courts, contracts, and regulatory institutions (Kaufmann et al., 2010). The composite Governance Index is included in the model and explains the variation formerly explained by the Control of Corruption coefficient, making it insignificant. The implication for practice is that it is not the anti-corruption policy per se that determines the dynamics of informality in high-income countries, but rather it is overall institutional quality. Both the log GDP per capita coefficient ($\beta = -6.021$, $p < 0.001$) and the composite Governance Index ($\beta = -0.279$, $p < 0.001$) are negative and highly significant. The coefficient of the Tax-to-GDP ratio is negative and significant ($\beta = -0.046$, $p < 0.001$), which is in line with the theoretical prediction that in a country with a good tax administration, a higher fiscal burden should make informal operation less attractive (Johnson et al., 1997). All these findings point to the fact that in high-income countries fiscal compliance and general institutional quality are the factors driving informality, while corruption control is integrated into the wider dimension of institutional quality.

Control Variables

The size of the shadow economy is best and most consistently predicted by log GDP per capita for both income groups and all specifications. The model coefficients of the PCA (-7.062 for low-income, $p < 0.001$; -6.021 for high-income, $p < 0.001$) suggest that development-driven formalisation is the main structural driver of informality. The income-informality relationship in the cross-country literature is well documented, and a 10 percent rise in real per capita income is associated with about a 0.7 percentage-point drop in the share of the shadow economy, as in the other literature. The income-informality relationship in the cross-country literature is well established, and a 10 percent increase in real per capita income is associated with approximately a 0.7 percentage-point fall in the share of the shadow economy, as in the other literature (Schneider et al., 2010). The composite Governance Index has a negative and significant impact on informality for both samples (low-income: $\beta = -0.157$, $p = 0.025$; high-income: $\beta = -0.279$, $p < 0.001$), thus indicating that the overall institutional quality has a negative effect on informality beyond the impact of corruption control and income. The higher absolute value of the governance range in the high-income income sample implies that the economic impact of the range of governance in high-income countries is more significant than in the low-income country sample, which is more institutionally homogeneous. The Tax-to-GDP ratio is statistically insignificant in the low-income sample and negative and significant in the high-income sample, in line with the theory that fiscal burden is only a constraint on informality decisions when tax administration is strong enough to make it a meaningful constraint (Dreher & Schneider, 2010).

Conclusion

This study examined whether the relationship between corruption and the shadow economy is complementary or substitutive in high-income and low-income countries. Fixed Effects models were estimated separately for 24 low-income and 74 high-income countries, covering the period 2010–2020, using a balanced panel of countries. Driscoll–Kraay standard errors were used to correct for heteroscedasticity, serial correlation, and cross-sectional dependence. The World Bank Worldwide Governance Indicators were aggregated into a composite governance index through Principal Component Analysis to resolve multicollinearity. The results validate the income-conditional nature of the link between corruption and shadow economy. The Control of Corruption coefficient is negative and statistically significant for all the specifications in low-income countries. In the preferred PCA model, there is a 0.714 percentage-point decrease in the shadow economy as a percentage of GDP. This finding of Johnson et al. (1997) and Schneider (2010) supports the complementary hypothesis that shadow economy and corruption work together. In institutionally weak settings, corrupt officials take rents from informal producers in return for impunity, thus fostering a vicious cycle where more corruption leads to a larger shadow economy. In high income countries, the Control of Corruption coefficient has a positive sign, but is not statistically significant ($\beta = +0.085$, $t = 1.907$, $p = 0.202$), meaning that the size of the shadow economy is not driven by corruption control alone, once institutional quality is considered. The most important factors of informality in this sample are the composite Governance Index and log GDP per capita, which confirms the independence hypothesis and the idea that corruption is "institutionalized" in the larger governance picture, rather than a unique factor.

Limitations

Shadow economy estimates are not directly observed but rather model based (MIMIC method), so the coefficients are not necessarily capturing the true level of unobserved informal activity instead variations in the informality estimates (Medina and Schneider, 2018). In addition to this, the balanced panel requirement to ensure consistency of the estimation process prevented the inclusion of some countries in South Asia and the Middle East, and this reduced the generalisability of findings to these regions. Finally, the analysis is designed by purpose for the period 2010–2020, since the economic changes caused by the pandemic would generate structural breaks that are not considered in this work.

Future Research Suggestions

The next step in this research would be to repeat the analysis for a post-COVID panel that is 2020 onwards. Shadow economic activity was contracted in some countries but increased in others during the pandemic, and the fiscal and institutional reactions differed markedly by income group. This suggests that the income conditional corruption-informality relationship found here may not be robust in the face of acute shocks to the external sector in the post-2020 period.

Policy Recommendations

The findings have practical policy implications on fighting corruption and formalization. In low-income countries, the fight against corruption needs to go together with improving the easy entrance in the formal sector otherwise the corrupt protection network could be taken away without an institutional replacement, thus depriving informal producers of access to markets and regulatory tolerance. The evidence from high-income countries suggests that the drivers of formalization are primarily more general institutional reforms and development, rather than specific anti-corruption measures. Corruption and the shadow economy are not necessarily connected in a single universal way by income level. They are bound together by the bribery arrangements that replace the legal protection offered by weak institutions in low-income countries. In high-income countries, that same relation breaks down and is encompassed in a more institutional manner.

Policymakers must make the formal sector easy to access, such as when the corrupt arrangements are demolished, all parties from the informal sector have a legitimate institutional alternative to shift to. There is no justification for the presence of corruption; however, the best way forward for both low and high-income countries is producing reforms like tax amnesty programmes to pair with anti-corruption policies. Indonesia's 2016 tax amnesty is a good example to follow. Initiatives like business registration simplification and targeted social protection can help cater to this economic problem. Corruption and shadow economy might not always go hand in hand, but they can be influenced by the institutional environment, thus policies can be built to tackle this issue.

References

1. Baklouti, N., & Boujelbene, Y. (2020). Shadow economy, corruption, and economic growth: An empirical analysis. *Review of Black Political Economy*, 47(3), 276–294. <https://doi.org/10.1177/0034644619885349>
2. Buehn, A., & Schneider, F. (2012). Corruption and the shadow economy: Like oil and vinegar, like water and fire? *International Tax and Public Finance*, 19(1), 172–194. <https://doi.org/10.1007/s10797-011-9175-y>
3. Choi, J. P., & Thum, M. (2005). Corruption and the shadow economy. *International Economic Review*, 46(3), 817–836. <https://doi.org/10.1111/j.1468-2354.2005.00347.x>
4. Dang, V. C., Nguyen, Q. K., & Tran, X. H. (2023). Corruption, institutional quality and shadow economy in Asian countries. *Applied Economics Letters*, 30(21), 3039–3044. <https://doi.org/10.1080/13504851.2022.2118959>
5. Dreher, A., & Schneider, F. (2010). Corruption and the shadow economy: An empirical analysis. *Public Choice*, 144(1–2), 215–238. <https://doi.org/10.1007/s11127-009-9513-0>
6. Dreher, A., Kotsoyiannis, C., & McCorriston, S. (2007). Corruption around the world: Evidence from a structural model. *Journal of Comparative Economics*, 35(3), 443–466. <https://doi.org/10.1016/j.jce.2007.05.001>

7. Dreher, A., Kotsogiannis, C., & McCorriston, S. (2009). How do institutions affect corruption and the shadow economy? *International Tax and Public Finance*, 16(6), 773–796. <https://doi.org/10.1007/s10797-008-9089-5>
8. Driscoll, J. C., & Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, 80(4), 549–560. <https://doi.org/10.1162/0034653985557825>
9. Elgin, C., & Oztunali, O. (2012). Shadow economies around the world: Model based estimates (Bogazici University Department of Economics Working Papers No. 2012/05). Bogazici University.
10. Gyls, P. (2005). Economy, anti-economy, underground economy: Conceptual and terminological problems. *Ekonomika*, 72, 1–11.
11. Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251–1271. <https://doi.org/10.2307/1913827>
12. Hibbs, D. A., Jr., & Piculescu, V. (2005). Institutions, corruption and tax evasion in the unofficial economy (Discussion paper). Göteborg University, Department of Economics.
13. Hindriks, J., Keen, M., & Muthoo, A. (1999). Corruption, extortion and evasion. *Journal of Public Economics*, 74(3), 395–430. [https://doi.org/10.1016/S0047-2727\(99\)00030-4](https://doi.org/10.1016/S0047-2727(99)00030-4)
14. Hsiao, C. (2014). *Analysis of panel data* (3rd ed.). Cambridge University Press.
15. Huynh, C. M. (2020). Shadow economy and air pollution in developing Asia: What is the role of fiscal policy? *Environmental Economics and Policy Studies*, 22(3), 357–381. <https://doi.org/10.1007/s10018-019-00260-8>
16. Huynh, C. M., & Nguyen, T. L. (2020). Fiscal policy and shadow economy in Asian developing countries: Does corruption matter? *Empirical Economics*, 59(4), 1745–1761. <https://doi.org/10.1007/s00181-019-01700-w>
17. Johnson, S., Kaufmann, D., & Shleifer, A. (1997). The unofficial economy in transition. *Brookings Papers on Economic Activity*, 28(2), 159–240. <https://doi.org/10.2307/2534688>
18. Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). Springer-Verlag.
19. Kaufmann, D., Kraay, A., & Mastruzzi, M. (2010). The worldwide governance indicators: Methodology and analytical issues (World Bank Policy Research Working Paper No. 5430). World Bank. <https://doi.org/10.2139/ssrn.1682130>
20. Medina, L., & Schneider, F. (2018). Shadow economies around the world: What did we learn over the last 20 years? (IMF Working Paper No. 18/17). International Monetary Fund. <https://doi.org/10.5089/9781484338636.001>
21. O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673–690. <https://doi.org/10.1007/s11135-006-9018-6>
22. Schneider, F. (2012). The shadow economy and work in the shadow: What do we (not) know? (IZA Discussion Paper No. 6423). Institute for the Study of Labor.
23. Schneider, F., Buehn, A., & Montenegro, C. E. (2010). Shadow economies all over the world: New estimates for 162 countries from 1999 to 2007 (World Bank Policy Research Working Paper No. 5356). World Bank.

24. Schneider, F., & Enste, D. H. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, 38(1), 77–114. <https://doi.org/10.1257/jel.38.1.77>
25. Shleifer, A., & Vishny, R. W. (1993). Corruption. *Quarterly Journal of Economics*, 108(3), 599–617. <https://doi.org/10.2307/2118402>
26. Tanzi, V. (1983). The underground economy in the United States: Annual estimates, 1930–80. *IMF Staff Papers*, 30(2), 283–305.
27. Tanzi, V. (1998). Corruption around the world: Causes, consequences, scope, and cures. *IMF Staff Papers*, 45(4), 559–594.
28. Torgler, B., & Schneider, F. (2009). The impact of tax morale and institutional quality on the shadow economy. *Journal of Economic Psychology*, 30(2), 228–245. <https://doi.org/10.1016/j.joep.2008.11.007>
29. Transparency International. (2023). Corruption perceptions index 2023. <https://www.transparency.org/en/cpi/2023>
30. Wong, M. Y. H. (2023). Economic development, corruption, and income inequality: The role of the informal sector. *Politics*, 45(2), 163–184. <https://doi.org/10.1177/02633957221148951>
31. Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.
32. Wooldridge, J. M. (2019). *Introductory econometrics: A modern approach* (7th ed.). Cengage Learning.
33. Yefimenko, A. Yu., & Dubovenko, M. M. (2024). Trends in corruption and shadow economy in European countries. *Ekonomika. Finansy. Pravo*, 2024(5), 112–117. <https://doi.org/10.37634/efp.2024.5.21>

Appendix A

Diagnostic Tests and Standard Error Correction:

Breusch-Pagan Test for Heteroscedasticity

Dependent variable: Shadow Economy (% of GDP) for Pooled OLS Residuals

Sample	BP Statistic	df	p-value	Decision
Low-Income Countries	29.52	8	0.0003	Homoscedasticity rejected
High-Income Countries	125.55	8	$< 2.2 \times 10^{-16}$	Homoscedasticity strongly rejected

Breusch-Godfrey-Wooldridge Test for Serial Correlation

Dependent variable: Shadow Economy (% of GDP) for Fixed Effects Residuals

Sample	χ^2 Statistic	df	p-value	Decision
Low-Income	98.91	11	2.94×10^{-16}	Serial correlation confirmed

Countries				
High-Income	254.9	11	$< 2.2 \times 10^{-16}$	Serial correlation confirmed
Countries				

Breusch-Pagan LM Test for Cross-Sectional Dependence

Dependent variable: Shadow Economy (% of GDP) for Fixed Effects Residuals

Sample	χ^2 Statistic	df	p-value	Decision
Low-Income	848.37	276	$< 2.2 \times 10^{-16}$	Cross-sectional dependence confirmed
Countries				
High-Income	8,806.6	2,701	$< 2.2 \times 10^{-16}$	Cross-sectional dependence confirmed
Countries				