Energy Economics 140 (2024) 107970

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneeco

Regulatory impact of informality on gasoline consumption efficiency in Africa: A proposed two-part complementary hypothesis test

Philip Kofi Adom

School of Economics and Finance, University of The Witwatersrand, Johannesburg, South Africa

ARTICLE INFO

Keywords:
Gasoline efficiency
Regulation
Informal economy
Africa
Two-part complementary hypothesis

ABSTRACT

This study provides new evidence for the regulatory impact of informality on gasoline efficiency in Africa. I propose a Two-Part Complementary Hypothesis (hereafter referred to as the TPCH test), advocating a differential approach to promoting gasoline efficiency: (1) an inverted U-shaped relationship between informality and gasoline inefficiency, and (2) a U-shaped relationship between government regulation and informality, with a significant level effect. The findings indicate an inverted U-shaped effect of informality on gasoline efficiency and a level-negative effect of regulation on informality. These results suggest a differential strategy for enhancing gasoline efficiency. Government regulation is more effective in economies at the pre-saturation stage (characterized by normal growth levels of informality) but proves ineffective in economies at the post-saturation stage (characterized by abnormal growth levels of informality), where energy-saving behaviors may be self-motivated. This is corroborated by the inefficiency equation, where indicators of good governance, such as the rule of law, control of corruption, and regulatory quality, are statistically significant in advancing energy efficiency goals. Gasoline efficiency performance varies across countries, with the higher performers also being the continent's most economically advanced. However, these estimates risk downward bias if outliers or unobserved/observed heterogeneity are not considered.

1. Introduction

Studies examining the relationship between environmental quality and the size of the informal economy have often argued that strengthening regulations can effectively reduce the size of informal sector, thereby improving environmental outcomes (Swain et al., 2020; Loayza et al., 2006). This perspective presents a straightforward hypothesis: a linear, inverse relationship between the size of the informal economy and the level of government regulation, implying that continuous regulatory improvements will consistently enhance environmental quality.

In this study, I challenge this view by proposing a more nuanced, two-part hypothesis: (1) an inverted U-shaped relationship between informality and gasoline inefficiency, and (2) a U-shaped relationship between government regulation and informality, with a significant level effect. This dual hypothesis suggests that government regulation is effective in advancing environmental goals only during the pre-saturation stage (when the informal economy is growing at a normal rate) and becomes less effective in the post-saturation stage (when informal sector growth becomes abnormal). Therefore, the TPCH test advocates for a differentiated approach to promoting environmental goals.

Promoting energy efficiency initiatives can alleviate the tension between economic growth and commitment to sustainable development (Fowler and Meeks, 2020). Energy efficiency is a key strategy for transitioning to a green economy (Regina et al., 2016). The OECD/IEA joint report "Green Growth Studies: Energy" (OECD/IEA, 2011) concluded that promoting energy efficiency policies, advancing low-carbon energy technologies, and eliminating fossil fuel subsidies could reduce global carbon emissions by half by 2050. Additionally, promoting energy efficiency policies is crucial for achieving energy security (Akorli and Adom, 2023).

Achieving the Sustainable Development Goal (SDG) target 7.3 can also facilitate progress toward SDG 11 (making cities and human settlements inclusive, safe, resilient, and sustainable) (Agradi et al., 2022) and SDG 13 (climate change mitigation). As noted by Fuso-Nerini et al. (2018), SDG 7 is interconnected with all other SDGs. According to the authors, decisions related to SDG 7 influence human welfare, the development of physical and social infrastructure, and the attainment of environmental sustainability.

Despite the significance of energy efficiency for sustainable

E-mail address: philip.adom@wits.ac.za.

https://doi.org/10.1016/j.eneco.2024.107970
Received 4 April 2024; Received in revised form 29 September 2024; Accepted 11 October 2024
Available online 24 October 2024

development, improvements in this area have generally been slow and uncertain. Globally, energy intensity improved by 2.3 % between 2011 and 2016 but declined to 1.8 % in 2017 and further to 1.1 % in 2018. Although there was a resurgence in 2019 to 2 %, energy intensity decreased to 1 % in 2020, then rose again to 1.9 % in 2021 (International Energy Agency [IEA], 2020/2021). Even though energy intensity improved by 2 % in 2019, energy savings due to technical improvements declined by 5 % (IEA, 2020b). These statistics indicate that consistently reducing energy intensity levels (i.e., enhancing energy efficiency) globally has proved difficult. Achieving sustainable improvements in energy efficiency in the future can be particularly challenging, especially in developing economies (Adom, 2024). A recent study by Adom (2024), using global data, found that while energy efficiency efforts are more sustainable in the short, medium, and long term in developed, high-income OECD countries, this is not the case for lower-middle- and upper-middle-income countries. Akorli and Adom (2023) predicted that African countries have a 21 %–24 % chance of transitioning from a low energy-efficient state to a high energy-efficient state in the medium to long term. Adom (2016) and Adom and Adams (2018) highlighted the low energy efficiency trap in Cameroon and Nigeria, respectively, explaining that conditions such as underdeveloped technology markets, inefficient energy pricing, and a lack of proper regulation and financial incentives make it difficult for Africa to escape this trap. While Europe has been the most active in announcing financial support for energy efficiency programs, amounting to about USD 57 billion, there have been no such announcements in Africa (IEA, 2020a). Hernandez (2020) provided evidence of the prevalence of inefficient products in Africa. Approximately 35 % of air conditioners in Africa recorded an efficiency ratio of less than 3 *W*/W (Hernandez, 2020), raising legitimate concerns about governance and regulatory frameworks in the region.

Addressing energy consumption inefficiency is essential, but it presents a complex challenge. Achieving this goal requires integrating technology, market-based solutions, and formalizing informal activities or strengthening regulation, among other approaches. This study aims to investigate the role of informal markets—defined as economic activities (in law or practice) not officially or insufficiently accounted for in the formal sector—in gasoline consumption efficiency in Africa. Fossil fuel consumption significantly contributes to atmospheric greenhouse gas emissions (IEA, 2019; Bölük and Mert, 2014). In Africa, fossil fuel dependency is increasing (Adom and Adams, 2018), and the pollution from these fuels poses severe health risks, leading to higher mortality rates (Marais et al., 2019). Improving the consumption efficiency of these fuels can help reduce air pollution and enhance environmental sustainability (Lanzi et al., 2011; Rashid et al., 2019).

The current literature is replete with findings on the roles of technology and market mechanisms in influencing the consumption efficiency of specific energy products and aggregate energy indicators (Jollands and Ellis, 2009; Jollands et al., 2011; inter alia). However, there is limited literature on the role of the informal sector in stimulating energy consumption efficiency. Existing studies have primarily focused on aggregate energy indicators such as energy consumption and energy intensity (Karanfil, 2008; Abid, 2015; Benkraiem et al., 2019; Hosier, 1994; and Basbay et al., 2016). Karanfil (2008), Abid (2015), and Benkraiem et al. (2019) considered the effect of informality on energy consumption, with Benkraiem et al. (2019) investigating the relationship while accounting for asymmetry. Hosier (1994) used microdata to examine the informal economy and energy efficiency in Tunisia, whereas Basbay et al. (2016) considered the macro effects of the informal economy on energy intensity (a measure of energy efficiency) and energy consumption using global data. This leaves an important gap in the literature concerning the role of informal markets in the consumption efficiency of specific energy products like gasoline fuel. Moreover, existing studies have not examined the mechanisms underlying the relationship between the informal economy and gasoline energy efficiency. Understanding the channels through which informal activities affect gasoline efficiency has important policy implications, as it could provide valuable insights into the design of targeted energy policies aimed at improving energy efficiency outcomes, a crucial driver of environmental sustainability. In this study, I aim to address two important questions: (1) Does the size of the informal market affect gasoline fuel consumption efficiency? (2) Are regulation and good governance channels for the gasoline consumption efficiency and informality nexus? To this end, I propose a two-part complementary hypothesis to identify the specific conditions under which regulation effectively reduces informality and promotes environmental goals.

The study of informal markets is interesting and important for the following reasons. First, the informal sector, defined as economic activity not under the purview of formal regulation, has grown significantly in developing economies, especially in Africa (Medina et al., 2017). The ratio of the informal sector to gross domestic product in Africa averaged 38.4 % between 1995 and 2015. Mauritius and Zimbabwe recorded the lowest and highest average values of 21.9 % and 60.98 %, respectively (Data from Medina and Schneider, 2018). The United Nations predicted that approximately 2 billion people would live in the informal sector by 2022 (Figueroa, 2016). Blackman (2000) cites population growth, rural–urban migration, and the lack of regulation as factors responsible for the growth of the informal sector. Second, the informal sector contributes significantly to poverty reduction (Dawa and Kinyanjuim, 2012) and employment creation (International Labour Organization [ILO], 2012a, 2012b). In sub-Saharan Africa, the informal sector contributes 55 % of the national GDP and 80 % of the labor force (African Development Bank, 2013). Additionally, the informal sector provides an economic cushion and adjustments during crises (Colombo et al., 2013). During the Latin American economic crisis of the 1980s, the Eurozone crisis, the financial crisis of 2008, and the implementation of structural adjustment policies, informal economies increased significantly in rich, poor, and rapidly developing countries. A similar situation occurred in Asia in the 1990s when millions of formal workers who lost their positions in the formal sector sought refuge in the informal sector (Lee, 1998). Between 2007 and 2009, the informal sector provided a safety net for about 27 million people who became redundant in the formal sector (ILO, 2012c).

However, due to the complex nature of activities in the informal sector, it is either very difficult or virtually impossible to regulate these activities formally. This poses a problem, as it becomes challenging to influence behaviors in the informal sector toward promoting goals such as environmental quality and resource use efficiency. Benson et al. (2014) note that the unorganized nature of the informal sector often makes it costly and complex to extend national regulations to track and enforce environmental standards. Consequently, the implied hypothesis is that economies with large informal sectors, indicative of weak regulation and governance, may face significant challenges in promoting energy efficiency goals. However, the relationship could be more complex than postulated. Environmental challenges impact informal sector activities, so participants in informal markets might naturally promote pro-environmental outcomes. Based on this argument, I hypothesize that there is a saturation point in the nexus between environmental/energy outcomes and the size of informal activities. This also motivates another important hypothesis tested in this study, which relates to the mechanisms that might explain gasoline energy use behavior in the informal sector before reaching saturation.

In this study, I contribute to the literature by exploring the regulatory and good governance channels using seven indicators: rule of law, regulatory quality, control of corruption, political stability and absence of violence, voice and accountability, and government effectiveness. A well-established relationship exists between the size of a nation's informal economy and the quality of regulation and governance. Countries with stricter regulation and good governance tend to have a larger formalized sector, whereas those with weak regulation tend to have a larger informal sector (Loayza et al., 2006). The informal sector itself might also operate under a set of non-state norms and conventions that influence behavior to yield accountability (Swain et al., 2020).

However, this study focuses on formal rules and their impact on gasoline energy efficiency. The implied hypothesis here is that strengthening regulation and improving governance (i.e., limiting the size of the informal economy) are associated with higher gasoline consumption efficiency. Since formalization helps promote and institute standards, improving the regulatory environment can facilitate investment in gasoline energy innovation.

In this regard, this study bears a resemblance to the study by Duffield and Hankla (2011), which focused on the role of political institutions in oil energy intensity. Their study examined the direct effects of different governance systems. Contrary to Duffield and Hankla (2011), I consider regulation and governance to be a pass-through conduit for the effect of informal activities on gasoline energy efficiency, thus focusing on the indirect effects of regulation and good governance. In a market economy, the government might have little direct control over gasoline consumption levels (Duffield and Hankla, 2011). However, indirectly, the government can influence gasoline use through incentive designs such as enacting energy efficiency regulations for motor vehicles, banning over-aged cars, and promoting the importation of fuel-efficient cars. The relevance of regulation and good governance in the design of energy policy has been emphasized. In 2009, the IEA recognized regulation and good governance as critical to delivering energy efficiency outcomes (IEA, 2010). Moreover, experiences have shown that energy efficiency policies are more likely to be successful if an effective system of good governance and regulation exists (Jollands et al., 2011).

Methodologically, this study makes an important contribution to the existing literature. Previous macro-level studies on the effects of informality on energy efficiency often rely on simplistic indicators such as energy intensity (Basbay et al., 2016), which lack a solid economic foundation and have been criticized as poor measures of energy efficiency (Filippini and Hunt, 2011). While some studies employ parametric stochastic and non-parametric techniques to estimate energy efficiency (Akorli and Adom, 2023; Adom et al., 2023; Edziah and Opoku, 2024 inter alia), they did not consider the dual aspects of informality and the mechanisms through which it impacts gasoline energy consumption efficiency.

In this study, I employ the parametric endogenous stochastic frontier analysis (ESFA) technique developed by Karakaplan and Kutlu (2017b). The ESFA method addresses the dual endogeneity problem (endogeneity in both the frontier and inefficiency equations) simultaneously, a major challenge in the standard SFA literature on energy. In typical energy–economy analyses, the income coefficient is susceptible to bias. This issue is exacerbated in African countries, where national income data are generally of poor quality and likely underreported (Adom and Amoani, 2021). Such measurement problems can bias the income coefficient estimates in the frontier equation within the SFA framework (Adom et al., 2023). Moreover, the interdependency between the economy and the environment can also introduce bias in the income coefficient estimate. Addressing endogeneity concerns is crucial to accurately identifying the frontier equation, estimating fuel consumption efficiency, and identifying the inefficiency equation (Adom et al., 2023). To mitigate these identification concerns, I used the share of arable land and the country's legal system origin (common law versus civil law) as instrumental variables for real gross domestic product per capita. Numerous studies link economic performance to land (Roe, 2006; Barbier, 2004; Alesina et al., 2005; Ding and Lichtenberg, 2010; Deninger, 2003; Parkinson et al., 2014) and legal systems (Damaska, 1986; Cross, 2002; La Porta et al., 2008; Xu, 2011; Adom et al., 2023; Adom, 2021a).

Additionally, discrepancies in gasoline consumption data collection across countries can lead to unreasonably low recorded values, potentially skewing efficiency estimates downward. To address this issue, the data were winsorized at various percentiles. Also, unobserved/observed heterogeneity, which could affect efficiency estimates, was also considered. I applied an SFA model that simultaneously addresses endogeneity and unobserved/observed heterogeneities (Kutlu et al., 2019). Finally, I conducted several robustness checks to address issues related to the non-temporal variation of instruments, omitted variable bias, and measurement errors in the size of the informal economy.

The rest of the study is organized as follows: section 2 discusses the theoretical channels linking informality to energy. Section 3 reviews the empirical literature. Section 4 discusses the method and the data. Section 5 presents and discusses the findings. Section 6 concludes with policy recommendations.

2. The energy efficiency and informality nexus: Channels

The link between energy efficiency and informal economy size is complex and not definite. In this section, I explain the six possible transmission mechanisms identified.

2.1. Income effect channel

Regulating the informal sector is difficult. Consequently, participants in the informal sector suffer from widespread low-quality and low-paid jobs. Informal employers, due to lack of regulation, are not compelled to pay minimum wages, severance payments, or insurance premiums for their employees, unlike formal employers (Basbay et al., 2016). As a result, income levels are quite low in the informal sector, driving up the level of income poverty within it. Macroeconomic evidence suggests a strong association between poverty and informality; countries with lower levels of informality see a decline in the number of working poor, and vice versa (AfDB, 2013). The OECD (2009) also notes that growth in informal employment significantly increases the level of poverty, making it challenging for countries to achieve sustainable development goals.

High poverty induces risk aversion and discounts future benefits (Haushofer and Fehr, 2014), making it very challenging for the poor to purchase energy-efficient and low-carbon energy technologies, which have uncertain future costs and benefits. Figueroa (2016) revealed that informal settlements have been very slow in adopting energy-efficient technologies. For example, there is a widespread presence of old and inefficient vehicles in the urban transport sector. Moreover, these vehicles are poorly managed and do not meet environmental quality standards (Biswas et al., 2012). Recent evidence also suggests the dumping of inefficient air conditioners in Africa (Hernandez, 2020).

However, it is important to note that informality might not always be synonymous with poverty. For many people, choice, not just necessity, drives the decision to stay in the informal sector (Benson et al., 2014). A study by the Swedish International Development Agency (Sida) revealed that informal market operators assess the relative benefits of formalization based on cost-benefit analyses and choose to remain informal (Becker, 2004). A small business project survey of informal entrepreneurs in South Africa found similar results (USAID, 2005:4). Powerful groups or cartels exist in the informal wholesale, retail, and transport sectors, controlling large amounts of capital. For example, attempts to formalize informal sectors have been highly contested in Peru. Consequently, the informal sector can be a game changer in terms of investing in energy-efficient and low-carbon technologies. Thus, based on the income effect alone, increasing the size of the informal sector can reduce or increase gasoline energy efficiency depending on which effect is dominant.

2.2. Input/technology mix channel

This channel reflects the composition of labor and capital in the productive activities of the informal sector. Labor is the dominant input used in the informal sector (Basbay et al., 2016; Ihrig and Moe, 2004; Celestin, 1989). The high labor intensity, combined with low capital intensity, reduces the energy requirements of the sector compared to the formal sector, resulting in lower energy consumption in the informal sector. It is important to note that this reduction in energy consumption is due to the input mix composition, which may not be related to efficiency improvements. Nonetheless, the input mix composition may correlate with energy efficiency behaviors in the informal sector. Energy efficiency is highly correlated with capital intensity (Hosier, 1994). Consequently, given the low capital intensity nature of the informal sector, growth in informal economy activities can lead to gasoline energy consumption inefficiency.

2.3. Regulatory effect channel

This perspective asserts that the highly unorganized and complex nature of the informal sector makes it difficult to implement and enforce regulations that control emissions and improve resource use efficiency (Benson et al., 2014; Elgin and Oztunali, 2014). Consequently, pollution and resource use inefficiency may be very high in informal economies. The implication is that in economies with large informal markets, strengthening regulation and governance might improve resource use efficiency and reduce pollution. Regulation ensures standards, creating the institutional framework required to promote innovation (Porter, 1991; Gann et al., 1998; De Vries and Verhagen, 2016), investment in energy R&D (Bruns et al., 2019; Kirk and Besco, 2021), and infrastructure and capacity (Cebula and Mixon, 2014).

2.4. Behavioural effect channel

This operates in two ways. First, many operators in the informal sector avoid paying the full cost of energy or pay a flat rate. For grid electricity, operators in the informal economy either consume electricity illegally or use informal connections (i.e., illegally established, unmetered connections to the electrical grid) due to the lack of bank accounts, proof of address, cumbersome procedures, and to avoid payment (Figueroa, 2016). This removes the precautionary motive—saving on energy bills—that facilitates the implementation of energy efficiency (Figueroa, 2016). Consequently, the growth of the informal economy strains an already overburdened electrical system, as there is limited incentive to rely on energy efficiency (Basbay et al., 2016). It is important to note that this first effect may result from regulatory lapses.

Second, climate change and resource scarcity threaten the activities of the informal economy, making the informal sector not just contributors to but also victims of climate change. Therefore, informal operators may become more alert to opportunities that promote emission reduction and energy efficiency improvements. Tandon (2012) revealed that the informal sector has become more responsive to resource scarcity and innovative solutions than their formal counterparts. Significant green activities are occurring in the informal sector, largely driven by international organizations, community organizations, non-governmental organizations, and researchers. These activities include the use of renewable energy and other fuel-saving innovations.

2.5. Product durability channel

In most developing economies, poor infrastructure adversely affects the lifespan of end-use equipment (Figueroa, 2016). In the transport sector, the poor condition of road infrastructure causes less durable motor vehicles to wear down quickly. As a result, investing in durable equipment is considered the best and safest option. Therefore, given that fuel-efficient vehicles tend to be durable, they become the preferred choice for informal market players. In other words, informal market players might prioritize product durability as the most critical factor when investing in technology.

2.6. Structural effect channel

The informal economy is predominantly composed of the manufacturing and service sectors, which are less energy-intensive compared to the industrialized firms of the formal sector (Chaudhuri and Mukhopadhyay, 2006). Therefore, the growth of the informal sector might indicate structural shifts in product composition from more energy-intensive sectors to less energy-intensive ones.

3. Empirical literature: Informality and energy nexus

There are very few empirical studies that investigate the nexus between energy and the informal economy. This review is organized into two categories: energy consumption and the informal economy, and energy efficiency and the informal economy. For the former, Karanfil (2008) is the first to consider the role of the informal economy in energy consumption in Turkey using a time series technique. The author's analytical procedure involved examining the link between officially recorded GDP and energy consumption, and then controlling for unrecorded economic activities. The results show that when not accounting for underground economic activities, there is a long-run relationship between official GDP and energy consumption, as well as unidirectional causality from official GDP to energy consumption. However, these results disappear once the author accounts for the size of the unrecorded economic activities. According to the author, this suggests the possibility of implementing energy conservation policies without adverse effects on the economy.

Adopting a somewhat similar analytical procedure, Abid (2015) reached a different conclusion in the short run for Tunisia. In the long run, the author finds evidence of causality between energy consumption and both formal GDP and total GDP (formal and informal). The short-term evidence of neutrality between formal GDP and energy consumption vanishes once the author considers the informal sector. The short-term evidence shows causality from total GDP to energy consumption, implying that energy conservation policies might harm the economy.

Benkraiem et al. (2019) extended the argument, suggesting that the relationship might not be symmetric, as proposed by Abid (2015) and Karanfil (2008). They applied the asymmetric autoregressive distributed lag technique to analyze the effect of official GDP, true GDP, and the underground economy on energy consumption in Bolivia. For the underground economy, the results show a positive and symmetric effect, indicating that increases in underground activities are associated with higher energy consumption in the long run.

Using the energy consumption indicator, it is challenging to determine whether higher informal activity is associated with increased or decreased energy inefficiency based on the above studies. This complexity arises because countervailing factors may be at play that do not have a direct link to technical efficiency improvements. Additionally, energy efficiency and energy consumption are not synonymous. Energy efficiency pertains to process innovation, while energy consumption reflects production outcomes, which can be influenced by the level of process innovation.

Studies examining the direct effect of informal activity on energy efficiency are limited. Hosier (1994) explored energy use behavior in the informal sector from a micro perspective, focusing on energy efficiency and employment potential in Tanzania. The study revealed that energy use in the informal sector is constrained by capital limitations. Capital-intensive subsectors tend to use modern fuels, such as electricity, and therefore require less energy. In contrast, less capital-intensive subsectors, which rely on traditional fuels, exhibit much higher energy consumption. Specifically, food preparation subsectors were identified as the least efficient, while welders and carpenters were found to be the most efficient.

In a similar micro-level approach, Women in Informal Employment Globalizing and Organizing (WIEGO, n.d.) conducted a study on urban informal workers (street traders and home-based workers) and the green economy in India. The results indicated that street traders use little or no electricity compared to their formal sector counterparts. For home-based workers, the study found that they use less space and utilities, are more cautious with electricity consumption, and require less transportation compared to those in the formal sector.

The only macro-level study addressing energy efficiency and informality is that by Basbay et al. (2016). They employed a panel regression approach across 159 countries to investigate the relationship between energy efficiency and the informal economy, using energy intensity as an indicator of energy efficiency. Their results indicated that countries with larger informal sectors had lower energy intensity. This finding was robust across oil-importing, emerging, OECD, and G20 countries. Additionally, the authors identified a non-linear and symmetric relationship between energy intensity and the size of the informal sector.

However, there are significant concerns with using energy intensity as a measure of energy efficiency at the macro level. Energy intensity can be influenced by several factors unrelated to technological improvements, such as population density, economic activities, climate conditions, energy prices, and production structures (Filippini and Hunt, 2011). Empirical studies, including those by Filippini and Hunt (2011), Filippini and Zhang (2016), and Zhang and Adom (2018), have demonstrated that energy intensity is only weakly correlated with energy efficiency.

In addition to potential measurement flaws, I found no evidence in the literature linking the size of the informal sector to gasoline energy efficiency, nor examining the role of regulatory and governance channels. Given the significant contribution of the informal sector to national GDP in Africa, it is plausible that informal sector players could have a considerable impact on gasoline consumption levels. This study tests the hypothesis that there is a point of saturation beyond which the informal sector becomes too large for government control (see Panel A and B of Fig. 1 demonstrating the TPCH test).[1] Before reaching this saturation point, the size of the informal sector is assumed to grow at a level within a controllable range for the government. At this pre-saturation stage, energy-saving behavior (ESB) reflects either weak or absent government regulation.

However, once the sector surpasses the saturation point, its growth exceeds the government's capacity for control. At this stage, the environmental consequences of the informal sector's activities become excessively high, potentially motivating informal market participants to engage in pro-environmental actions. Panel B illustrates that as long as countries have not reached the crisis stage, the government can address environmental issues through regulation. However, beyond the crisis stage, extending regulation to improve environmental outcomes becomes economically unfeasible, as enforcement, compliance, and monitoring become increasingly difficult. In this study, I test the TPCH. The presence of the TPCH requires an inverted U-shaped relationship between the environmental indicator and the size of the informal economy, followed by a statistically insignificant U-shaped relationship between informality and regulation, with only the level effect showing a statistically significant negative impact.

4. Method and data

4.1. Theoretical and empirical model

According to Malghan (2019), estimating efficiency involves defining a normative benchmark for a phenomenon, comparing it with the actual observed state, and measuring the deviation from this benchmark. In this study, I modeled gasoline fuel consumption efficiency using Kopp's (1981) non-radial measure of efficiency, which focuses on input-specific efficiency. Since this study is concerned with input-specific efficiency, Kopp's approach is appropriate. Given that energy is a derived demand, maximizing consumer satisfaction involves choosing the level of energy and capital that optimizes energy services (Filippini and Hunt, 2015).

[1] Panel A demonstrates relationship between gasoline efficiency (or measure of environmental outcome) and size of informality while Panel B shows the relationship between Size of informality and level of government regulation.

Fig. 2 illustrates the utility-maximizing condition for the energy user. It demonstrates that for any given level of energy service produced, the energy user must select a corresponding level of energy and capital that maximizes that service. The isoquant, denoted as IQ, represents the technically efficient level, referred to here as the normative benchmark. Point Q1 exemplifies a technically efficient allocation, where the chosen levels of energy (E) and capital (K) maximize energy services. Achieving a point like Q1 is quite rare.

Now, assume that energy services are produced at point Q2. For the same level of capital, more energy input (i.e., 'E2') is required. Since this point lies outside the production frontier, it is technically inefficient. To enhance technical efficiency of the energy input while holding output, technology, and capital constant, the energy input should be reduced from 'E2' to 'E1'. This is analogous to improving building insulation and replacing outdated equipment and appliances (Filippini and Hunt, 2011 & 2015).

Let Q1 represent the normative benchmark for energy use and Q2 denote the actual observed energy use. The segment Q1Q2 measures the deviation of the actual observed state from the normative benchmark. Technical efficiency in energy input, according to Kopp's (1981) approach, is calculated as the ratio of the distance from KQ1 to KQ2.

Based on Kopp's (1981) non-radial approach, Filippini and Hunt (2011, 2015) proposed a stochastic energy demand frontier to empirically estimate energy consumption efficiency. Their method involves estimating a conditional stochastic energy input demand frontier. Eq. 1 illustrates this frontier and includes the following components:

$f(p_{it}, y_{it}, X_{it} : \beta)e^{v_{it}}$ represents the optimal or benchmark energy use.

$e^{-u_{it}}$ denotes the deviation from the normative benchmark.

gc_{it} is the gasoline energy consumption by country 'i' at time t.

pg_{it} is the price of gasoline energy in country 'i' at time t.

$rgdp_{it}$ is the real income per capita of country 'i' at time t.

X_{it} represents other factors affecting gasoline energy consumption.

β is the vector of parameters for the frontier equation.

u_{it} and v_{it} capture the inefficiency and noise terms, respectively.

A log transformation of Eq. 1 yields Eq. 2.

$$gc_{it} = f\left(pg_{,it}, rgdp_{it}, X_{it} : \beta\right)e^{v_{it}} * e^{-u_{it}} \tag{1}$$

$$lngc_{it} = ln\, f(pg_{it}, rgdp_{it}, X_{it} : \beta) + v_{it} - u_{it} \tag{2}$$

A key consideration is the inclusion of independent variables in the frontier eq. I observe considerable variability in empirical specifications for gasoline demand across the literature, driven by contextual relevance, data availability, and the specific focus of each study. This flexibility is justified by the fact that multiple factors—economic, structural, demographic, environmental, and political—are associated with gasoline demand at the macro or national level. Gasoline plays a crucial role in the production, distribution, consumption, and transportation of goods and services across national, firm, and household levels, indicating that both micro and macro factors can influence gasoline consumption. For both micro- and macro-level studies, gasoline price and end-user income have been identified as critical factors influencing gasoline consumption (Dahl, 2012; Burke and Nishitateno, 2011; Basso and Oum, 2007; Tirkaso and Gren, 2020; Filippini and Heimsch, 2016; Adom et al., 2016; Steven, 2020, among others). Higher gasoline prices can reduce per capita gasoline consumption in the road sector or prompt substitution toward more fuel-efficient vehicles, reflecting negative price elasticity (Dahl, 2012; Burke and Nishitateno, 2011).

The relationship between income and gasoline consumption is more complex, with the literature reporting both positive and negative effects. In Dahl's systematic review, she noted that approximately 50 % of the studies reported income elasticity values between 0.25 and 0.99, while about 10 % reported negative income elasticity. Higher income can lead to increased gasoline consumption intensity; however, it can also result in lower gasoline intensity as consumers opt for more fuel-efficient vehicles.

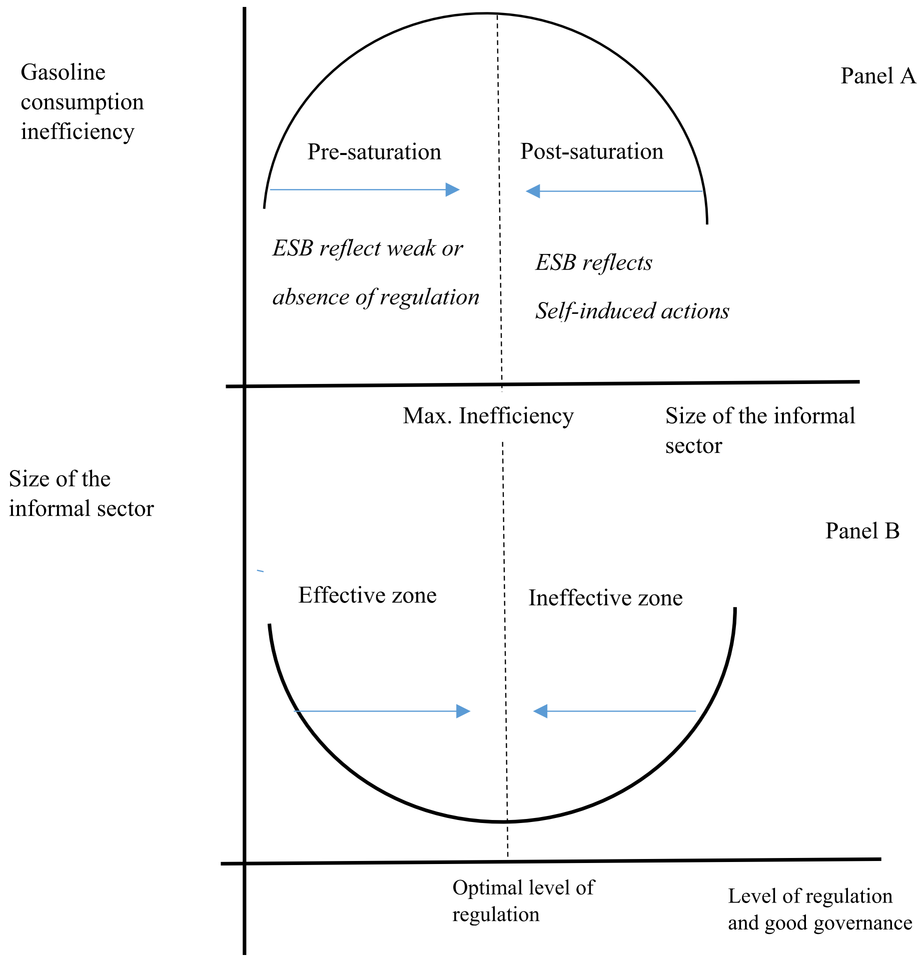

Fig. 1. Schematic display of ESB, informal activity, and government regulation.
(Source: Author's own construction)

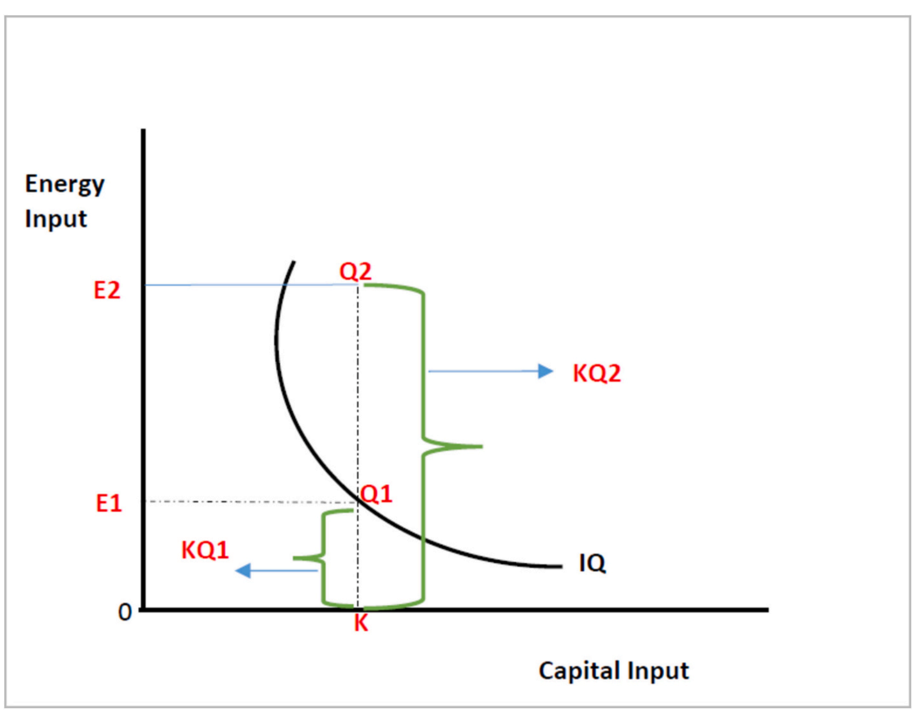

Fig. 2. Diagrammatic illustration of technical efficiency in energy input (adopted from Adom, 2021b).

Gasoline consumption levels can vary significantly among countries with similar gasoline prices and income levels. Such variations may be attributed to factors like population density (see Burke and Nishitateno, 2011; Filippini and Heimsch, 2016; Duffield and Hankla, 2011), temperature variation (Zhang et al., 2022; Wenz et al., 2017; Davis and Gertler, 2015; Li et al., 2019), technological innovation (Tirkaso and Gren, 2020; Raghoo and Surroop, 2020; Bakhat et al., 2017; Adom et al., 2016), and structural factors (Duffield and Hankla, 2011), among others.

Densely populated areas might suggest reduced travel time and, consequently, lower gasoline consumption. Burke and Nishitateno (2011) and Duffield and Hankla (2011) found that densely populated areas are associated with lower per capita gasoline consumption. Conversely, high population density can lead to heavy traffic, particularly in developing countries where public transportation is unreliable. In such economies, owning a vehicle becomes a necessity rather than a luxury. Temperature variation can affect the number of cooling and heating hours required in the transport sector. Additionally, monotonic changes such as increased travel demand, efficiency gains, and technological progress in the transport sector can influence gasoline consumption levels. Lastly, the structure of the production mix can also impact gasoline consumption intensity. For example, the agricultural and industrial sectors are typically more gasoline-intensive than the service sector (Duffield and Hankla, 2011).

Motivated by the aforementioned studies, I controlled for the effects of climate/weather changes (measured using temperature changes (temp)), demographic effects (measured using population density (pd)), and structural shifts in production (measured using the share of the service sector value added as a percent of gross domestic product [sva]). To account for monotonic changes in the transport sector, following Bakhat et al. (2017), Tirkaso and Gren (2020), and Raghoo and Surroop (2020), I included a trend term and its square. Another critical factor is vehicle characteristics. However, as pointed out by Basso and Oum (2007), incorporating vehicle characteristics in long-run estimates is cumbersome.

4.2. Econometric concerns and estimation

In a standard consumption model, income is not independent of the two-sided error term (Filippini and Zhang, 2016; Adom, 2021b; Adom et al., 2023 inter alia). This is because several factors that correlate with gasoline consumption may also correlate with income. Omitting these important variables can lead to an inconsistent estimate of the income coefficient and, consequently, incorrect identification of the frontier equation and inefficiency term. While including these control variables improves identification, it does not fully address endogeneity concerns. Measurement error is another potential issue. In Africa, it is well established that income data are of very poor quality (Adom and Amoani, 2021). Additionally, economic growth is intertwined with the environment. This interdependency between the economy and the environment suggests possible reverse causality from gasoline consumption to real income changes (Adom et al., 2023).

I addressed the identification concern by using two instrumental variables for real income per capita: the legal system origin of the country and the share of arable land. The link between the economy and law dates back to Coase's seminal 1960 work on social cost, which later became known as the Coase theorem. Coase's problem of social cost paved the way for numerous studies on the role of law in economic growth. Before the 1990s, research focused on how law influences individual incentives and firm behavior using micro-level data. However, by the end of the 1990s, a macro approach emerged to study the interaction between the economy and the legal system.

A prominent scholar in this field is La Porta et al. (1997, 1998, 1999), who used cross-country data to examine how different legal rules or systems affect economic growth. Their study identified two primary legal systems: common law (originating in England) and civil law (originating in France). They found that common law is more protective of investors compared to civil law. This protection reduces the risk of expropriation by corporate insiders, making investors more willing to finance firms, thereby promoting the growth of financial markets and the economy. La Porta et al. (2002) noted that state ownership of banks is more common in civil law countries and is usually associated with poor economic and financial performance. Similarly, Levine (2002) observed that the efficiency of legal systems in enforcing legal rights has a strong positive relationship with long-term economic growth.

The superiority of common law over civil law is attributed to several factors: First, common law provides judges with greater independence from other branches of government, empowering them to conduct constitutional reviews of legislation and administrative acts (La Porta et al., 1997, 1998; Xu, 2011). Second, as Damaska (1986) noted, civil law is "policy-implementing" while common law is "dispute-resolving." Common law relies on judicial resolution of private disputes rather than legislative solutions to social problems. Third, countries with common law have greater respect for jurisprudence as a source of law, allowing judges broader interpretation powers. This enables courts to adapt and create laws as needed. Consequently, common law countries enjoy more economic freedom than civil law countries (La Porta et al., 2004).

An efficient legal system can influence gasoline consumption only indirectly through the economic activity it facilitates. Since effective gasoline consumption is primarily driven by economic activity, an efficient legal system will support, but not directly drive, gasoline consumption. Therefore, it is expected that the legal system origin of a country will not have a direct effect on gasoline consumption.

Similarly, several studies have explored the connection between economic performance and land availability (Barbier, 2004; Roe, 2006; Alesina et al., 2005; Alouini and Hubert, 2019, inter alia), though there is no consensus on whether land directly drives economic growth. Barbier (2004) indicated that agricultural land might correlate with the cyclical patterns of economic growth. Besides being a productive resource, abundant land could signal significant natural resource endowment potential (Alesina et al., 2005), which might include energy resources. However, substantial natural resource endowments could also entail higher management costs for public services and transportation (Alesina et al., 2005).

Moreover, land availability is not expected to directly affect gasoline consumption. While large land size may indicate substantial energy potential, this potential does not equate to effective demand. Effective demand arises from the economic activities that occur. Therefore, I also expect the share of arable land to be externally valid as an instrumental variable. Testing for the external exclusion restriction of instruments is not currently feasible within the Stochastic Frontier Analysis (SFA) framework. Consequently, I relied on a pseudo-maximum likelihood regression and D'Haultfoeuille and Sasaki (2021) exclusion test for instruments to test the external validity of the instrumental variables.

In the inefficiency equation, the size of the underground/shadow economy (se) and its square were included as explanatory variables in the first stage of analysis (see Basbay et al., 2016) to test the first part of TPCH test. The validity of the first claim requires a statistically significant inverted U-shaped relationship between the environmental outcome indicator and the size of the informal economy. Subsequently, I controlled for additional independent variables, including the price of gasoline, trend, and the square of the trend, as well as measures of financial sector development [bcp] (see Adom, 2021a) and the cost of starting a business [csb]. The rationale is that in the inefficiency equation, the square term of the underground economy may capture structural shifts or other unobserved/observed heterogeneities.

Methodologically, addressing endogeneity within the Stochastic Frontier Analysis (SFA) framework has been challenging until recently. As noted by Kutlu (2010), Greene (2011), and Mutter et al. (2013), the complex nature of the SFA specification and the special nature of the error term make it very difficult to model endogeneity in SFA. Endogeneity causes the maximum likelihood estimator to produce inconsistent parameter estimates. To address this problem, the standard procedure involves modeling the joint distribution of the endogenous and dependent variables before maximizing the log-likelihood function (Karakaplan and Kutlu, 2017a).

Some previous studies have relied on a pseudo-econometric approach (Filippini and Zhang, 2016), where fitted values are used for endogenous variables within the SFA framework. However, this approach is considered inappropriate (see Amsler et al., 2016) and inefficient as it may require a bootstrap procedure to correct standard errors (Karakaplan and Kutlu, 2017a). Kutlu (2010) proposed a maximum likelihood procedure allowing some regressors in the frontier equation to correlate with the two-sided error term. Tran and Tsionas (2013) recommended estimating Kutlu's model using the generalized method of moments, assuming the one-sided error term is independent of the two-sided error term.

Tran and Tsionas (2015) allowed the endogenous input to correlate with the composed error term (i.e., noise and inefficiency term). Amsler et al. (2016) further refined this by allowing the endogenous input variable to correlate separately with the statistical noise and inefficiency term. This model was later extended by Amsler et al. (2017) to allow both the endogenous input and environmental variables to correlate with the statistical noise and inefficiency terms. Karakaplan and Kutlu (2017a) proposed a maximum likelihood procedure to handle endogeneity in the frontier equation, inefficiency equation, or both for cross-sectional data, showing that their approach outperforms standard SFA procedures that do not consider endogeneity.

The above procedures primarily used a cross-sectional data approach. Karakaplan and Kutlu (2017b) extended their earlier model to include panel data, addressing the dual endogeneity problem but not the heterogeneity problem (see Wang and Ho, 2010; Kutlu et al., 2019). According to the authors, efficiency estimates are likely to be more reliable with panel data than with cross-sectional data. In this study, I adopted the panel SFA approach of Karakaplan and Kutlu (2017b), implemented in Stata as xtsfkk (see Karakaplan, 2017 and Karakaplan, 2022).

Specifying Eq. 2 in a more explicit form and following the procedure of Karakaplan and Kutlu (2017b), this study presents the empirical model as follows:

$$gc_{it} = \boldsymbol{x}_{gcit}^{'}\boldsymbol{\beta} + v_{it} - su_{it} \tag{3}$$

$$\boldsymbol{x}_{it} = \boldsymbol{Z}_{it}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_{it} \tag{4}$$

where

$$s = \begin{cases} 1 \ \textit{for production function} \\ \quad -1 \ \textit{for cost function} \end{cases} \tag{5}$$

$\boldsymbol{x}_{gcit}$: represents a vector of endogenous and exogenous variables.

$\boldsymbol{x}_{it}$: denotes a vector of all endogenous variables.

$\boldsymbol{Z}_{it}$: denotes a vector of all exogenous variables.

v_{it} & $\boldsymbol{\varepsilon}_{it}$ are two-sided noise terms assumed to be normally distributed.

u_{it}: denotes a one-sided error term that captures the inefficiency term.

One implication of the model is that the inefficiency term (u_{it}) and the two-sided error term (v_{it}) can correlate with the vector of endogenous variables ($\boldsymbol{x}_{it}$) but they are conditionally independent given the vector of endogenous ($\boldsymbol{x}_{it}$) and exogenous variables ($\boldsymbol{z}_{it}$). Similarly, u_{it} and ε_{it} are conditionally independent given the vector of endogenous and exogenous variables (Karakaplan and Kutlu, 2017b). Eq. 6 can be used to predict energy efficiency, where ϕ is the standard normal probability density function.

$$EEF_{it} = exp(-\mathrm{E}[u_{it}|e_i) = exp\left(-h_{it}\left(\mu_{i^*} + \frac{\sigma_{i^*}\phi\left(\frac{\mu_{i^*}}{\sigma_{i^*}}\right)}{\Phi\left(\frac{\mu_{i^*}}{\sigma_{i^*}}\right)}\right)\right) \tag{6}$$

The method relies on the standard Durbin–Wu–Hausman test to assess endogeneity. Rejecting the null hypothesis suggests a preference for endogenous SFA over traditional SFA. To test the validity of the instrument choice, the Cragg–Donald F-statistics is used to determine whether the instruments are weak. These results are reported in the results tables.

The current approach by Karakaplan and Kutlu (2017b) does not provide a direct way to test the exclusion restriction of the instruments, which states that the instrumental variable should not directly influence the outcome variable (i.e., gasoline consumption) but only indirectly through the endogenous variable (i.e., real GDP per capita). To test the exclusion restriction, I adopted a pseudo regression approach. First, I included the instrumental variables in the frontier part of the SFA and then applied the random effect maximum likelihood estimator with bootstrap standard errors to estimate the regression. I then imposed a zero restriction on the coefficient of the instrumental variable (i.e., legal system origin of the country and share of arable land). Failing to reject the null hypothesis of zero effect implies that the instrument choice satisfies the exclusion restriction. This was complemented with a formal instrument exclusion test by D'Haultfoeuille and Sasaki (2021), where the null hypothesis is that the exclusion restriction is satisfied. External validity is confirmed by failing to reject the null hypothesis.

There are two data-related issues that, if not properly handled, can introduce bias in the estimate of gasoline consumption efficiency: outliers and unobserved/observed heterogeneity. To address outliers, I chose winsorizing over data trimming (Wilcox, 1998), as the former retains critical information, which is particularly important for a data-driven technique like SFA. Wilcox and Keselman (2003) recommend retaining at least 60 % of the original data for good statistical properties (i.e., using the 20th and 80th percentiles for the lower and upper data points). For the independent variables in the frontier and inefficiency equations, I winsorized the data using the 2.5th and 97.5th percentiles for the lower and upper ends, respectively, thus retaining 95 % of the original data. This approach was consistently applied across various models using the winsorized data.

For the dependent variable (gasoline consumption), I initially used the 2.5th and 97.5th percentiles and then adjusted by 2.5 percentile intervals, with the final winsorizing occurring at the 12.5th and 87.5th percentiles. Consequently, I retained 95 %, 90 %, 85 %, 80 %, and 75 % of the original data across different winsorized windows. The varying winsorizing windows for gasoline consumption were necessary due to the highly skewed nature of the data.

To address unobserved and observed heterogeneity bias in the data, I adopted the model by Kutlu et al. (2019), which handles both endogeneity and unobserved/observed heterogeneities. This counterfactual model uses the original, non-winsorized data. Given that the instrumental variables do not exhibit rapid temporal variation, I assume that countries are in their long-run position and therefore interpret the coefficients as long-run coefficients.

4.3. Other robustness checks

I conducted five additional robustness checks on the results. For all tests of robustness, except the last one, I applied winsorized data at the 12.5th and 87.5th percentiles. In the first robustness check, I explored the sensitivity of the instruments to determine if they influence the outcome. Second, using the size of the informal economy, already estimated data in a regression raises concerns about the validity of the standard errors. To address this, I weighted the regressors by the inverse of the standard deviation of the informality size.

Third, given that energy efficiency is unlikely to change rapidly in the short term due to challenges in adopting new technologies, differing discount factors among energy users, and various market imperfections (Adom et al., 2023), and considering that the instrumental variables lack significant temporal variation, relying on short-term data might lead to inaccurate estimates of long-term effects. To address these issues, I used a five-year average regression. This approach helps account for potential rigidity in energy efficiency and instruments, common time-varying shocks or cross-sectional dependence, and business cycle effects.

Fourth, informality can stem from impediments faced by the formal sector or constraints imposed on it. For example, high costs of starting a business and limited access to credit can push businesses into the informal sector. I used the cost of starting a business as a proxy for the size of the informal economy. This approach is intended to address potential measurement error issues with the size of the informal economy.

Finally, the inefficiency term might capture unobserved or observed heterogeneity, potentially introducing bias into the efficiency estimates. Additionally, winsorizing data can create artificiality, which may complicate the identification of the frontier and inefficiency equations. This is because addressing extreme values can reduce the variability in the data, which might be crucial for accurately identifying inefficiency.

To address these issues—both the omitted unobserved/observed heterogeneity and the artificiality introduced by winsorizing—I adopted the approach by Kutlu et al. (2019), which uses the original data values. The model by Kutlu et al. addresses both endogeneity and heterogeneity biases (observed and unobserved). By using the original data, this approach helps to obtain more accurate efficiency estimates and assess the extent of bias introduced by winsorizing.

4.4. Regulation and good governance channel of informality

To examine regulation and good governance as channels for influencing the informal economy, I estimated Eq. 7, which postulates a non-linear relationship between the size of the informal economy and regulation and good governance. Several studies have documented the negative effect of regulation and good governance on the size of the informal economy (see Barra and Papaccio, 2024). In this study, however, I anticipate a U-shaped relationship, with a statistically significant negative linear effect and an insignificant positive quadratic effect. This suggests that the effectiveness of regulation is limited to periods when the informal economy is experiencing a normal growth rate, within a range that remains manageable by the government. In this equation, I controlled for country-fixed effects, as well as time-varying and time-invariant effects. Since governance and regulation, along with the informal economy, may not exhibit temporal variation in the short term, I first performed a five-year average regression and then varied the regression window within five-year intervals. This latter approach improves the causal claim by examining whether subsequent changes in the size of the informal economy are accompanied by successive changes in regulation and good governance.

$$lnse_{it} = \alpha + \beta lngovindex_{it} + \delta lngovindex_{it}^2 + + \delta_i + \tau_t + e_{it} \quad (7)$$

In these regressions, I postulate that regulation and governance indicators exert a statistically significant and negative effect on the size of the informal economy. For the second part of TPCH test, $\beta < 0$ (statistically significant) and $\delta > 0$ (statistically insignificant). Thus, the validity of the TPCH requires the simultaneous satisfaction of the first and second claims. The validity of these claims provide the basis for including measures of regulation and good governance indicators in the inefficiency equation as a channel of effect for the informal economy. Recognizing that the size of the informal economy can indicate the strength of regulation and good governance, I later introduced six direct measures of regulation and good governance into the inefficiency equation (see Jollands et al., 2011; IEA, 2010).

4.5. Data

This study utilized unbalanced panel data from 44 African countries, covering the period from 1995 to 2019[2] (see Fig. 3 for a list of the included countries). Table 1 provides definitions of the variables, units of measurement, data sources, and descriptive statistics. The size of the shadow economy was estimated using the multiple indicators multiple causes (MIMIC) approach (Medina and Schneider, 2018). The original MIMIC approach, which uses GDP per capita or the GDP growth rate as both cause and indicator variables, has limitations. Medina and Schneider (2018) addressed this issue by employing a light intensity approach instead of GDP as the indicator variable. To calibrate the relative MIMIC estimates of the informal economy, they used the predictive mean matching (PMM) method proposed by Rubin (1987). As noted by Medina and Schneider (2018), this represents the first attempt to incorporate a light intensity indicator within the MIMIC framework and to use a full alternative method, such as PMM, for calibrating the relative estimates of the shadow economy.

Fig. 3 displays the average size of the informal economy as a percentage of GDP from 1995 to 2019. The reference line represents the group sample average. The plot indicates that in Africa, Mauritius and South Africa have relatively smaller informal economies compared to Zimbabwe, Nigeria, Gabon, Tanzania, Benin, and Gambia, where the informal sector constitutes a larger share of the economy. While some countries show similar informal market sizes, there is considerable variation across countries. Two implications arise from Fig. 3 that could affect model estimates. First, the substantial differences between the extremes suggest that extremely high or low values might influence the mean results. Second, the variability among countries may be attributed to both observed and unobserved heterogeneity, which could also affect the results. To address these issues, I tackle potential outliers and heterogeneity. First, I winsorize the data to manage outliers. Then, I apply the model proposed by Kutlu et al. (2019) to account for heterogeneity in the inefficiency term. I tested for heterogeneity effects of panel unit level.

Six different indicators were used in this study to capture regulation and good governance: rule of law *(rl)*, regulatory quality *(rq)*, control of corruption *(ccr)*, government effectiveness *(ge)*, political stability and absence of violence/terrorism *(psav)*, and voice and accountability *(va)*. These variables originally range from -2.5 to $+2.5$ in standard normal units. For this study, they were rescaled to range from 0 to 1, with higher values indicating better outcomes.

The principal component analysis (PCA) technique was used to derive a composite index of regulation and good governance from the six indicators. This approach helps address concerns about potential high correlations between the indicators. Before applying PCA, I tested the suitability of the data using the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy. The overall KMO measure yielded a value of 0.9114, which is well above the threshold value of 0.5 required for PCA. Additionally, the KMO value for each variable exceeded the 0.5 threshold.

Applying PCA produced six components. Among these, the first component was the only one positively correlated with all six indicators (see Stata profile codes). This first component accounted for approximately 80.16 % of the total variance. Its eigenvalue was significantly

[2] This dataset represents the most updated period available. Extending the data beyond 2019 could be problematic due to the impact of COVID-19. The total to partial lockdowns implemented globally halted economic activity, especially in the transport sector, where movements were highly restricted, and trade in the informal sector was limited. Consequently, data from 2020 to 2022 are likely to deviate significantly from the normal long-term trend, potentially affecting long-term inferences.

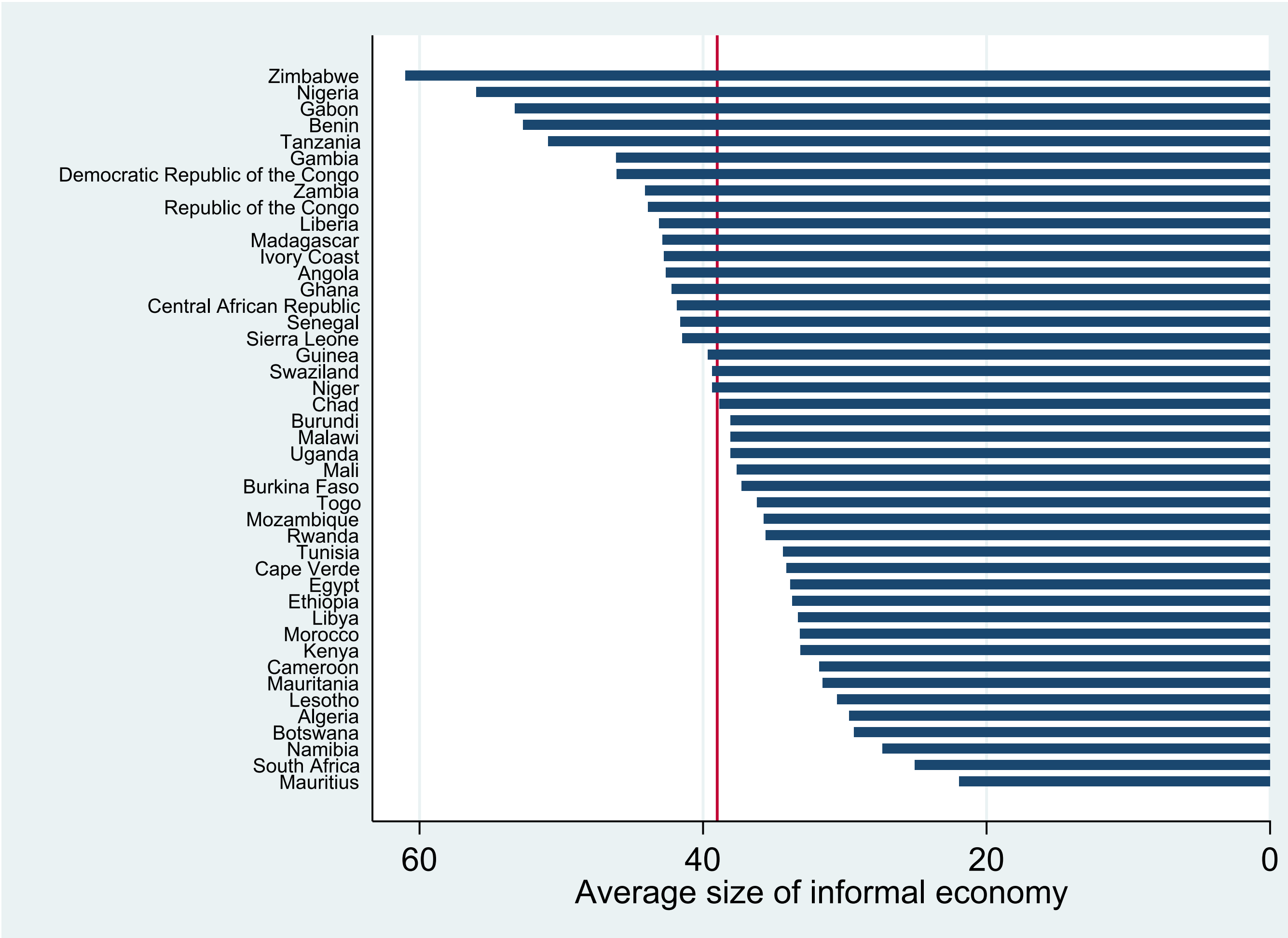

Fig. 3. Plot of mean size of informal economy % GDP (1995 to 2019).
(Source: Author's construction)

higher than those of the other components (see Stata profile codes). Based on these findings, the first component was used as the composite index of regulation and good governance.

4.6. Model diagnostics

Several preliminary diagnostic tests were performed. First, the log-likelihood ratio test was used to choose between the Cobb-Douglas and Translog functional forms. It is important to note that the strong requirement for model preference is evidence of endogeneity. Thus, the choice between functional forms becomes valid only when the first requirement of endogeneity is satisfied. This implies that while the log-likelihood test may prefer Translog over Cobb-Douglas, if endogeneity is only present in the Cobb-Douglas case, the overriding condition for model selection should be evidence of endogeneity. This consideration becomes redundant only if there is endogeneity in both functional forms. In such a case, the log-likelihood test is used to choose the best functional type.

In this study, both the Cobb-Douglas and Translog functional forms show evidence of endogeneity. For the Cobb-Douglas model, the eta endogeneity test statistic is 35.17, which is statistically significant at the 1 % level. Similarly, the eta endogeneity test statistic for the Translog model is 29.78, also statistically significant at the 1 % level. Therefore, I rely on the log-likelihood test to choose the appropriate functional form. The LR chi-square statistic is 159.65 with a p-value of 0.000. Additionally, the Akaike and Bayesian information criteria are smaller for the Translog model than for the Cobb-Douglas model, suggesting that the Translog model better fits the data. Consequently, the Translog functional form is estimated in this study.

One major drawback of the Translog function is multicollinearity, which could lead to nonsensical results. Multicollinearity was tested using the variance inflation factor (VIF). The mean VIF is 4.38, and none of the individual VIFs exceed the threshold value of 10, suggesting a lack of serious multicollinearity in the model. Next, the Coelli (1995) test was applied to choose between a cost or production Translog SFA. The summary statistics show that the least square residual is positively skewed. The Coelli (1995) test produces a statistic of 11.726, confirming the positive skewness of the least square residuals. Therefore, the production Translog SFA was estimated.

Estimating the translog function helps control for omitted variable bias that might originate from variable interactions and nonlinearities. In the energy literature, studies have indicated the presence of nonlinearities and interactions between price, income, temperature, demography, and structural changes in output (see: Fouquet, 2015 & 2012; Gately and Huntington, 2002; De Cian et al., 2007; Petrick et al., 2010; Henley and Peirson, 1996 & 1998; Adom et al., 2023; Akorli and Adom, 2023; Marquez, 1994, inter alia).

5. Results and discussion

The presentation of the empirical results proceeds as follows. First, I present the results for the frontier determinants. I then discuss the

Table 1
Variables, Measurements, Definitions, and Descriptive Statistics.

Variables	Unit of measurement	Definition of the variables	obs	mean	sd	min	max
gc	*Gasoline consumption in thousand barrels per day (bpd).*	Gasoline consumption includes the consumption of conventional gasoline, all types of oxygenated gasoline, including gasohol, and reformulated gasoline, but excludes the consumption of aviation gasoline. Volumetric data on blending components, such as oxygenates, are not included in the data on finished motor gasoline until the blending components are blended into the gasoline. Source: US Energy Information Administration	1125	17.707	42.983	0.1	337.74
gp	*Gasoline price in US$.*	Fuel prices refer to pump prices of the most widely sold grade of gasoline. Prices have been converted from local currency to U.S. dollars. Source: German Agency for International Cooperation	768	0.9235	0.3903	0.05	2.3
rgdp	*Real gross domestic product* per capita. *In 2015, constant prices (US$)*	GDP per capita is the gross domestic product divided by the midyear population. GDP is the sum of the gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for the depreciation of fabricated assets or for the depletion and degradation of natural resources. Data are in constant 2015 US dollars. Source: World Bank Development Indicators	1116	1959.5	2210.0	217.62	13,729.3
temp	*Temperature*	*Annual temperature in degrees Celsius. Source:* WB Climate Change Knowledge Portal	1034	24.301	3.313	12.628	30.477
se	*Shadow economy percentage of GDP*	The shadow economy as a percentage of total annual GDP. Detailed methodology for the estimations can be obtained from the following International Monetary Fund working paper by Leandro Medina and Friedrich Schneider (2018): Shadow Economies Around the World: What Did We Learn Over the Last 20 Years?. Source: Medina and Schneider (2018)	924	38.847	8.9405	19.23	69.08
sva	*Service val. Added percentage of GDP.*	Services correspond to ISIC divisions 50–99. They include value added in wholesale and retail trade (including hotels and restaurants), transport, and government, financial, professional, and personal services such as education, health care, and real estate services. Also included are imputed bank service charges, import duties, and any statistical discrepancies noted by national compilers and discrepancies arising from rescaling. Value added is the net output of a sector after adding all outputs and subtracting intermediate inputs. It is calculated without making deductions for the depreciation of fabricated assets or the depletion and degradation of natural resources. The industrial origin of value added is determined by the International Standard Industrial Classification (ISIC), revision 3 or 4. Data are in constant 2010 US dollars. Source: World Bank Development Indicators	1080	45.161	10.102	10.88	77.84
pd	*Population density, persons per square kilometer of land area*	Population density is the midyear population divided by land area in square kilometres. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship, except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin. Land area is a country's total area, excluding the area under inland water bodies, national claims to continental shelf, and exclusive economic zones. In most cases, the definition of inland water bodies includes major rivers and lakes. Source: Food and Agriculture Organization	1125	80.123	112.41	2	624
rl	*Rule of law. Index rescaled from − 2.5 to + 2.5 to range from 0 to 1.*	The index for Rule of Law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, particularly the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Source: World Bank Governance Indicators	945	0.43471	0.2076	0	1
ge	*Government effectiveness. Index rescaled from − 2.5 to + 2.5 to range from 0 to 1.*	The index of government effectiveness captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. Source: World Bank Governance Indicators	942	0.3935	0.1883	0	1
ccr	*Control of corruption. Index rescaled from − 2.5 to + 2.5 to range from 0 to 1.*	The Index for Control of Corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture of the state by elites and private interests. Source: World Bank Governance Indicators	945	0.3613	0.2042	0	1
rq	*Regulatory quality. Index rescaled from − 2.5 to + 2.5 to range from 0 to 1.*	The Regulatory Quality Index captures perceptions of the ability of the government to develop and implement sound policies and regulations that permit and promote private	945	0.477	0.1677	0	1

(*continued on next page*)

Table 1 (*continued*)

Variables	Unit of measurement	Definition of the variables	obs	mean	sd	min	max
		sector development. Source: World Bank Governance Indicators					
va	*Voice and Accountability. Index rescaled from − 2.5 to + 2.5 to range from 0 to 1.*	The index for Voice and Accountability captures perceptions of the extent to which citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media. Source: The World Bank Governance Indicators	945	0.4672	0.235	0	1
psav	*Political stability and absence of violence/ terrorism measures. Index rescaled from − 2.5 to + 2.5 to range from 0 to 1.*	The index of Political Stability and Absence of Violence/ Terrorism measures perceptions of the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including politically motivated violence and terrorism. The index is an average of several other indexes from the Economist Intelligence Unit, the World Economic Forum, and the Political Risk Service. Source: World Bank Governance Indicators	945	0.5584	0.2189	0	1
csb	*Cost of starting a business % of income* per capita.	This indicator includes all official fees and fees for legal or professional services if such services are required by law. The company law, commercial code, and specific regulations and fee schedules are used as sources for calculating costs. The indicator excludes bribes. World Bank Development Indicators	733	104.2975	181.26	0	1540.2
bcp	*Bank credit to the private sector percentage GDP*	Domestic credit to the private sector by banks refers to financial resources provided to the private sector by other depository corporations (deposit taking corporations except central banks), such as through loans, purchases of nonequity securities, trade credits, and other accounts receivable, that establish a claim for repayment. For some countries, these claims include credit to public enterprises. Source: World Bank Governance Indicators	1086	19.10519	17.70634	0.45	104.85
arl	*Arable land percentage total land.*	Arable land includes land defined by the FAO as land under temporary crops (double-cropped areas are counted once), temporary meadows for mowing or pasture, land under market or kitchen gardens, and land temporarily fallow. Land abandoned because of shifting cultivation is excluded. Source: World Bank Development Indicators	1125	13.76551	12.908	0.3	48.7
lso_fr	*Legal system origin of the country*	1 if France (Civil law) and 0 if Britain (Common Law)	1125			0	1

Source: Authors' own construction.

estimated gasoline consumption efficiency and the gasoline energy savings potential in energy units for the countries in the sample. Next, I analyze the drivers of gasoline consumption inefficiency, followed by robustness tests for the results. Finally, I conclude this section with a discussion on the regulation and good governance channels concerning the gasoline consumption inefficiency and informality nexus. For the SFA regression, I present the full results tables in the appendix, which includes frontier equation, inefficiency equation, and estimate of gasoline efficiency.

5.1. Energy efficiency and informality

5.1.1. Frontier determinants

Table 2A originally has three parts. The first part contains information about the determinants of gasoline consumption (i.e., frontier equation). The second part provides information about the drivers of gasoline consumption inefficiency. The third part includes model diagnostics. In all models presented in Table 2A, the legal system origin and the share of arable land were used as instrumental variables. Table 2A presented here includes only the last two sections to minimize the table's textual content (see Appendix A for the full version).

Model 1 is the baseline regression, which does not involve winsorized data. For the remaining models, the regression involved winsorized data for both dependent and independent variables. Specifically, for models II–VII, the independent variables were winsorized at the 2.5th and 97.5th percentiles. The dependent variable was initially winsorized at the 2.5th and 97.5th percentiles in Model II, with subsequent adjustments of both the lower and upper limits of the data distribution by 2.5 percentile intervals. The final winsorizing window occurred at the 12.5th and 87.5th percentiles in Models VI and VII. The highly skewed nature of the gasoline data motivated these differential levels of winsorizing.

Consistently, the eta endogeneity test in Table 2A confirms the presence of endogeneity, validating the choice of endogenous SFA over traditional exogenous SFA. Additionally, the test for instrument relevance indicates that the instrumental variables are not weak, with F-statistics exceeding the threshold value at the 10 % critical value (see the bottom of Table 2A). The instrumental variables also satisfy the exclusion restriction, as indicated by the exclusion restriction test at the bottom of the table. The D'Haultfoeuille et al. test for the instruments produced a KS statistic of 2.442 (with a *p*-value of 0.721) for the share of arable land and a KS statistic of 3.219 (with a p-value of 0.119), suggesting a failure to reject the null hypothesis of valid exclusion restriction of instruments.

See appendix A for the frontier results. Consistently, the level effects of real GDP per capita, temperature, gasoline prices, and population density (except in the baseline regression) are negative and statistically significant. Conversely, the level effects of service sector value added are positive and statistically significant, all other things being equal. The coefficients for the interactive variables are largely statistically significant. Notably, the effect of temperature depends not only on itself but also on other factors, such as the level of income (De Cian et al., 2007; Petrick et al., 2010). The interaction between income and temperature is negative and statistically significant, similar to findings by Petrick et al. (2010). Comparatively, richer countries or households can better adapt to temperature changes and afford more energy-efficient technologies than poorer ones. Therefore, with rising temperatures, the decline in energy consumption is expected to be steeper for richer countries or households. Consequently, an increase in income should increase (in absolute terms) the temperature elasticity of demand.

The response of gasoline energy demand to price also depends on income and temperature (Henley and Peirson, 1996 & 1998). The interaction between price and income is negative and statistically significant, which confirms the findings of Adom et al. (2023). A

Table 2A
Gasoline consumption inefficiency and informality.

Inefficiency equation	MI Baseline	MII	MIII	MIV	MV	MVI	MVII
Dep var.: $ln\left(\delta^2_{u}\right)$							
constant	−14.42***	−12.86**	−7.616	−9.690	−14.45*	−18.62**	−3.126
	(4.089)	(6.547)	(7.932)	(7.750)	(8.500)	(8.503)	(10.22)
lnse	9.468***	8.458**	5.568	6.601	9.135*	10.07**	0.774
	(2.259)	(3.591)	(4.338)	(4.244)	(4.673)	(4.586)	(5.597)
lnse^2	−1.337***	−1.204**	−0.818	−0.955*	−1.296**	−1.346**	−0.168
	(0.311)	(0.490)	(0.590)	(0.579)	(0.640)	(0.619)	(0.778)
trend	−0.0129	−0.0345**	−0.0594***	−0.0681***	−0.0766***	0.0781***	0.293***
	(0.0123)	(0.0141)	(0.0176)	(0.0182)	(0.0179)	(0.0257)	(0.0828)
trend^2	0.00107**	0.00177***	0.00271***	0.00294***	0.00313***	−0.00192**	−0.00734***
	(0.000481)	(0.000553)	(0.000695)	(0.000721)	(0.000718)	(0.000943)	(0.00254)
lngp	−0.182***	−0.231***	−0.333***	−0.348***	−0.327***	0.0578	0.0175
	(0.0560)	(0.0711)	(0.0792)	(0.0840)	(0.0894)	(0.125)	(0.159)
lncsb							0.195***
							(0.0622)
lnbcp							−0.325***
							(0.0888)
Dep var.: $ln\left(\delta^2_{w}\right)$							
constant	−2.674***	−2.749***	−2.773***	−2.816***	−2.852***	−3.279***	−3.752***
	(0.0573)	(0.0561)	(0.0561)	(0.0561)	(0.0562)	(0.0634)	(0.0700)
/eta1_lnrgdp	3.255***	2.269***	1.657***	1.609***	1.804***	2.366***	1.782***
	(0.646)	(0.541)	(0.529)	(0.491)	(0.547)	(0.526)	(0.499)
/e1	−0.170***	0.160***	−0.160***	−0.160***	−0.160***	−0.154***	0.108***
	(0.00458)	(0.00431)	(0.00431)	(0.00431)	(0.00431)	(0.00468)	(0.00355)
Eta_end test	25.39***	17.57***	9.81***	10.75***	10.88***	20.29***	12.76***
Weak IV test	99.45***	93.62***	92.22***	92.02***	89.82***	75.42***	48.69***
Exclusion test_P-val	0.4455	0.9334	0.8009	0.8248	0.6246	0.3519	0.3519
Average TEE (%)	21.41	26.52	27.92	30.32	32.35	42.37	55.62
Median TEE (%)	11.42	14.72	17.20	23.37	27.11	37.87	53.60
Log-likelihood	46.66	123.02	136.30	154.37	169.26	266.55	479.37
Observations	689	689	689	689	689	542	462
Countries	44	44	44	44	44	44	44

Standard errors are in parentheses. Symbols indicate significance at the 1 % (***), 5 % (**), and 10 % (*) levels. MI is the baseline model, where the original data are used in the regression. For MII–MVII, the variables (both dependent and independent) are winsorized. For the independent variables, a winsorized window of 2.5 and 97.5 percentiles was maintained, meaning 95 % of the original data for these variables remained unchanged for MII–MVII. For the dependent variable (gasoline consumption), 95 %, 90 %, 85 %, 80 %, and 75 % of the original data were retained in MII, MIII, MIV, MV, and MVI, respectively. Model VII is an extension of Model VI, where the cost of starting a business and bank credit to the private sector were controlled for in the inefficiency equation. In all models, the legal system origin of the country and the share of arable land were used as instruments for real GDP per capita.
Source: Author's construction.

simultaneous increase in income and the price of energy could either reduce or increase energy consumption, depending on the relative growth rates of income and energy prices.

Income, gasoline prices, temperature, and population density exhibit consistent monotonic negative effects on gasoline consumption, all other things being equal, ruling out the possibility of nonlinearities in these effects. Specifically, the case of income disputes the existence of an energy Kuznets curve, contrasting the findings of Adom et al. (2023) and Adom (2021b). These studies considered total final energy consumption, which differs technically from gasoline consumption. Additionally, the consistent monotonic negative effects of gasoline prices reject the hypothesis of a tolerable price range for energy products, as found in Adom (2017) for electricity consumption in Ghana and in Adom et al. (2016), Adom et al. (2018) for gasoline consumption in West Africa and Ghana, respectively.

In contrast, the share of the service sector shows an inverted U-shaped effect on gasoline consumption, suggesting that extensive growth in the service sector drives gasoline consumption down. This could be due to technological advancements, as the interaction of the trend and the share of service sector output produces a statistically significant and negative effect on gasoline consumption. This finding supports the claims of Suri and Chapman (1998) and Syrquin and Chenery (1988) that the growth of less energy-intensive sectors reduces overall energy consumption.

A major concern is the extent of artificiality introduced in the coefficient estimates due to winsorizing the data. I provide a qualitative assessment of this issue by considering the deviation of the regression coefficients in the winsorized model from the baseline regression (i.e., the coefficient size). In this context, Model MVI generally presents the minimum risk qualitatively, making it the preferred model in this study.[3]

5.1.2. *Energy efficiency and informality*

Table 2A shows the regression of gasoline consumption inefficiency on the size of the shadow/informal economy. The effect of the informal economy on gasoline consumption inefficiency consistently exhibits an inverted U-shaped relationship, supporting the findings of Basbay et al. (2016). This implies a point of saturation in the relationship between gasoline inefficiency and the size of the informal economy. According to Model VI, this saturation point occurs when the informal economy reaches 42 % of the national GDP. Beyond this point, the informal sector may become too large to be effectively controlled by the government. Consequently, regulations aimed at promoting pro-environmental behaviors may become ineffective due to difficulties in enforcement and monitoring.

At the post-saturation stage, it is expected that pro-environmental

[3] I adopted the recommendation of 20 % winsorizing by Wilcox and Keselman (2003), but the model failed to converge.

behavior might become self-motivated due to the impact of inefficient gasoline use and climate change on the sector's operations. During this stage, informal sector operators might naturally shift toward gasoline conservation by investing in fuel-efficient vehicles and adopting other conservation behaviors to secure production and minimize losses. In other words, the saturation point might be a crucial juncture where the transition to a low-carbon economy becomes self-motivated for informal market operators.

Before reaching saturation, however, the behavior of informal market players may largely reflect weak or absent regulation and standardization. Consequently, the growth of the informal sector could threaten gasoline consumption efficiency and environmental sustainability. Government regulation might be more effective and critical at the pre-saturation stage. In the next section, I discuss the regulation channel of the informality effect on gasoline consumption inefficiency.

Omitted variable bias can confound the above relationship. Informality emerges due to constraints faced by the formal sector, such as the high cost of starting a business and lack of access to private sector credit. This claim is supported by simple correlation and fixed-effect regression analyses. The correlation between the size of the informal economy and the cost of starting a business is 60.89 %, while the correlation between the size of the informal economy and access to credit is −55.5 %.

Table 2B presents a simple fixed-effect regression with bootstrapped standard errors, showing the relationship between informality, the cost of starting a business, and bank credit to the private sector. The results indicate a positive elasticity of informality with respect to the cost of starting a business and a negative elasticity with respect to financial development, both statistically significant. Specifically, a 10 % increase in the cost of starting a business is associated with a 0.43 % increase in the size of the informal sector, while a 10 % increase in bank credit to the private sector corresponds to a 0.734 % reduction in the size of the informal sector. Ignoring these variables in the inefficiency equation introduces omitted variable bias.

In Model MII of Table 2A, both the cost of starting a business and bank credit to the private sector are controlled for. The cost of starting a business positively affects gasoline consumption inefficiency, whereas financial sector development has a negative effect. The negative effect of bank credit on gasoline consumption inefficiency corroborates findings by Amuakwa-Mensah et al. (2018), Adom et al. (2020), Adom (2021b), and Adom et al. (2023), which suggest that financial sector development stimulates energy efficiency improvement.

Including the cost of starting a business and bank credit to the private sector reduces the coefficients' size for both the level and square terms of informality and diminishes the statistical power of the informality coefficient, yet the postulated relationship remains intact. Throughout Table 2A, the effect of gasoline prices is inconsistent; it is largely negative but occasionally positive, with the negative effect remaining statistically significant. This suggests that a price incentive mechanism can be an effective tool for improving gasoline consumption efficiency in Africa. Given that fuels are often subsidized on the continent, reducing or removing these subsidies could help achieve environmental goals.

Table 2B
Informality, cost of starting a business, and financial sector development.

Dependent variable: log of informal economy			
	MI	MII	MIII
VARIABLES			
lncsb	0.0599***		0.0431***
	(0.0108)		(0.0114)
lnbcp		−0.127***	−0.0734***
		(0.0228)	(0.0239)
constant	3.337***	3.954***	3.603***
	(0.0595)	(0.0623)	(0.101)
Observations	541	894	528
R-squared	0.217	0.217	0.302
Number of id	44	44	44

Bootstrap standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: Author's construction.

5.1.3. Estimation of gasoline consumption efficiency

From the baseline regression, the estimated mean gasoline consumption efficiency is 21.4 %, indicating significant inefficiency in gasoline consumption in Africa, similar to the findings of Adom and Adams (2020) and Akorli and Adom (2023). However, as indicated earlier in this study, the extremely low values of gasoline consumption data introduce a downward bias in the efficiency estimate. One major underlying reason for these extremely low national values in Africa could be poor data quality.

As shown in Table 2A, the efficiency estimates improve with different winsorized windows, confirming that outliers in the data might have introduced bias in the efficiency estimate. The log-likelihood values also increase with different winsorized windows, reaching a maximum in Model MVII. In that model, the estimated mean gasoline efficiency is about 56 %, which is higher than reported in Akorli and Adom (2023). Compared with the baseline estimate, this represents a downward bias in the baseline regression's efficiency estimate. The kernel density plot of efficiency distribution based on original and winsorized data in Fig. 4 confirms this observation.

Fig. 5 plots the average gasoline consumption efficiency performance at the country level from 1995 to 2019. Most of the well-performing countries are also the continent's top economic performers, suggesting that improving gasoline consumption efficiency may be driven by the level of development in the region. Due to the capital-intensive nature of energy efficiency investments, larger economies are more likely to support such investments (Adom et al., 2023; Adom, 2024). Fig. 6 illustrates the average maximum potential number of gallons of gasoline per day that countries can save through efficiency improvements alone.[4] The least-performing countries possess the greatest potential to save more energy in terms of gallons of gasoline per day.

The variation in gasoline energy efficiency performance across countries suggests that least-performing countries must conditionally converge to higher-performing countries to achieve the goals of universal access to energy and zero emission targets in Africa.

5.2. Robustness checks

In this section, I conduct several robustness checks on the results. Table 2C contains these results (see appendix B for full table). Except for Model MV, which uses the original data, all the models in this table are based on winsorized data, where 75 % of the original data for gasoline consumption and 95 % of the original data for the independent variables are retained. Consistently in Table 2C, the instrument choice proves to be valid, as indicated by the weak IV and external exclusion restriction tests. First, the results may be sensitive to the instrument choice. I address this concern in Model MI, where the share of arable land is used as the instrument. The results remain robust, discounting the influence of the choice of instrumental variable on the outcome.

Another concern is the validity of the estimated standard error in the regression, given that the informal data are estimated. I addressed this by running a weighted SFA regression, in which I weighted each regressor by the inverse of the standard deviation of the informal economy. This is reported in Model MII. The standard errors associated with the coefficient of the informal economy are not distinctly different from those obtained in Table 2A under MVII. The results remain robust, with the effect of informality on gasoline consumption inefficiency showing

[4] The energy saving potential is computed following the procedure of Zhang (2017) and Evans et al. (2013): $gsp_{it} = \left(\frac{1}{eff\_mean} + \frac{1}{eff_{it}} - 2\right)$ where gsp denotes the maximum gasoline saving potential due to efficiency improvement alone and eff denote efficiency in gasoline consumption.

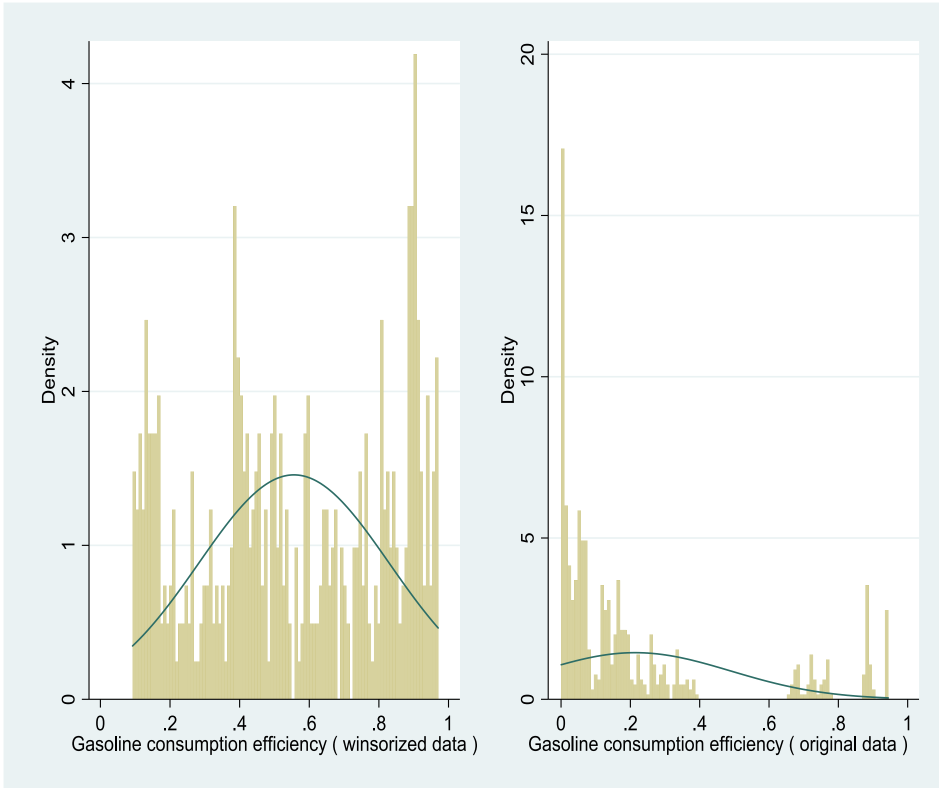

Fig. 4. Kernel density plot of gasoline consumption efficiency.
(Source: Author's construction)

an inverted U-shape, although statistically weak.

The relationship between inefficiency and informality can be complicated by the fact that the inefficiency term and informality do not exhibit rapid changes in the short term (Adom et al., 2023). As discussed elsewhere in this paper, rigidity in inefficiency could stem from the difficulty in switching to new technology in the short term, exposure to different market imperfections, and the use of different discounting rates by various energy users. Moreover, if the level of development impacts both energy efficiency (Adom, 2024) and the size of the informal economy, as found in this study, and development itself exhibits rigidity in the short term, then both energy efficiency and informality should not exhibit significant temporal variation in the short term. Failing to address this concern could impair the causal link between inefficiency and informality and affect the size of the long-run coefficient. Additionally, the instrumental variables used do not exhibit strong temporal variation.

To address this concern, I performed a five-year average regression (see Models MIII - MIV). This approach helps address possible measurement error problems and business cycle and heterogeneity effects, improving the estimation of the long-run coefficient. Additionally, a five-year average regression helps address the issue of partial adjustment of gasoline demand and inefficiency to price changes, which may not be fully captured by a static model.

The effect of the informal economy still exerts an inverted U-shaped effect on gasoline consumption inefficiency. This implies that the postulated relationship, rather than the coefficient size, is not driven by possible rigidity in inefficiency and the informal economy (influenced by the level of development), the choice of instrumental variables, measurement error in informality, business cycle and heterogeneity effects, or the partial adjustment nature of gasoline prices. However, the coefficients increase compared to those obtained in Model MVII in Table 2A, which was expected.

Next, measuring the size of the informal economy is fraught with measurement errors and is most likely underestimated. Here, I used the cost of starting a business as a proxy for the size of the informal economy. As shown in this study, a statistically significant positive association exists between the size of the informal economy and the cost of starting a business. Model MV contains the results, demonstrating that the cost of starting a business has a positive effect on gasoline consumption inefficiency.

Additionally, the effect of gasoline price in the inefficiency equation is statistically significant only when the effect is negative, confirming the earlier claim that the pricing tool can effectively promote gasoline consumption efficiency. Across the models from MI to MV, the estimated mean and median gasoline consumption efficiencies range between 47 % and 58 % and 47 % and 64 %, respectively. These values are not qualitatively different from those obtained in Model MVII in Table 2A, at least for the winsorized data.

Finally, I address concerns about omitting unobserved and observed heterogeneity and the problem of artificiality introduced by winsorizing the data, which can bias the estimate of gasoline consumption efficiency. Using the original data, the population effect size is expected to match the sample effect size well, providing a good counterfactual case to detect bias introduced through data winsorizing. I adopted the Kutlu et al. (2019) model, which allows for the control of unobserved and observed country heterogeneity (i.e., fixed effects and time-varying factors) and endogeneity simultaneously.[5] The F-statistics for heterogeneous effects of panel unit level is 141.79 with a p-value of 0.000, suggesting panel heterogeneity effects. In this application, I controlled for unobserved and observed country fixed effect heterogeneity. Model VI shows the results.

Regarding the effect of the size of the informal sector, the relationship is an inverted U-shape, indicating a saturation point at a 38 % share

[5] I estimate this model in Matlab. See the attached codes.

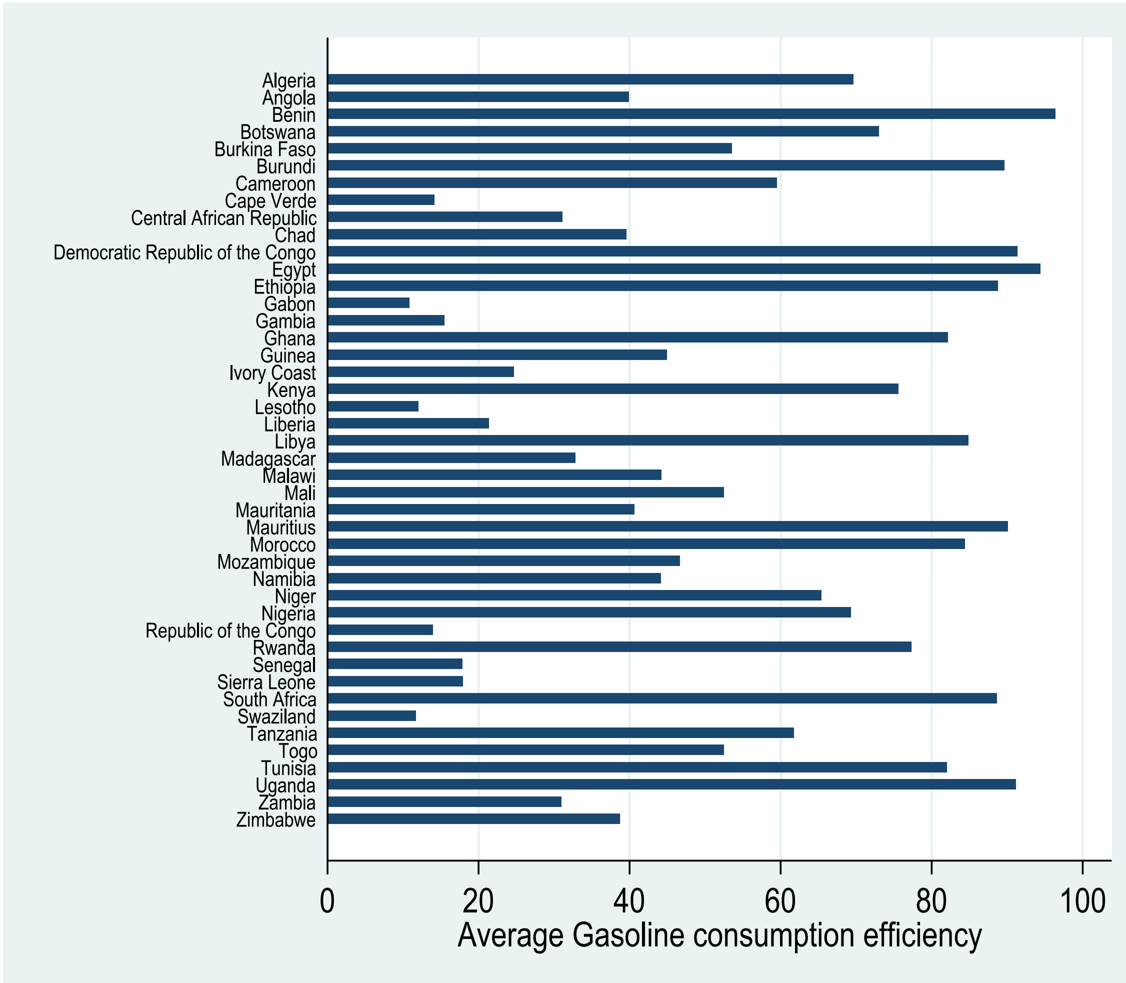

Fig. 5. Plot of average gasoline consumption efficiency by country.
(Source: Author's construction)

of informal activity in national GDP, compared to 34.5 % and 42 % shares of informal activity for the baseline regression and the regression with winsorized data, respectively. In terms of computed gasoline consumption efficiency, controlling for heterogeneity leads to an efficiency value of approximately 42 %, compared to 21.4 % for the baseline model and 55.6 % for the regression with winsorized data. The message here is clear: failing to account for outliers and country heterogeneity introduces a downward bias in the efficiency estimate. In the case of winsorizing, some of this bias might be driven by the artificiality introduced in the data. Given that the original data was used in the latter case, it might be safe to say that the extent of bias introduced in the efficiency estimate due to winsorizing could be the difference between 55.6 % and 42 %. Regarding the postulated relationship between gasoline inefficiency and the size of the informal economy, the result remains robust, at least in terms of the sign of the effect.

Additionally, I explored alternate estimators such as the Greene (2005) true fixed effect model, the Chen et al. (2014) consistent true fixed effect model, and the Kumbhakar et al. (2014) fixed effect version using the original data. All three models address unobserved heterogeneity, with Kumbhakar et al. (2014) distinguishing between persistent and transient efficiency. In these models, I applied a pseudo regression where the predicted value from an IV regression of the endogenous variable on the other independent and instrument variables replaced the endogenous variable, ensuring robustness to endogeneity.

The efficiency values for the transient versions vary significantly from the baseline estimate, while the persistent version of the efficiency is similar to the baseline. The effect of the informal sector shows an inverted U-shaped relationship (see Stata codes for these results). In the next section, I explore the regulatory effect channel for the size of the informal sector.

5.3. Informality, regulation, and governance nexus

This section identifies regulation and good governance as transmission mechanisms through which the effect of informality on energy efficiency could materialize. It is hypothesized that strengthening regulation and good governance reduces the size of the informal economy. To confirm this, I first performed a simple correlation analysis between informality and both aggregate and specific indicators of regulation and good governance. The results consistently show a negative association (see Stata codes). However, since correlation does not imply causation, I performed a simple regression of informality on the aggregate index of regulation and good governance.

Table 3A presents the results for both the full sample and the five-year average period regression. The latter approach addresses concerns about business cycle effects that could bias our coefficient estimates. For the full sample, year effects were accounted for to handle aggregate time trends and common time-varying shocks in the data, which might not indicate causation. In all scenarios, the regulation and good governance index has a negative and statistically significant effect on the size of informality. This confirms the earlier claim that stricter regulation and good governance can reduce the size of the informal

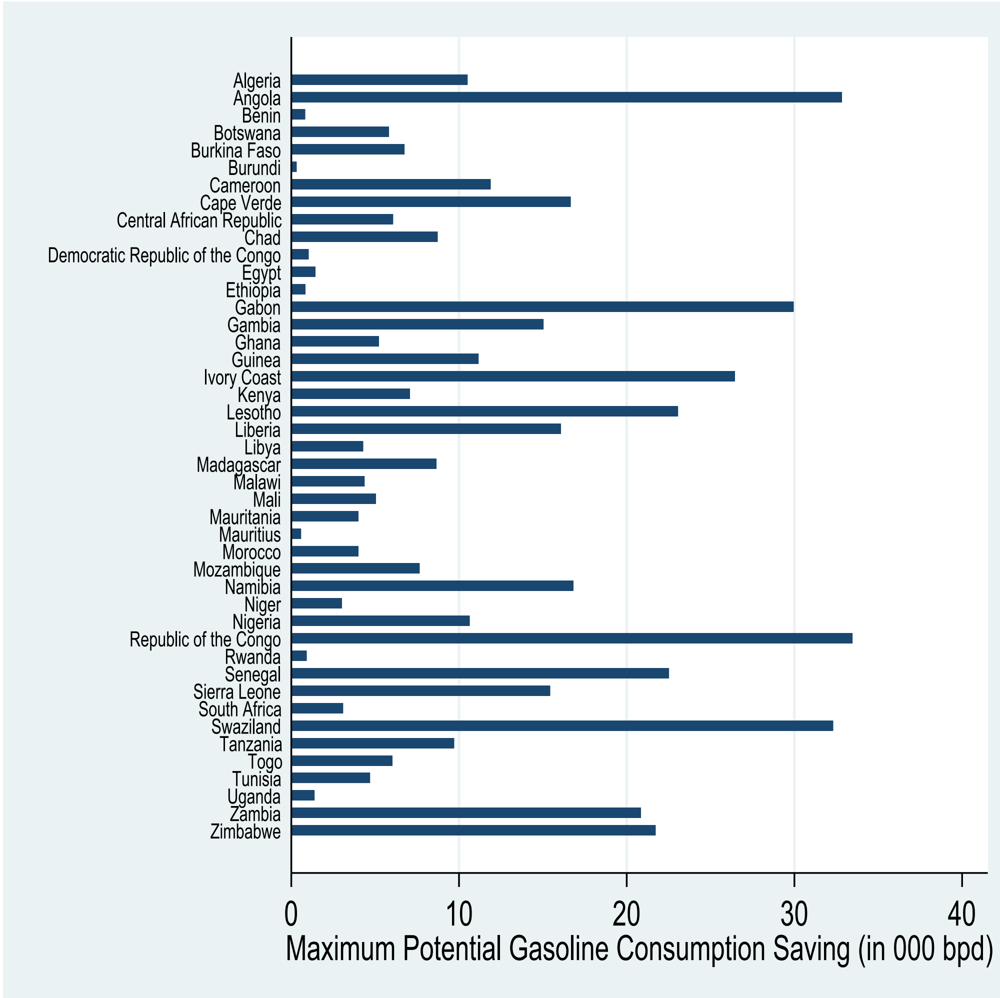

Fig. 6. Plot of average maximum potential gasoline energy saving.
(Source: Author's construction)

economy. Barra and Papaccio (2024) found that improving regulatory quality reduces the size of informal economy.

I do not expect all economies to experience rapid changes in the size of the informal sector, as it is connected to the level of development. Therefore, as a further robustness check to improve causal inferences, I considered long-term changes over five-year windows. This transformation reduces the amount of information, which could affect the degrees of freedom. Nevertheless, it helps strengthen the claim of causation because it allows for the observation of whether changes in the size of the informal sector over time are associated with changes in regulation and good governance. Additionally, this approach addresses concerns about business cycle effects.

Table 3B shows the results. In column DLSE_95/99, I examine whether changes in informality observed between 1995 and 1999 are associated with changes in regulation and good governance indicators. The same approach applies to the other columns. Consistently, the results show that regardless of the time window considered, increasing regulation and good governance has a negative and statistically significant effect on informality. This indicates that the evolution of the size of the informal economy in Africa is significantly associated with regulation and good governance.

Earlier, I posited that during the pre-saturation stage, behaviors in the informal sector might reflect an absence or inadequacy of government regulation. This suggests that the effectiveness of government regulation in achieving environmental sustainability and enhancing energy efficiency is stronger in the pre-saturation stage. Beyond this optimal level, the size of the informal economy becomes less sensitive to government regulation due to its large size, which complicates enforcement, compliance, and monitoring of regulations.

To test this hypothesis, I expect that if government regulation is below the optimal level, further strengthening of regulation should be associated with a reduction in the size of the informal sector. However, if regulation exceeds this optimal level, additional regulation would likely be ineffective in reducing the size of the informal economy in an economy with a very large informal sector. Empirically, this would manifest as a negative and statistically significant level effect of government regulation on the size of the informal sector, while the effect of the squared term of government regulation should be positive but statistically insignificant. Table 3C presents the regression results, confirming this hypothesis. This corroborates the findings of Fredstrom et al. (2021), who observed that informality diminishes entrepreneurial ability and that, when the informal sector becomes excessively large, government efforts to enhance governance quality may become counterproductive. Thus, based on Tables 2A and 3C, TPCH is confirmed in this study.

5.4. Gasoline consumption efficiency, regulation, and good governance

Having established the causal impact of regulation and good governance on informality, I now examine how regulation and good

Table 2C
Gasoline consumption inefficiency and informality – robustness checks.

Inefficiency equation Dep var.: $ln\left(\delta^2_{-u}\right)$	MI	MII	MIII	MIV	MV	MVI
constant	−1.741	−3.755	−38.64	−17.21	−2.028***	−16.467***
	(11.75)	(11.18)	(28.99)	(23.29)	(0.721)	(0.1279)
lnse	0.388	1.197	20.28	8.147		8.429***
	(6.607)	(6.256)	(16.04)	(12.74)		(0.4147)
lnse^2	−0.0725	−0.187	−2.918	−1.268		−1.158***
	(0.927)	(0.877)	(2.238)	(1.788)		(0.1020)
trend	0.0497	0.106	0.182***	0.183***	0.225***	0.284**
	(0.0999)	(0.0996)	(0.0322)	(0.0363)	(0.0787)	(0.1436)
trend^2	0.00122	−0.000593			−0.00479**	−0.016***
	(0.00310)	(0.00309)			(0.00234)	(0.0057)
lngp	0.158	0.152	−2.353***	−2.086***	−0.327***	0.579
	(0.177)	(0.170)	(0.447)	(0.490)	(0.0894)	(0.4980)
lncsb	0.245***	0.269***	0.304**	0.413***	0.201***	−0.028
	(0.0567)	(0.0570)	(0.131)	(0.145)	(0.0628)	(0.0452)
lnbcp	−0.165*	−0.127	−0.0956		0.0560	−0.242**
	(0.0979)	(0.0980)	(0.242)		(0.168)	(0.0963)
Dep var.: $ln\left(\delta^2_{-w}\right)$						
constant	−3.729***	−3.739***	−4.173***	−3.997***	−3.517***	−3.984***
	(0.0780)	(0.0782)	(0.190)	(0.184)	(0.0758)	(0.0770)
/eta1_lnrgdp	1.901***	2.237***	1.732***	1.153*	2.294***	3.210***
	(0.538)	(0.605)	(0.624)	(0.610)	(0.813)	(1.033)
/eta1_lnrgdp^2						1.587*
						(0.8423)
/e1	−0.107***	−0.103***	−0.115***	0.115***	−0.115***	
	(0.00392)	(0.00378)	(0.00822)	(0.00807)	(0.00403)	
Mu						
constant						0.0005***
						(0.0001)
Eta_end test	12.48***	13.67***	7.71***	3.57*	7.97***	10.705***
Weak IV test	19.89***	51.30***	7.01**	10.68***	37.18***	13.801***
Exclusion test_P-val	0.1731	0.3519	0.3519	0.3519	0.3519	0.3519
Average TEE (%)	48.84	48.61	58.34	57.40	47.94	41.7
Median TEE (%)	46.93	47.20	64.43	60.76	52.45	38.6
Log-likelihood	384.92	399.22	71.28	67.63	356.47	1626.595
Observations	378	378	101	102	416	462
Countries	44	44	44	44	44	44

Standard errors are in parentheses. Symbols indicate significance at the 1 % (***), 5 % (**), and 10 % (*) levels. For all the regressions in Table 2C, I retained 75 % of the original data for the dependent variable and 95 % of the original data for the independent variables except for MVI which use the original data. MI uses the share of arable land as the instrumental variable. MII weights the regressors by the inverse of the standard deviation of informality. MIII and MIV perform 5-year average regressions. MV uses the cost of starting a business as a proxy for informality. MVI controls for unobserved heterogeneity. Models MII-MVI use the legal system origin and the share of arable land as instrumental variables.
Source: Author's construction.

Table 3A
Impact of Regulation and Good Governance on Informality.

	MI	MII	MIII	MIV
VARIABLES	POLS	FE	POLS_5G*	FE_5G*
govindex	−0.0548***	−0.0416***	−0.0532***	−0.0415**
	(0.00326)	(0.0147)	(0.00685)	(0.0192)
constant	3.705***	3.536***	3.632***	3.631***
	(0.0256)	(0.0166)	(0.0152)	(0.00146)
Observations	745	745	175	175
R-squared	0.329	0.481	0.247	0.062
Number of id		44		44
Year Effect	YES	YES	NO	NO

Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. OLS is pooled ordinarily least squares and FE is a fixed effect. 5G* implies a five-year generation regression.
Source: Author's construction.

governance mediate the effect of informality on gasoline energy efficiency. Table 4 presents the results of the impact of regulation and good governance indicators on gasoline energy inefficiency (see appendix C for full table results). In all models shown in this table, the share of arable land and the legal system origin of the country are used as instrumental variables for real income per capita.

Except for models MVII and MIX, which utilize the original, non-winsorized data, the remaining models employ winsorized data. Specifically, the winsorizing procedure retains 75 % of the original data for gasoline consumption and 95 % of the original data for the independent variables. The validity of the instrumental variables is confirmed by passing both the weak IV and external exclusion restriction tests (see bottom part of Table 4).

Models MI and MII reveal that the composite measure of regulation and the good governance indicator have a negative effect on gasoline consumption inefficiency. Similarly, model MVII, which accounts for unobserved heterogeneity, also shows a negative and statistically significant effect of the composite governance index on gasoline consumption inefficiency. This result suggests that regulation can be an effective channel for enhancing gasoline consumption efficiency within the informal sector. Given the strong correlation between fossil fuels and pollution, these findings imply that improving regulation could help reduce environmental pollution by boosting efficiency. This supports the conclusions of Swain et al. (2020) and Biswas et al. (2012), who found that better regulation leads to reduced environmental pollution.

The use of composite measures of regulation and good governance complicates the identification of which specific indicators drive this effect. Including all indicators simultaneously can be problematic due to their high correlation, with an initial examination revealing a correlation coefficient greater than 70 %. While this raises concerns about

Table 3B
Impact of Regulation and Good Governance on Informality – different time windows.

	Model I	Model II	Model III	Model IV	Model V
VARIABLES	DLSE_95/99	DLSE_99/03	DLSE_03/07	DLSE_07/11	DLSE_11/15
lnse_95	−0.00777***				
	(0.00294)				
lnse_99		−0.00156			
		(0.00141)			
lnse_03			−0.00296		
			(0.00304)		
lnse_07				−0.0221***	
				(0.00397)	
lnse_11					−0.00742***
					(0.00178)
govindex	−0.00964***	−0.00473***	−0.00814***	−0.00489***	−0.00739***
	(0.000738)	(0.000775)	(0.00112)	(0.00138)	(0.00134)
Constant	0.0169	−0.0214***	−0.0754***	0.0496***	−0.0107*
	(0.0107)	(0.00413)	(0.0109)	(0.0145)	(0.00585)
R-squared	0.113	0.021	0.042	0.029	0.037

Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. DLSE measures the difference in informality between the ending year and the initial year. For example, DLSE_95/99 measures the difference in informality between 1995 and 1999. lnse_95, lnse_99, lnse_2003, lnse_2007, and lnse_2011 indicate the initial size of informality in those years.
Source: Author's construction.

Table 3C
Test of potency of regulation in reducing size of the informal economy.

Dependent variable: size of informality				
Variables	MI	MII	MIII	MIV
govindex	−0.0360*	−0.0393***	−0.0360*	−0.0393***
	(0.0199)	(0.0133)	(0.0231)	(0.0125)
govindex^2	0.00574	0.00192	0.00574	0.00192
	(0.00470)	(0.00401)	(0.00531)	(0.00387)
constant	3.590***	3.527***	3.590***	3.527***
	(0.0230)	(0.0237)	(0.0230)	(0.0398)
Year effect	NO	YES	NO	YES
Observations	745	745	745	745
R-squared	0.078	0.483	0.078	0.483
Number of id	44	44	44	44

Standard errors are in parentheses. MI – MII uses robust standard errors while MIII– MIV uses bootstrap standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: Author's construction.

multicollinearity, it is not necessarily a definitive indicator. Non-experimental or observational data often cannot be perfectly orthogonal, suggesting some tolerance for multicollinearity in regression analysis. To address this, I used the variance inflation factor (VIF) to test for multicollinearity. The overall VIF is 3.84, below the common threshold of 10, and none of the individual VIFs exceed this threshold. Based on these results, I included all indicators of regulation and good governance in a single regression. Model MIII presents the results, highlighting that among the indicators, the rule of law, political stability and absence of violence/terrorism, and regulatory quality are the key factors that reduce inefficiency in gasoline consumption.

Relying solely on Model MIII can be problematic because the contribution of each indicator to the composite measure varies significantly. The component loading plot from the Principal Component Analysis (PCA) (see Stata code for plot) reveals that indicators such as voice and accountability, political stability, and absence of violence/terrorism contribute the least to the aggregate index. Consequently, in Model MIV, these low-contributing indicators were omitted. A similar approach was adopted in Models MV and MVI. In these models, the rule of law indicator, which has the highest component contribution, shows a negative and statistically significant effect on gasoline consumption inefficiency.

In the analogous models (MVIII and MIX), where heterogeneity was controlled for, the rule of law continues to significantly reduce gasoline consumption inefficiency, alongside indicators for controlling corruption and regulatory quality. These findings align with the observations of Akorli and Adom (2023), Porter (1991), Fredstrom et al. (2021), and Babie and Wawryk (2021). The rule of law supports the procedural principles necessary for the effective implementation of market-based policies (Babie and Wawryk, 2021) and thereby aids in achieving energy efficiency improvements. This suggests that a lack of procedural principles under the rule of law could undermine the integrity of market-based schemes designed to enhance energy efficiency.

High levels of corruption could create political risks that impede technological transfers through foreign direct investment and increase uncertainties around investment returns. Given the high initial costs of energy-efficient technologies, particularly for poorer households and nations, corruption could deter investment in these technologies. Fredriksson et al. (2004) noted that corruption and lobbying obstruct developing economies from implementing stringent energy efficiency policies. Furthermore, Gennaioli and Tavoni (2011) highlighted that corruption and rent-seeking behavior inhibit the adoption of clean energy.

Effective regulation can guide behavior and foster innovation in the informal sector. By establishing clear rules and standards, individuals can better respond to market and technological conditions. Porter (1991) emphasized that stricter regulation could compel firms to invest in innovative technologies. In many developing economies, the prevalence of counterfeit or low-quality energy-efficient products undermines consumer confidence. Enhancing the regulatory framework to ensure the supply of only genuine, high-quality energy-efficient equipment could boost confidence and investment in energy efficiency. For instance, a randomized controlled trial in Kibera, Nairobi, found that respondents were more inclined to purchase energy-efficient products, even at higher prices, if the government ensured the availability of only high-quality, non-counterfeit products (Figueroa, 2016). The foregoing suggests that strengthening regulation and good governance within the appropriate range can improve gasoline consumption efficiency in Africa by reducing informality.

Regarding other drivers of gasoline consumption inefficiency, the cost of starting a business and bank credit to the private sector show a positive and negative effect, respectively, which corroborates previous findings. However, the influence of price remains weak in this context.

For the estimated gasoline consumption efficiency values, the mean efficiency ranges from 48 % to 54 % in the model that addresses outliers, and from 61 % to 64 % in the model that controls for heterogeneity bias. Comparing these estimates to the baseline regression highlights that not accounting for unobserved/observed heterogeneity or outliers can

Table 4
Gasoline consumption inefficiency and governance indicators.

	MI	MII	MIII	MIV	MV	MVI	MVII	MVIII	MIX
Inefficiency equation									
Dep var.: $ln\left(\delta^2_{-u}\right)$									
constant	−0.00837	−2.087***	−0.718	−0.721	−1.498*	−1.273*	−3.056	−0.984	−1.739
	(0.0452)	(0.686)	(0.817)	(0.842)	(0.823)	(0.755)	(3.0837)	(1.1630)	(1.0889)
govindex	−0.00837	−0.128**	–	–			−0.198***		
	(0.0452)	(0.0580)					(0.0691)		
rl	–		−1.193	−2.694***	−1.527*	−1.454***		−1.852***	−1.597**
			(0.863)	(0.837)	(0.811)	(0.491)		(0.7099)	(0.6369)
ge	–		−0.131	0.236	0.395			0.810	0.881
			(0.612)	(0.658)	(0.589)			(0.7363)	(0.7872)
ccr	–		2.581***	2.324***	0.0451			−1.354*	−1.382**
			(0.552)	(0.605)	(0.988)			(0.7380)	(0.6713)
rq	–		0.461	−0.896				−0.633	−0.123
			(0.885)	(0.987)				(0.9390)	(0.9523)
va	–		−1.433***	–				0.300	
			(0.551)					(0.6234)	
psav	–		−1.148**	–				0.292	
			(0.465)					(0.3906)	
lngp	−0.00837	−0.0254	0.209	0.169	−0.0188	−0.00788	0.467	0.462	0.449
	(0.0452)	(0.166)	(0.175)	(0.189)	(0.183)	(0.169)	(0.645)	(0.3088)	(0.2932)
l.lngp		0.211	0.233	0.327	0.275*	0.266*			
		(0.148)	(0.153)	(0.163)	(0.152)	(0.145)			
trend	−0.00837	0.267***	0.268***	0.243***	0.265***	0.246***	0.408	0.363***	0.426***
	(0.0452)	(0.0714)	(0.0711)	(0.0850)	(0.0828)	(0.0717)	(0.6025)	(0.1404)	(0.1150)
trend^2	−0.00837	−0.0065***	−0.00565**	−0.00620**	−0.00626**	−0.00581***	−0.021	−0.020***	−0.022***
	(0.0452)	(0.00212)	(0.00255)	(0.00257)	(0.00251)	(0.00214)	(0.0237)	(0.0053)	(0.0043)
lncsb	0.183***	0.292***	0.268***	0.247***	0.287***	0.286***	0.099	−0.032	0.006
	(0.0689)	(0.0677)	(0.0711)	(0.0758)	(0.0725)	(0.0690)	(0.0920)	(0.0747)	(0.0772)
lnbcp	−0.242***	−0.317***	−0.410***	−0.569***	−0.423***	−0.345***	−0.334**	−0.0405***	−0.405***
	(0.0904)	(0.115)	(0.122)	(0.122)	(0.117)	(0.113)	(0.1432)	(0.1385)	(0.1438)
Dep var.: $ln\left(\delta^2_{-w}\right)$									
constant	−0.242***	−3.671***	−3.761***	−3.727***	−3.646***	−3.668***	−3.870***	−3.902***	−3.893***
	(0.0904)	(0.0798)	(0.0800)	(0.0808)	(0.0824)	(0.142)	(0.0788)	0.0699)	(0.0693)
/eta1_lnrgdp	1.854**	4.949***	3.538***	3.074***	3.228***	4.585***	−9.310***	0.784	−0.549
	(0.783)	(1.131)	(0.728)	(0.881)	(1.090)	(1.063)	(1.6221)	(1.1467)	(1.0714)
/eta1_lnrgdp^2							−7.844***	−1.350*	−1.928***
							(0.8267)	(0.7295)	(0.5869)
/e1	0.115***	0.107***	0.103***	0.106***	0.106***	−0.107***			
	(0.00401)	(0.00397)	(0.00384)	(0.00396)	(0.00396)	(0.0040)			
Mu									
constant							0.344	1.075	0.709*
							(1.0766)	(0.7463)	(0.4084)
Eta_end test	5.61**	19.14***	23.59***	12.17***	8.77***		–	11.403***	19.344***
Weak IV test_front eq	26.14***	72.23***	71.29***	55.32***	53.86***	67.86***	8.090**	12.588***	10.900***
Exclusion test_P_value	0.4618	0.4618	0.4618	0.4618	0.4618	0.4618	0.4618	0.4618	0.4618
Average TEE (%)	49.11	47.81	47.48	54.49	50.09	48.96	62.89	61.05	63.35
Median TEE (%)	52.67	45.06	42.42	54.63	49.89	48.39	65.25	63.984	66.36
Log-likelihood	357.27	361.86	395.23	384.04	365.18	362.69	1714.832	1756.44	1753.96
Observations	416	369	369	369	369	369	511	511	511
Countries	44	44	44	44	44	44	44	44	44

Standard errors are in parentheses. Symbols denote significance at the 1 % (***) level, 5 % (**) level, and 10 % (*) level. All models use the share of arable land and legal system origin as instrumental variables. Models MI-MVI are based on winsorized data, retaining 75 % of the original gasoline consumption data and 95 % of the original independent variables data. Models MI-MVI do not account for unobserved heterogeneity. Models MVII-MIX, which use the original data, account for unobserved heterogeneity.
Source: Author's construction.

introduce a downward bias in the efficiency estimate, reinforcing earlier observations.

6. Conclusion and policy implications

The literature extensively explores the role of technology and market mechanisms in influencing gasoline energy efficiency. However, there is a notable gap concerning the impact of informality, regulation, and good governance on gasoline consumption efficiency. Establishing a robust system of good governance is crucial for the successful implementation and achievement of energy efficiency outcomes. Evidence suggests that many failed energy efficiency policies neglected the importance of strong regulation and good governance.

In this study, I propose TPCH test, which estimates gasoline consumption efficiency and investigate how informality, regulation, and good governance affect gasoline consumption efficiency. I propose that there is a saturation point in the relationship between gasoline consumption inefficiency and the size of the informal economy, reflected by an inverted U-shaped curve. This implies that during the pre-saturation stage, energy-saving behaviors in the informal sector may be indicative of weak or absent government regulation. Conversely, in the post-saturation stage, such behaviors might be driven by self-motivation due to the impacts of climate change on informal sector activities. Therefore, government regulation might be more effective in promoting gasoline consumption efficiency during the pre-saturation stage than in the post-saturation stage, where enforcement, compliance, and monitoring become challenging.

Empirically, this translates into (1) an inverted U-shaped effect of the size of the informal sector on gasoline inefficiency and (2) a U-shaped effect of government regulation on the size of the informal sector, with only the level effect being statistically significant. I tested these hypotheses using a stochastic frontier technique that accounts for endogeneity, outlier effects, and heterogeneity bias. The study uses the legal system origin of the country and the share of arable land as instrumental variables for real gross domestic product. The data covers 44 African countries from 1995 to 2019.

In the first part of the analysis, I consistently confirm that the relationship between the size of the informal sector and gasoline consumption inefficiency is inverted U-shaped, indicating the presence of a saturation point. This led to an examination of the second claim, which posits that government regulation has a U-shaped effect on the size of the informal sector. However, only the level effect coefficient of government regulation is statistically significant.

When integrating these findings, the results affirm that government regulation is more effective in enhancing gasoline consumption efficiency during the pre-saturation stage compared to the post-saturation stage. This conclusion is supported by the introduction of governance indicators into the inefficiency model. Regulation and good governance were found to have a statistically significant and negative impact on gasoline consumption inefficiency, with these results remaining robust after controlling for outliers and heterogeneity bias.

Among the governance indicators, the rule of law, control of corruption, and regulatory quality emerged as the key factors that promote gasoline consumption efficiency. Additionally, the study highlights the important role of bank credit to the private sector and the cost of starting a business—rather than price—in promoting gasoline consumption efficiency.

Finally, the estimated gasoline consumption efficiency varies significantly across countries. Generally, better-performing economies are also among the continent's top economic performers, suggesting a strong link between economic development and gasoline consumption efficiency. Conversely, the countries with the least efficient performance exhibit greater potential for savings in gasoline energy consumption. Moreover, the analysis reveals that failing to control for outliers and heterogeneity bias in the data can introduce a downward bias in the estimate of gasoline consumption efficiency.

The findings of this study carry significant implications for designing energy efficiency policies. They suggest a differentiated approach. In economies at the pre-saturation stage, the effectiveness of gasoline energy efficiency policies relies heavily on regulating informal markets. For these policies to succeed, it is crucial to establish robust regulatory frameworks and good governance systems. This includes strengthening institutions to enhance regulatory quality, control corruption, and uphold the rule of law.

For economies in the post-saturation stage, policy design should focus on rewarding self-initiated pro-environmental actions. For instance, subsidizing the cost of energy-efficient appliances for the informal sector could be a key incentive for improving energy efficiency investments. Additionally, offering soft loans specifically for purchasing energy-efficient appliances could alleviate income constraints and stimulate investment in such technologies. Providing tax rebates to informal firms investing in energy-efficient technologies could also be beneficial.

While this study sets a valuable precedent for future research, there are noticeable limitations. Although this study addressed some impediments faced by the informal sector, other factors may still need to be considered. For example, due to data constraints, the study does not account for the effects of local infrastructure and environmental policies, which could be crucial factors. Since these variables may correlate with long-term economic development, the study's treatment of rigidity in both gasoline consumption efficiency and the size of the informal economy offers only partial insights. Future research should explicitly include these variables to provide a more comprehensive analysis.

Also, the study lacks detailed micro-level insights. Future research could benefit from a micro-level approach to better understand specific activities within the informal sector and their energy use behavior. Finally, future studies analyzing gasoline consumption data should be cautious of outliers and heterogeneity bias, as these can introduce downward bias in coefficient estimates.

CRediT authorship contribution statement

Philip Kofi Adom: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The author declares that this manuscript has not been published or presented elsewhere, either in part or in its entirety, and is not under consideration by any other journal. I have read and understood the journal's policies, and neither the manuscript nor the study violates any of these policies. There are no conflicts of interest to declare.

Data availability

The author declares that the data and software codes used to generate the results in this study have been included in the article for replication purposes.

Acknowledgment

This study benefited from the insightful comments provided by two anonymous reviewers and the journal editor, Richard Tol. The author is also grateful to the Environment for Development (EfD Global) and the Sustainable Energy Transformation Initiative (SETI) for their support. Additionally, this work was presented at a capacity training workshop at the Ghana Institute of Management and Public Administration (GIMPA), and the author appreciates the valuable contributions of the participants.

Appendix A

Table 2A
Gasoline consumption inefficiency and informality.

Frontier det. Dep var. lngc	MI Baseline	MII	MIII	MIV	MV	MVI	MVII
Lnrgdp	−2.443***	−1.401**	−0.594	−0.544	−0.767	−1.389**	−1.058**
	(0.690)	(0.598)	(0.590)	(0.537)	(0.577)	(0.557)	(0.471)
Lntemp	−5.806***	−3.835*	−1.977	−2.270	−3.080	−7.374***	−1.654
	(1.739)	(1.988)	(1.856)	(1.783)	(1.875)	(1.890)	(1.183)
Lngp	−2.495***	−2.139***	−1.754***	−1.849***	−2.086***	−2.233***	−1.662***
	(0.543)	(0.535)	(0.516)	(0.483)	(0.533)	(0.550)	(0.392)
Lnsva	3.314***	2.270***	1.596**	1.672***	1.919***	2.348***	2.096**
	(0.775)	(0.688)	(0.650)	(0.618)	(0.684)	(0.731)	(0.821)
Lnpd	1.049***	−0.0553	−0.111	−0.139	−0.158	−0.610***	−0.287***
	(0.373)	(0.210)	(0.201)	(0.181)	(0.168)	(0.136)	(0.0921)
Trend	0.147***	0.0839*	0.0189	0.0211	0.0373	0.199***	0.230***
	(0.0535)	(0.0472)	(0.0456)	(0.0421)	(0.0435)	(0.0478)	(0.0656)
*lnrgdp*lntemp*	−5.613***	−4.803***	−4.237***	−4.708***	−5.675***	−4.275***	−0.695
	(1.696)	(1.577)	(1.507)	(1.467)	(1.748)	(1.656)	(1.346)
*lnrgdp*lngp*	−1.754***	−1.540***	−1.141**	−1.176***	−1.379***	−2.197***	−0.381
	(0.472)	(0.496)	(0.478)	(0.450)	(0.503)	(0.589)	(0.251)
*lnrgdp*lnsva*	1.691***	1.119**	0.637	0.590	0.746*	1.082**	0.392
	(0.551)	(0.480)	(0.437)	(0.418)	(0.453)	(0.523)	(0.398)
*lnrgdp*lnpd*	−1.233***	−1.235***	−0.911***	−0.850***	−0.838***	−0.339**	−0.266***
	(0.214)	(0.182)	(0.164)	(0.151)	(0.144)	(0.149)	(0.0828)
*lntemp*lngp*	−8.847***	−8.273***	−6.676***	−6.220***	−6.364***	−5.388***	−2.023*
	(1.788)	(1.798)	(1.709)	(1.606)	(1.738)	(1.896)	(1.099)
*lntemp*lnsva*	−5.145	−4.222	−2.153	−2.441	−3.741	−4.987	−1.383
	(3.839)	(3.656)	(3.422)	(3.286)	(3.527)	(3.427)	(2.892)
*lntemp*lnpd*	4.235***	3.791***	4.934***	5.069***	5.353***	−3.608***	4.206***
	(1.352)	(1.306)	(1.261)	(1.167)	(1.064)	(0.947)	(0.991)
*lngp*lnsva*	−0.469	1.335*	0.792	0.749	0.871	1.319*	0.426
	(0.539)	(0.689)	(0.636)	(0.607)	(0.659)	(0.781)	(0.596)
*lngp*lnpd*	−0.336**	−0.378**	−0.260	−0.254*	−0.287*	−1.158***	−0.0315
	(0.157)	(0.156)	(0.159)	(0.149)	(0.160)	(0.247)	(0.0886)
*lnsva*lnpd*	2.356***	1.512***	0.888**	0.820**	0.931**	1.912***	1.069***
	(0.565)	(0.438)	(0.400)	(0.373)	(0.407)	(0.524)	(0.393)
*lnrgdp*trend*	0.150***	0.109***	0.0728**	0.0693***	0.0779***	0.124***	0.0675***
	(0.0324)	(0.0295)	(0.0287)	(0.0265)	(0.0291)	(0.0291)	(0.0239)
*lntemp*trend*	0.258***	0.249***	0.156*	0.165**	0.205**	0.427***	0.114**
	(0.0757)	(0.0841)	(0.0808)	(0.0764)	(0.0858)	(0.0849)	(0.0562)
*lngp*trend*	0.111***	0.0831***	0.0604**	0.0695***	0.0872***	0.0852***	0.0945***
	(0.0285)	(0.0276)	(0.0266)	(0.0248)	(0.0271)	(0.0246)	(0.0208)
*lnsva*trend*	−0.181***	−0.146***	−0.117***	−0.119***	−0.128***	−0.117***	−0.131***
	(0.0395)	(0.0348)	(0.0328)	(0.0313)	(0.0350)	(0.0356)	(0.0444)
*lnpd*trend*	0.0139*	0.0251***	0.0170***	0.0156***	0.0162***	0.0268***	0.00229
	(0.00747)	(0.00615)	(0.00583)	(0.00548)	(0.00569)	(0.00701)	(0.00520)
lnrgdp^2	−1.504***	−1.102***	−0.694***	−0.607***	−0.671***	−1.059***	−0.840***
	(0.295)	(0.249)	(0.251)	(0.227)	(0.235)	(0.252)	(0.162)
lntemp^2	−1.271	−0.000142	1.978	−5.690	−12.58*	−11.19	−1.078
	(5.232)	(8.555)	(7.914)	(7.446)	(7.202)	(8.373)	(5.231)
lngp^2	−1.169***	−1.622***	−1.249***	−1.260***	−1.444***	−1.009***	−0.431**
	(0.244)	(0.347)	(0.324)	(0.305)	(0.337)	(0.314)	(0.177)
lnsva^2	−0.193	−3.589***	−2.907***	−2.465***	−2.318**	−1.978*	−0.400
	(0.713)	(1.089)	(0.959)	(0.919)	(0.970)	(1.147)	(0.945)
lnpd^2	0.423**	−0.399***	−0.401***	−0.357***	−0.281***	−0.247***	−0.244***
	(0.176)	(0.128)	(0.120)	(0.108)	(0.102)	(0.0851)	(0.0633)
trend^2	−0.000915	0.000955	0.00262**	0.00231**	0.00174	−0.00209*	−0.00464**
	(0.00146)	(0.00125)	(0.00121)	(0.00112)	(0.00111)	(0.00120)	(0.00181)
Constant	2.937***	3.236***	3.385***	3.159***	2.835***	0.998**	0.275
	(0.561)	(0.392)	(0.369)	(0.354)	(0.380)	(0.434)	(0.575)
Inefficiency equation Dep var.: $ln\left(\delta^2_{-u}\right)$							
Constant	−14.42***	−12.86**	−7.616	−9.690	−14.45*	−18.62**	−3.126
	(4.089)	(6.547)	(7.932)	(7.750)	(8.500)	(8.503)	(10.22)
Lnse	9.468***	8.458**	5.568	6.601	9.135*	10.07**	0.774
	(2.259)	(3.591)	(4.338)	(4.244)	(4.673)	(4.586)	(5.597)
lnse^2	−1.337***	−1.204**	−0.818	−0.955*	−1.296**	−1.346**	−0.168
	(0.311)	(0.490)	(0.590)	(0.579)	(0.640)	(0.619)	(0.778)
Trend	−0.0129	−0.0345**	−0.0594***	−0.0681***	−0.0766***	0.0781***	0.293***
	(0.0123)	(0.0141)	(0.0176)	(0.0182)	(0.0179)	(0.0257)	(0.0828)
trend^2	0.00107**	0.00177***	0.00271***	0.00294***	0.00313***	−0.00192**	−0.00734***
	(0.000481)	(0.000553)	(0.000695)	(0.000721)	(0.000718)	(0.000943)	(0.00254)

(*continued on next page*)

Table 2A (*continued*)

Frontier det. Dep var. lngc	MI Baseline	MII	MIII	MIV	MV	MVI	MVII
Lngp	−0.182***	−0.231***	−0.333***	−0.348***	−0.327***	0.0578	0.0175
	(0.0560)	(0.0711)	(0.0792)	(0.0840)	(0.0894)	(0.125)	(0.159)
Lncsb							0.195***
							(0.0622)
Lnbcp							−0.325***
							(0.0888)
Dep var.:$ln\left(\delta^2_{-w}\right)$							
Constant	−2.674***	−2.749***	−2.773***	−2.816***	−2.852***	−3.279***	−3.752***
	(0.0573)	(0.0561)	(0.0561)	(0.0561)	(0.0562)	(0.0634)	(0.0700)
/eta1_lnrgdp	3.255***	2.269***	1.657***	1.609***	1.804***	2.366***	1.782***
	(0.646)	(0.541)	(0.529)	(0.491)	(0.547)	(0.526)	(0.499)
/e1	−0.170***	0.160***	−0.160***	−0.160***	−0.160***	−0.154***	0.108***
	(0.00458)	(0.00431)	(0.00431)	(0.00431)	(0.00431)	(0.00468)	(0.00355)
Eta_end test	25.39***	17.57***	9.81***	10.75***	10.88***	20.29***	12.76***
Weak IV test	99.45***	93.62***	92.22***	92.02***	89.82***	75.42***	48.69***
Test of instrument exclusion_P-value	0.4455	0.9334	0.8009	0.8248	0.6246	0.3519	0.3519
Average TEE (%)	21.41	26.52	27.92	30.32	32.35	42.37	55.62
Median TEE (%)	11.42	14.72	17.20	23.37	27.11	37.87	53.60
Log-likelihood	46.66	123.02	136.30	154.37	169.26	266.55	479.37
Observations	689	689	689	689	689	542	462
Countries	44	44	44	44	44	44	44

Standard errors are in parentheses. Symbols indicate significance at the 1 % (***), 5 % (**), and 10 % (*) levels. MI is the baseline model, where the original data are used in the regression. For MII–MVII, the variables (both dependent and independent) are winsorized. For the independent variables, a winsorized window of 2.5 and 97.5 percentiles was maintained, meaning 95 % of the original data for these variables remained unchanged for MII–MVII. For the dependent variable (gasoline consumption), 95 %, 90 %, 85 %, 80 %, and 75 % of the original data were retained in MII, MIII, MIV, MV, and MVI, respectively. Model VII is an extension of Model VI, where the cost of starting a business and bank credit to the private sector were controlled for in the inefficiency equation. In all models, the legal system origin of the country and the share of arable land were used as instruments for real GDP per capita.

Appendix B

Table 2C
Gasoline consumption inefficiency and informality – robustness checks.

Frontier det. Dep var. lngc	MI	MII	MIII	MIV	MV	MVI
lnrgdp	−1.134*	−1.405**	−1.666***	−1.145*	−1.586*	−2.069**
	(0.580)	(0.617)	(0.594)	(0.593)	(0.821)	(0.9514)
lntemp	−1.389	−1.983	−4.232**	−3.098**	−1.132	−0.873
	(1.486)	(1.628)	(1.662)	(1.537)	(1.818)	(1.7182)
lngp	−2.045***	−2.209***	−1.866**	−2.051**	−2.193***	0.439
	(0.533)	(0.563)	(0.876)	(0.844)	(0.616)	(0.4740)
lnsva	1.982**	2.182**	6.305***	5.306***	3.181***	−0.232
	(0.908)	(0.944)	(1.746)	(1.745)	(1.214)	(0.4770)
lnpd	−0.245**	−0.222*	−0.757***	−0.674***	−0.310**	−0.187
	(0.122)	(0.127)	(0.199)	(0.196)	(0.140)	(0.9822)
lrend	0.172**	0.217**	0.145	0.178**	0.287***	0.085
	(0.0793)	(0.0884)	(0.0962)	(0.0875)	(0.0878)	(0.0852)
*lnrgdp*lntemp*	−1.236	−1.056	−6.171***	−4.877***	1.711	0.470
	(1.911)	(1.846)	(1.948)	(1.854)	(1.855)	(1.8462)
*lnrgdp*lngp*	−0.630**	−0.699**	−1.114	−0.808	−0.508	0.121
	(0.319)	(0.331)	(0.699)	(0.682)	(0.317)	(0.2011)
*lnrgdp*lnsva*	0.237	0.453	5.071***	4.483***	0.641	−0.980**
	(0.495)	(0.538)	(1.372)	(1.560)	(0.649)	(0.4325)
*lnrgdp*lnpd*	−0.207	−0.227*	−0.752***	−0.682***	−0.430**	−1.138***
	(0.127)	(0.132)	(0.181)	(0.200)	(0.189)	(0.3931)
*lntemp*lngp*	−1.812	−1.877	−7.938**	−5.158	−2.357*	−1.016
	(1.261)	(1.366)	(3.609)	(3.253)	(1.376)	(0.7937)
*lntemp*lnsva*	0.0441	−2.727	−10.66	−13.10*	−6.039	1.595
	(3.199)	(3.612)	(8.273)	(7.862)	(4.352)	(2.5973)
*lntemp*lnpd*	3.173***	3.404***	0.846	1.471	3.718***	−0.188
	(1.068)	(1.124)	(1.493)	(1.562)	(1.233)	(1.4657)
*lngp*lnsva*	−0.235	0.0383	4.065*	3.154*	0.670	−0.969
	(0.747)	(0.810)	(2.114)	(1.917)	(0.919)	(0.6855)
*lngp*lnpd*	−0.396**	−0.457**	−0.962**	−0.821**	−0.341*	0.092
	(0.174)	(0.188)	(0.404)	(0.385)	(0.178)	(0.0899)
*lnsva*lnpd*	0.827	1.063*	2.752**	2.042	2.009**	−0.047
	(0.510)	(0.560)	(1.284)	(1.316)	(0.820)	(0.2723)
*lnrgdp*trend*	0.0871***	0.0961***	0.0916***	0.0654*	0.101***	0.054***
	(0.0282)	(0.0295)	(0.0347)	(0.0341)	(0.0374)	(0.0199)
*lntemp*trend*	0.127**	0.136**	0.330***	0.254***	0.0948	0.068

(*continued on next page*)

Table 2C (*continued*)

Frontier det. Dep var. lngc	MI	MII	MIII	MIV	MV	MVI
	(0.0626)	(0.0653)	(0.106)	(0.0983)	(0.0711)	(0.0612)
*lngp*trend*	0.111***	0.120***	0.0162	0.0390	0.110***	−0.015
	(0.0263)	(0.0282)	(0.0560)	(0.0582)	(0.0304)	(0.0176)
*lnsva*trend*	−0.129***	−0.132***	−0.330***	−0.262***	−0.161***	−0.017
	(0.0477)	(0.0477)	(0.0913)	(0.0839)	(0.0548)	(0.0277)
*lnpd*trend*	0.0134**	0.0120*	0.0408***	0.0356***	0.0146**	0.017
	(0.00665)	(0.00670)	(0.0134)	(0.0137)	(0.00731)	(0.0125)
lnrgdp^2	−0.768***	−0.886***	−1.185***	−0.848***	−0.992***	1.325*
	(0.241)	(0.262)	(0.257)	(0.242)	(0.334)	(0.7820)
lntemp^2	6.902	4.605	−10.02	−8.302	3.581	−0.419
	(6.581)	(7.283)	(8.184)	(7.794)	(8.221)	(5.2507)
lngp^2	−0.637**	−0.717***	−0.0630	−0.487	−0.937***	−0.181
	(0.254)	(0.274)	(0.788)	(0.785)	(0.301)	(0.1182)
lnsva^2	−1.371	−1.585	−8.738***	−6.917**	−0.681	−1.241
	(1.119)	(1.193)	(3.307)	(3.161)	(1.248)	(1.0864)
lnpd^2	−0.204**	−0.165	−0.295***	−0.360***	−0.192	−0.267
	(0.101)	(0.102)	(0.102)	(0.0940)	(0.124)	(0.3054)
trend^2	−0.00140	−0.00269	0.00135	−0.000445	−0.00482**	−0.003
	(0.00220)	(0.00244)	(0.00313)	(0.00284)	(0.00220)	(0.0027)
constant	0.646	0.233	0.278	0.162	−0.247	
	(0.689)	(0.764)	(0.754)	(0.715)	(0.820)	
Inefficiency equation						
Dep var.:$ln\left(\delta^{2}_{-u}\right)$						
constant	−1.741	−3.755	−38.64	−17.21	−2.028***	−16.467***
	(11.75)	(11.18)	(28.99)	(23.29)	(0.721)	(0.1279)
lnse	0.388	1.197	20.28	8.147		8.429***
	(6.607)	(6.256)	(16.04)	(12.74)		(0.4147)
lnse^2	−0.0725	−0.187	−2.918	−1.268		−1.158***
	(0.927)	(0.877)	(2.238)	(1.788)		(0.1020)
trend	0.0497	0.106	0.182***	0.183***	0.225***	0.284**
	(0.0999)	(0.0996)	(0.0322)	(0.0363)	(0.0787)	(0.1436)
trend^2	0.00122	−0.000593			−0.00479**	−0.016***
	(0.00310)	(0.00309)			(0.00234)	(0.0057)
lngp	0.158	0.152	−2.353***	−2.086***	−0.327***	0.579
	(0.177)	(0.170)	(0.447)	(0.490)	(0.0894)	(0.4980)
lncsb	0.245***	0.269***	0.304**	0.413***	0.201***	−0.028
	(0.0567)	(0.0570)	(0.131)	(0.145)	(0.0628)	(0.0452)
lnbcp	−0.165*	−0.127	−0.0956		0.0560	−0.242**
	(0.0979)	(0.0980)	(0.242)		(0.168)	(0.0963)
Dep var.:$ln\left(\delta^{2}_{-w}\right)$						
constant	−3.729***	−3.739***	−4.173***	−3.997***	−3.517***	−3.984***
	(0.0780)	(0.0782)	(0.190)	(0.184)	(0.0758)	(0.0770)
/eta1_lnrgdp	1.901***	2.237***	1.732***	1.153*	2.294***	3.210***
	(0.538)	(0.605)	(0.624)	(0.610)	(0.813)	(1.033)
/eta1_lnrgdp^2						1.587*
						(0.8423)
/e1	−0.107***	−0.103***	−0.115***	0.115***	−0.115***	
	(0.00392)	(0.00378)	(0.00822)	(0.00807)	(0.00403)	
Mu						
Constant						0.0005***
						(0.0001)
Eta_end test	12.48***	13.67***	7.71***	3.57*	7.97***	10.705***
Weak IV test	19.89***	51.30***	7.01**	10.68***	37.18***	13.801***
Test of instrument exclusion_P-value	0.1731	0.3519	0.3519	0.3519	0.3519	0.3519
Average TEE (%)	48.84	48.61	58.34	57.40	47.94	41.7
Median TEE (%)	46.93	47.20	64.43	60.76	52.45	38.6
Log-likelihood	384.92	399.22	71.28	67.63	356.47	1626.595
Observations	378	378	101	102	416	462
Countries	44	44	44	44	44	44

Standard errors are in parentheses. Symbols indicate significance at the 1 % (***), 5 % (**), and 10 % (*) levels. For all the regressions in Table 2C, I retained 75 % of the original data for the dependent variable and 95 % of the original data for the independent variables except for MVI which use the original data. MI uses the share of arable land as the instrumental variable. MII weights the regressors by the inverse of the standard deviation of informality. MIII and MIV perform 5-year average regressions. MV uses the cost of starting a business as a proxy for informality. MVI controls for unobserved heterogeneity. Models MII-MVI use the legal system origin and the share of arable land as instrumental variables.

Appendix C

Table 4
Gasoline consumption inefficiency and governance indicators.

	MI	MII	MIII	MIV	MV	MVI	MVII	MVIII	MIX
Frontier det. lngc									
lnrgdp	−0.367	−3.642***	−2.557***	−2.178***	−2.032**	−3.265***	10.622***	0.666	1.928*
	(1.788)	(1.079)	(0.676)	(0.804)	(1.007)	(0.989)	(1.8058)	(1.0696)	(1.0164)
lntemp	−0.367	−7.273**	−6.648***	−3.704*	−4.035	−6.491**	1.317	−1.441	−0.845
	(1.788)	(2.894)	(2.180)	(2.045)	(2.676)	(2.664)	(3.4316)	(1.7215)	(1.7107)
lngp	−1.907***	−3.099***	−2.110***	−1.977***	−2.067***	−2.856***	−0.594	−0.037	−0.130
	(0.585)	(0.934)	(0.651)	(0.683)	(0.791)	(0.873)	(0.4349)	(0.2645)	(0.2448)
lnsva	2.540**	6.664***	4.647***	4.272***	4.075**	6.053***	0.993	0.395	0.478
	(1.147)	(1.972)	(1.287)	(1.433)	(1.685)	(1.815)	(0.8076)	(0.4348)	(0.4245)
lnpd	−0.298**	−0.381*	−0.321**	−0.370***	−0.385**	−0.403**	3.553**	−0.110	0.371
	(0.136)	(0.195)	(0.153)	(0.141)	(0.157)	(0.182)	(1.3944)	(0.8986)	(0.8785)
trend	0.249***	0.519***	0.391***	0.381***	0.386***	0.497***	−0.339***	0.016	−0.023
	(0.0837)	(0.130)	(0.0881)	(0.0925)	(0.113)	(0.123)	(0.0514)	(0.0622)	(0.6871)
*lnrgdp*lntemp*	2.111	−0.661	−3.153	1.106	1.533	0.158	10.108***	2.393	3.501*
	(1.876)	(2.375)	(2.210)	(1.792)	(2.186)	(2.204)	(3.8301)	(2.0530)	(2.0216)
*lnrgdp*lngp*	−0.381	−0.489	−0.250	−0.362	−0.212	−0.462	−0.116	0.090	0.088
	(0.295)	(0.412)	(0.298)	(0.294)	(0.317)	(0.388)	(0.2629)	(0.1180)	(0.1162)
*lnrgdp*lnsva*	0.355	2.043**	1.183	0.994	0.986	1.758*	3.198***	0.164	0.482
	(0.616)	(0.986)	(0.733)	(0.762)	(0.824)	(0.904)	(0.7266)	(0.4802)	(0.4483)
*lnrgdp*lnpd*	−0.370*	−0.768***	−0.628***	−0.441**	−0.422**	−0.670***	2.049**	−0.800*	−0.360
	(0.189)	(0.242)	(0.184)	(0.185)	(0.208)	(0.207)	(0.8879)	(0.4425)	(0.42910)
*lntemp*lngp*	−1.996	−0.937	−1.219	−1.879	−0.760	−1.170	−2.980**	−1.775**	−1.710**
	(1.255)	(2.421)	(1.778)	(1.657)	(1.743)	(2.282)	(1.4766)	(0.7197)	(0.6998)
*lntemp*lnsva*	−4.764	−2.367	−0.505	−0.989	−0.344	−2.854	1.349	3.409	2.487
	(4.057)	(6.103)	(4.634)	(4.536)	(4.735)	(5.822)	(5.3368)	(2.8598)	(2.7692)
*lntemp*lnpd*	3.666***	2.826*	1.940	3.676***	2.662**	2.863**	−1.283	−1.337	−1.065
	(1.209)	(1.463)	(1.286)	(1.177)	(1.318)	(1.391)	(2.8665)	(1.4239)	(1.3723)
*lngp*lnsva*	0.344	3.386**	2.113*	2.645**	2.447*	3.348**	−6.204***	−2.290***	−2.561***
	(0.848)	(1.651)	(1.153)	(1.167)	(1.323)	(1.567)	(0.7242)	(0.5915)	(0.5072)
*lngp*lnpd*	−0.287*	−0.867**	−0.426*	−0.484**	−0.527*	−0.805**	−0.375**	−0.018	−0.061
	(0.161)	(0.339)	(0.225)	(0.239)	(0.272)	(0.318)	(0.1642)	(0.0834)	(0.0815)
*lnsva*lnpd*	1.589**	3.499***	2.622***	2.207***	2.008**	3.057***	−0.430	−0.367	−0.292
	(0.794)	(1.097)	(0.758)	(0.836)	(0.967)	(0.994)	(0.5911)	(0.3055)	(0.2920)
*lnrgdp*trend*	0.0803**	0.196***	0.143***	0.124***	0.120***	0.178***	−0.197***	−0.012	−0.034*
	(0.0365)	(0.0498)	(0.0311)	(0.0370)	(0.0460)	(0.0458)	(0.0370)	(0.0201)	(0.0188)
*lntemp*trend*	0.0646	0.274**	0.242***	0.157*	0.149	0.250**	0.028	0.047	0.041
	(0.0666)	(0.121)	(0.0882)	(0.0868)	(0.104)	(0.113)	(0.0849)	(0.0530)	(0.0499)
*lngp*trend*	0.0962***	0.151***	0.111***	0.0923**	0.0990**	0.138***	0.029	0.0002	0.004
	(0.0289)	(0.0499)	(0.0363)	(0.0364)	(0.0416)	(0.0469)	(0.0192)	(0.0124)	(0.0119)
*lnsva*trend*	−0.133**	−0.353***	−0.264***	−0.254***	−0.233***	−0.325***	0.018	−0.040	−0.030
	(0.0517)	(0.101)	(0.0649)	(0.0706)	(0.0834)	(0.0926)	(0.0390)	(0.0265)	(0.0260)
*lnpd*trend*	0.0117*	0.0209**	0.0143**	0.0145**	0.0148**	0.0203**	−0.070***	0.011	−0.0001
	(0.00690)	(0.00953)	(0.00668)	(0.00649)	(0.00726)	(0.00883)	(0.0221)	(0.0121)	(0.0117)
lnrgdpˆ2	−0.832**	−1.939***	−1.723***	−1.504***	−1.403***	−1.799***	8.039***	1.627**	2.166***
	(0.338)	(0.439)	(0.278)	(0.303)	(0.402)	(0.390)	(0.8000)	(0.7422)	(0.6053)
lntempˆ2	5.373	−27.12**	−28.92***	−23.95***	−17.47	−26.27**	−11.601	−5.387	−5.528
	(8.189)	(11.87)	(8.394)	(8.777)	(11.01)	(10.95)	(11.1310)	(5.2432)	(5.1602)
lngpˆ2	−0.905***	−0.746	−0.749	−0.646	−0.733	−0.752	−0.256	−0.270**	−0.277**
	(0.270)	(0.660)	(0.487)	(0.446)	(0.464)	(0.618)	(0.2260)	(0.1110)	(0.1084)
lnsvaˆ2	−0.723	3.338	1.205	0.972	1.778	2.769	5.204***	1.475*	1.705**
	(1.146)	(2.412)	(1.767)	(1.650)	(1.742)	(2.244)	(1.2513)	(0.8246)	(0.7222)
lnpdˆ2	−0.221*	−0.129	−0.228**	−0.210*	−0.240*	−0.160	1.400***	0.022	0.180
	(0.131)	(0.141)	(0.112)	(0.109)	(0.133)	(0.136)	(0.4798)	(0.2997)	(0.2899)
trendˆ2	−0.00404*	−0.0105***	−0.00799***	−0.00769***	−0.00770***	−0.0101***	0.003***	−0.001	−0.0002
	(0.00206)	(0.00345)	(0.00246)	(0.00243)	(0.00288)	(0.00325)	(0.0012)	(0.0017)	(0.0016)
constant	0.116	−2.066*	−0.608	−0.721	−0.730	−1.835*			
	(0.793)	(1.155)	(0.764)	(0.842)	(1.016)	(1.093)			
Inefficiency equation									
Dep var.:$ln\left(\delta^2_{u}\right)$									
constant	−0.00837	−2.087***	−0.718	−0.721	−1.498*	−1.273*	−3.056	−0.984	−1.739
	(0.0452)	(0.686)	(0.817)	(0.842)	(0.823)	(0.755)	(3.0837)	(1.1630)	(1.0889)
govindex	−0.00837	−0.128**	–	–			−0.198***		
	(0.0452)	(0.0580)					(0.0691)		
rl	–		−1.193	−2.694***	−1.527*	−1.454***		−1.852***	−1.597**
			(0.863)	(0.837)	(0.811)	(0.491)		(0.7099)	(0.6369)
ge	–		−0.131	0.236	0.395			0.810	0.881
			(0.612)	(0.658)	(0.589)			(0.7363)	(0.7872)
ccr	–		2.581***	2.324***	0.0451			−1.354*	−1.382**
			(0.552)	(0.605)	(0.988)			(0.7380)	(0.6713)
rq	–		0.461	−0.896				−0.633	−0.123

(*continued on next page*)

Table 4 (*continued*)

	MI	MII	MIII	MIV	MV	MVI	MVII	MVIII	MIX
			(0.885)	(0.987)				(0.9390)	(0.9523)
va	–		−1.433***	–				0.300	
			(0.551)					(0.6234)	
psav	–		−1.148**	–				0.292	
			(0.465)					(0.3906)	
lngp	−0.00837	−0.0254	0.209	0.169	−0.0188	−0.00788	0.467	0.462	0.449
	(0.0452)	(0.166)	(0.175)	(0.189)	(0.183)	(0.169)	(0.645)	(0.3088)	(0.2932)
l.lngp		0.211	0.233	0.327	0.275*	0.266*			
		(0.148)	(0.153)	(0.163)	(0.152)	(0.145)			
trend	−0.00837	0.267***	0.268***	0.243***	0.265***	0.246***	0.408	0.363***	0.426***
	(0.0452)	(0.0714)	(0.0711)	(0.0850)	(0.0828)	(0.0717)	(0.6025)	(0.1404)	(0.1150)
trend^2	−0.00837	−0.0065***	−0.00565**	−0.00620**	−0.00626**	−0.00581***	−0.021	−0.020***	−0.022***
	(0.0452)	(0.00212)	(0.00255)	(0.00257)	(0.00251)	(0.00214)	(0.0237)	(0.0053)	(0.0043)
lncsb	0.183***	0.292***	0.268***	0.247***	0.287***	0.286***	0.099	−0.032	0.006
	(0.0689)	(0.0677)	(0.0711)	(0.0758)	(0.0725)	(0.0690)	(0.0920)	(0.0747)	(0.0772)
lnbcp	−0.242***	−0.317***	−0.410***	−0.569***	−0.423***	−0.345***	−0.334**	−0.0405***	−0.405***
	(0.0904)	(0.115)	(0.122)	(0.122)	(0.117)	(0.113)	(0.1432)	(0.1385)	(0.1438)
Dep var.:$ln\left(\delta^2_{-w}\right)$									
constant	−0.242***	−3.671***	−3.761***	−3.727***	−3.646***	−3.668***	−3.870***	−3.902***	−3.893***
	(0.0904)	(0.0798)	(0.0800)	(0.0808)	(0.0824)	(0.142)	(0.0788)	0.0699)	(0.0693)
/eta1_lnrgdp	1.854**	4.949***	3.538***	3.074***	3.228***	4.585***	−9.310***	0.784	−0.549
	(0.783)	(1.131)	(0.728)	(0.881)	(1.090)	(1.063)	(1.6221)	(1.1467)	(1.0714)
/eta1_lnrgdp^2							−7.844***	−1.350*	−1.928***
							(0.8267)	(0.7295)	(0.5869)
/e1	0.115***	0.107***	0.103***	0.106***	0.106***	−0.107***			
	(0.00401)	(0.00397)	(0.00384)	(0.00396)	(0.00396)	(0.0040)			
Mu									
constant							0.344	1.075	0.709*
							(1.0766)	(0.7463)	(0.4084)
Eta_end test	5.61**	19.14***	23.59***	12.17***	8.77***		–	11.403***	19.344***
Weak IV test_front eq	26.14***	72.23***	71.29***	55.32***	53.86***	67.86***	8.090**	12.588***	10.900***
Instrument exclusion test_P_value	0.4618	0.4618	0.4618	0.4618	0.4618	0.4618	0.4618	0.4618	0.4618
Average TEE (%)	49.11	47.81	47.48	54.49	50.09	48.96	62.89	61.05	63.35
Median TEE (%)	52.67	45.06	42.42	54.63	49.89	48.39	65.25	63.984	66.36
Log-likelihood	357.27	361.86	395.23	384.04	365.18	362.69	1714.832	1756.44	1753.96
Observations	416	369	369	369	369	369	511	511	511
Countries	44	44	44	44	44	44	44	44	44

Standard errors are in parentheses. Symbols denote significance at the 1 % (***) level, 5 % (**) level, and 10 % (*) level. All models use the share of arable land and legal system origin as instrumental variables. Models MI-MVI are based on winsorized data, retaining 75 % of the original gasoline consumption data and 95 % of the original independent variables data. Models MI-MVI do not account for unobserved heterogeneity. Models MVII-MIX, which use the original data, account for unobserved heterogeneity.

Appendix D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eneco.2024.107970.

References

Abid, M., 2015. The close relationship between informal economic growth and carbon emissions in Tunisia since 1980: the (ir) relevance of structural breaks. Sustain. Cities Soc. 15, 11–21.

Adom, P.K., 2016. The transition between energy efficient and energy inefficient states in Cameroon. Energy Econ. 54, 248–262.

Adom, P.K., 2017. The long-run price sensitivity dynamics of industrial and residential electricity demand: the impact of deregulating electricity prices. Energy Economics 62, 43–60.

Adom, P.K., 2021b. Energy efficiency and financial depth nexus revisited: does the choice of instrumental variable and measure of financial depth matter? Environ. Sci. Pollut. Res. 28, 60080–60094.

Adom, P.K., 2024. Global energy efficiency transition tendencies: A development phenomenon or not? Energy Strategy Rev. 55, 1001524.

Adom, P.K., 2021a. Financial Depth and Electricity Consumption in Africa: Does Education matter? Empirical Economics 61, 1985–2039.

Adom, P.K., Adams, S., 2018. Energy savings in Nigeria: is there a way of escape from energy inefficiency? Renew. Sust. Energ. Rev. 81, 2421–2430.

Adom, P.K., Adams, S., 2020. Technical fossil fuel energy efficiency (TFFEE) and debt-finance government expenditure nexus in Africa. Journal of Cleaner Production 271, 122670.

Adom, P.K., Amoani, S., 2021. The role of climate adaptation readiness in economic growth and climate change relationship: an analysis of the output/income and productivity/institution channels. J. Environ. Manag. 293, 112923.

Adom, P.K., Agradi, M., Barnor, C., 2018. Road Transport Energy Demand in West Africa: A test of the Consumer-tolerable Price hypothesis? International Journal of Sustainable Energy 37 (10), 919–940.

Adom, P.K., Amakye, K., Barnor, C., Quartey, G., Bekoe, W., 2016. Shift in demand elasticities, road energy forecast, and the persistence profile of shocks. Econ. Model. 55, 189–206.

Adom, P.K., Amuakwa-Mensah, F., Akorli, C.D., 2023. Energy efficiency as a sustainability concern in Africa and financial development: how much bias is involved? Energy Econ. 120, 106577.

Adom, P.K., Amuakwa-Mensah, F., Amuakwa-Mensah, S., 2020. Degree of Financialisation and Energy Efficiency in sub-Sahara Africa: Do Institutions matter? Financial Innovation 6, 33.

AfDB, 2013. 'Recognizing Africa's Informal Sector', Group http://www.afdb.org/en/blogs/afdb-championing-inclusive-growth-acrossafrica/post/recognizing-africas-informal- sector-11645/.

Agradi, M., Adom, P.K., Vezzulli, A., 2022. Towards sustainability: does energy efficiency reduce unemployment in African societies? Sustain. Cities Soc. 79, 103683.

Akorli, C.D., Adom, P.K., 2023. The role of corruption control and regulatory quality in energy efficiency transition tendencies in Africa. iScience 26 (3), 106262.

Alesina, A., Spolaore, E., Wacziarg, R., 2005. "Trade, growth and the size of countries." in Handbook of Economic Growth, ed. by P. Aghion, and S. Durlauf. Handbook Econ. Growth 1 (23), 1499–1542. Elsevier.

Alouini, O., Hubert, P., 2019. Country size and economic performance and volatility. Revue de l'ofce No. 164, pp. 139–163.

Amsler, C., Prokhorov, A., Schmidt, P., 2016. Endogenous stochastic frontier models. J. Econ. 190, 280–288.

Amsler, C., Prokhorov, A., Schmidt, P., 2017. Endogenous environmental variables in stochastic frontier models. J. Econ. 199, 131–140.

Amuakwa-Mensah, F., Klege, R., Adom, P.K., Hagan, E., Amoah, A., 2018. Unveiling the energy saving role of banking performance in sub-Sahara Africa. Energy Econ. 74, 828–842.

Babie, Paul T., Wawryk, A., 2021. Climate Change, the Rule of Law, and Market-Based Schemes to Encourage Renewable Energy and Energy Efficiency (May 26, 2021). U.

of Adelaide Law Research Paper No. 2021-47, 29 Michigan State International Law Review 363, Available at SSRN: https://ssrn.com/abstract=3853400.

Bakhat, M., Labandeira, X., Labega, J.M., Lopez-Otero, X., 2017. Elasticities of trans- port fuels at times of economic crisis: an empirical analysis for Spain. Energy Econ. 68 (1), 66–80.

Barbier, E.B., 2004. Agricultural expansion, resource booms and growth in Latin America: implications for long-run economic development. World Dev. 32 (1), 137–157.

Barra, C., Papaccio, A., 2024. Does regulatory quality reduce informal economy? A theoretical and empirical framework. Soc. Indic. Res. 172, 543–567.

Basbay, M.M., Elgin, C., Torul, O., 2016. Energy consumption and the size of the informal economy. Econ. Open-Access, Open-Assess. E-J. 10 (2016–14), 1–28.

Basso, Leonardo J., Oum, Tae Hoon, 2007. Automobile fuel demand: a critical assessment of empirical methodologies. Transplant. Rev. 27 (4), 449–484.

Becker, K.F., 2004. The Informal Economy, SIDA Report. Swedish International Development Cooperation Agency, Stockholm.

Benkraiem, R., Lahiani, A., Miloudi, A., Shahbaz, M., 2019. The asymmetric role of shadow economy in the energy-growth nexus in Bolivia. Energy Policy 125, 405–417.

Benson, E., et al., 2014. Informal and Green? The Forgotten Voice in the Transition to a Green Economy. IIED Discussion Paper, IIED London.

Biswas, A.K., Farzanegan, M.R., Thum, M., 2012. Pollution, shadow economy and corruption: theory and evidence. Ecol. Econ. 75, 114–125.

Blackman, A., 2000. Informal sector pollution control: what policy options do we have? World Dev. 28 (12), 2067–2082.

Bölük, G., Mert, M., 2014. Fossil & renewable energy consumption, GHGs (greenhouse gases) and economic growth: Evidence from a panel of EU (European Union) Countries. Energy 74, 439–446.

Bruns, S.B., Moneta, A., Stern, D.I., 2019. Macroeconomic time series evidence that energy efficiency improvements do not save energy. CAMA Working Paper 21.

Burke, P.J., Nishitateno, S., 2011. Gasoline prices, gasoline consumption and new vehicle fuel economy: Evidence for a large sample of countries. Australian National University Working Papers-trade and development, WP No. 2011/15.

Cebula, R.J., Mixon, F.G., 2014. The roles of economic freedom and regulatory quality in creating a favourable environment for investment in energy R&D, infrastructure and capacity. Am. J. Econ. Sociol. 73, 299–324.

Celestin, Jean-Bernard, 1989. Urban Informal Sector Information: Needs and Meth- Ods. International Labour Office, Geneva.

Chaudhuri, S., Mukhopadhyay, U., 2006. Pollution and informal sector: a theoretical analysis. J. Econ. Integr. 21, 363–378.

Chen, Y, Wang, H, Schmidt, P, 2014. Consistent estimation of the fixed effects stochastic frontier model. J Economet 181, 65–76.

Coelli, T., 1995. Estimators and hypothesis tests for a stochastic frontier function: a Monte Carlo analysis. J. Prod. Anal. 6,,, 247–268.

Colombo, E., Onnis, L., Tirelli, P., 2013. Shadow economies at times of banking crises: Empirics and theory. No. 234 – February 2013 – Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa –Università degli Studi di Milano – Bicocca, Milan.

Cross, F.B., 2002. Law and economic growth. Texas Law Rev. 80, 1737–1775.

Dahl, C.A., 2012. Measuring global gasoline and diesel price and income elasticities. Energy Policy 41, 2–13.

Damaska, M.R., 1986. V the Faces of Justices and State Authority: A Comparative Approach to the Legal Process. Yale University Press, New Haven and London.

Davis, L.W., Gertler, P.J., 2015. Contribution of air conditioning adoption to future energy use under global warming. Proc. Natl. Acad. Sci. 112 (19), 5962–5967.

Dawa, F.O., Kinyanjuim, M.N., 2012. Green Economy and Sustainable Development: Which Way for the Informal Economy. United Nations Research Institute for Social Devel- opment, Geneva. http://www.unrisd.org/news/kinyanjui-dawa#sthash.sv2J3FVw.dpuf.

De Cian, E., Lanzi, E., Roson, R., 2007. The impact of temperature change on energy demand: a dynamic panel analysis. FEEM Nota Di Lavoro 46, 2007.

De Vries, H.J., Verhagen, W.P., 2016. Impact of changes in regulatory performance standards on innovation: a case of energy performance standards for newly-built houses. Technovation 48–49, 56–68.

Deninger, Klaus, 2003. Land Policies for Growth and Poverty Reduction. World Bank and Oxford University Press, New York.

D'Haultfoeuille, X., Hoderlein, S., Sasaki, Y., 2021. Testing and relaxing the exclusion restriction in the control function approach. Journal of Econometrics 240 (2), 105075.

Ding, C., Lichtenberg, E., 2010. Land and urban economic growth in China. J. Reg. Sci. 51 (2), 299–317.

Duffield, J.S., Hankla, C.R., 2011. The efficiency of institutions: political determinants of oil consumption. Comp. Polit. 43 (2), 187–205.

Edziah, B.K., Opoku, E.E.O., 2024. Enhancing energy efficiency in Asian-Pacific: comprehensive energy policy analysis. Energy Econ. 138, 107831.

Elgin, C., Oztunali, O., 2014. Pollution and informal economy. Econ. Syst. 38 (3), 333–349.

Evans, J., Filippini, M., Hunt, L.C., 2013. The contribution of energy efficiency towards meeting CO2 targets. In: Fouquet, Roger (Ed.), Handbook on Energy and Climate Change, pp. 175–223.

Figueroa, A., 2016. Behaviour matters: improving energy efficiency in informal settlements. Briefing paper, no. 7/2016, DIE, Bonn.

Filippini, M., Heimsch, F., 2016. The regional impact of CO2 tax on gasoline demand: a spatial econometric approach. Resour. Energy Econ. 46, 85–100.

Filippini, M., Hunt, L.C., 2011. Energy demand and energy efficiency in the OECD countries: a stochastic demand frontier approach. Energy J. 32 (2), 59–80.

Filippini, M., Hunt, L.C., 2015. Measurement of energy efficiency based on economic foundations. Energy Econ. 52, S5–16.

Filippini, M., Zhang, L., 2016. Estimation of the energy efficiency in Chinese provinces. Energ. Effic. 9 (6), 1315–1328.

Fouquet, R., 2012. Trends in income and price elasticities of transport demand (1850–2010). Energy Policy 50, 62–71.

Fouquet, R., 2015. Long-run demand for energy services: income and price elasticities over two hundred years. Rev. Environ. Econ. Policy 8 (2), 143–333.

Fowler, M., Meeks, R., 2020. Rethinking energy efficiency in the developing world. Retrieved on 13/04/2020 from. http://www.robynmeeks.com/wp-content/uploads/2020/0 3/Fowlie-and-Meeks-2020.pdf.

Fredriksson, Per, G., Vollebergh, Herman, R.J., Dijkgraaf, Elbert, 2004. Corruption and energy efficiency in OECD countries: theory and evidence. Journal of Environmental Economics and Management 47 (2), 207–231.

Fredstrom, A., Peltonen, J., Wincent, J., 2021. A country-level institutional perspective on entrepreneurship productivity: the effects of informal economy and regulation. J. Bus. Ventur. 36 (5), 106002.

Fuso-Nerini, F., Tomei, J., To, L.S., Bisaga, I., Parikh, P., Black, M., Borrion, A., Spataru, C., Broto, V.C., Anandarajah, G., et al., 2018. Mapping synergies and trade-offs between energy and the sustainable development goals. Nat. Energy 3, 10–15.

Gann, D.M., Wang, Y., Hawkins, R., 1998. Do regulations encourage innovations? The case of energy efficiency in housing. Build. Res. Inf. 26, 280–296.

Gately, D., Huntington, H.G., 2002. The asymmetric effects of changes in price and income on energy and oil demand. Energy J. 23 (1), 19–35.

Gennaioli, Caterina and Tavoni, Massimo, Clean or 'Dirty' Energy: Evidence on a Renewable Energy Resource Curse (August 31, 2011). FEEM Working Paper No. 63.2011, Available at SSRN: https://ssrn.com/abstract=1920145 or https://doi.org/10.2139/ssrn.1920145.

Greene, W.H., 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of Econometrics 126, 269–303.

Greene, W.H., 2011. Econometric Analysis, 7th ed. Prentice- Hall, Upper Saddle River.

Haushofer, J., Fehr, E., 2014. On the psychology of poverty. Science 344, 862–867.

Henley, A., Peirson, J., 1996. Energy pricing and temperature interaction: British experimental evidence. In: Aberystwyth Economic Research Paper, No. 9616. University of Wales Aberyst-wyth.

Henley, A., Peirson, J., 1998. Residential energy demand and the interaction of price and temperature: British experimental evidence. Energy Econ. 20, 157–171.

Hernandez, M., 2020. Environmentally harmful dumping of inefficient and obsolete air conditioners in Africa. In: CLASP. https://clasp.ngo/publications/environmentally-harmful-d umping-of-inefficient-and-obsolete-air-conditioners-in-africa.

Hosier, R., 1994. Informal Sector Energy Use in Tanzania: Efficiency and Employment Potential. Energy, Environment and Development Series – No. 25. SIDA, Stockholm. ISBN:91-88116-85-9.

IEA, 2010. Energy Efficiency Governance. IEA/OECD, Paris.

IEA, 2020a. Energy Efficiency 2020. IEA, U.S.A https://webstore.iea.org/download/direct/4259.

IEA, 2020b. Energy Efficiency 2020. IEA. https://webstore.iea.org/download/dir ect/4259.

IEA, Global CO2 emissions in 2019, Retrieved from, https://www.iea.org/articles/global-co2-emissions-in-2019, 2019, 28/02/2020.

Ihrig, J., Moe, K., 2004. Lurking in the shadows: the informal sector and government policy. J. Dev. Econ. 73 (2), 541–577.

ILO, 2012a. Working Towards Sustainable Development: Opportunities for Decent Work and Social Inclusion in a Green Economy, Geneva, Switzerland. http://www.ilo.org/wcmsp5/groups/public/—dgreports/—dcomm/publ/documents/publication/wcms_181836.pdf.

ILO, 2012b. Statistical Update on Employment in the Informal Economy, ILO Department of Statistics. See http://laborsta.ilo.org/applv8/data/INFOR MAL_ECONO MY/2012-06Statistical%20update%20-%20v2.pdf.

ILO, 2012c. Global Employment Trends 2012: Preventing a Deeper Jobs Crisis. International Labour Organization, Geneva.

Jollands, N., Ellis, M., 2009. Energy efficiency governance – an emerging priority. 2009 summer study of the European Council for an Energy-Efficient Economy. Eceee, Nice, France.

Jollands, N., Heffner, G., Pasquier, S., Saussay, A., 2011. Enabling Energy Efficiency through Good Governance. eceee, Nice, France.

Karakaplan, M.U., 2017. Fitting endogenous stochastic frontier models in Stata. Stata J. 17 (1), 39–55.

Karakaplan, M.U., 2022. Panel stochastic frontier models with endogeneity. Stata Journal 22, 643–663.

Karakaplan, M.U., Kutlu, L., 2017a. Handling endogeneity in stochastic frontier analysis. Econ. Bull. 37, 889–891.

Karakaplan, M.U., Kutlu, L., 2017b. Endogeneity in panel stochastic frontier mod- els: an application to the Japanese cotton spinning industry. Appl. Econ. 49, 5935–5939.

Karanfil, F., 2008. Energy consumption and economic growth revisited: does the size of unrecorded economy matter? Energy Policy 36 (8), 3029–3035.

Kirk, E.A., Besco, L., 2021. Improving energy efficiency: the significance of nor- mativity. J. Environ. Law 33, 669–695.

Kopp, R.J., 1981. The measurement of productive efficiency: a reconsideration. Q. J. Econ. 96, 477–503.

Kumbhakar, S.C., et al., 2014. Technical efficiency in competing panel data models: a study of Norwegian grain farming. J. Prod. Anal. 41 (2), 321–337.

Kutlu, L., 2010. Battese-Coelli estimator with endogenous Regressors. Econ. Lett. 109, 79–81.

Kutlu, L., Tran, K.C., Tsionas, M.G., 2019. A time-varying true individual effects model with endogenous Regressors. J. Econ. 211, 539–559.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1997. Legal determi- nants of external finance. J. Financ. 52, 1131–1150.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1998. Law and finance. J. Polit. Econ. 106, 1113–1155.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1999. The quality of government. J. Law Econ. Org. 15, 222–279.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 2002. Investor protection and corporate valuation. J. Financ. 57, 1147–1170.
La Porta, R., Lopez-de-Silanes, F., Pop-Eleches, C., Shleifer, A., 2004. Judicial checks and balances. J. Polit. Econ. 112, 445–470.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., 2008. The economic consequences of legal origins. J. Econ. Lit. 46, 285–322.
Lanzi, E., Verdolini, E., Haščič, I., 2011. Efficiency-improving fossil fuel technologies for electricity generation: Data selection and trends. Energy Policy 39 (11), 7000–7014.
Lee, E., 1998. The Asian Financial Crisis: The Challenge for Social Policy. ILO, Geneva, Switzerland. http://www.ilo.org/public/libdoc/ilo/1998/98B09_292_engl.pdf.
Levine, R., 2002. Bank-based or market-based financial systems: which is better. J. Financ. Intermed. 11, 398–428.
Li, Y., Pizer, W.A., Libo, Wu, 2019. Climate change and residential electricity consumption in the Yangtze River Delta, China. Proc Natl Acad Sci 116 (2), 472–477.
Loayza, N., Oviedo, A.M., Serven, L., 2006. The impact of regulation on growth and informality: Cross- country evidence. In: Guha- Khasnobis, B., Kanbur, R., Ostrom, E. (Eds.), Linking the Formal Economy: Concepts and Policies. Oxford University Press, Oxford, pp. 121–144.
Malghan, B., 2019. Efficiency. In: Kothari, A., et al. (Eds.), Pluriverse: A Post develop- ment Dictionary (50-53). Tukila books, New Delhi, India.Marquette, H. (2012). "Finding God" or "Moral Disengagement"in the Fight against Corruption in Developing Countries? Evidence from India and Nigeria. Public Adm. Dev. 32, 11–26.
Marais, E.A., et al., 2019. Air quality and health impact of future fossil fuel use for electricity generation and transport in Africa. Environmental Science and Technology 53, 13524–13534.
Marquez, J., 1994. The constancy of illusions or the illusion of constancies: income and price elasticities for U.S. imports, 1890–1992; international finance discussion papers 475. Board of Governors of the Federal Reserve System (U.S).
Medina, L., Schneider, F., 2018. Shadow economies around the world: what did we learn over the last 20 years? IMF Working Paper, WP/18/17.
Medina, L., Jonelis, A., Cangul, M., 2017. The informal economy in sub-saharan Africa: size and determinants. IMF Working Paper, WP/17/156.
Mutter, R.L., Greene, W.H., Spector, W., Rosko, M.D., Mukamel, D.B., 2013. Investigating the impact of endogeneity on inefficiency estimates in the application of stochastic frontier analysis to nursing homes. J. Prod. Anal. 39, 101–110.
OECD, 2009. Competition Policy and the Informal Economy. OECD, Paris.
OECD/IEA, 2011. OECD Green Growth Studies: Energy. Retrieved on. https://www.oecd.org/greengrowth/greening-energy/49157219.pdf2. (Accessed 3 June 2019).
Parkinson, M., Meegan, R., Karecha, J., 2014. City size and economic performance: is bigger better, small more beautiful or middling Marvellous? Eur. Plan. Stud. 23 (6), 1054–1068.
Petrick, S., Rehdanz, K., Tol, R.S.J., 2010. The Impact of Temperature Changes on Residential Energy Consumption. Kiel Working Paper No.1618.
Porter, M.E., 1991. America's green strategy. Sci. Am. 264, 168.
Raghoo, P., Surroop, D., 2020. Price and income elasticities of oil demand in Mauritius: an empirical analysis using cointegration method. Energy Policy 140, 11400.
Rashid, K., Safdarnejad, S.M., Powell, K.M., 2019. Process intensification of solar thermal power using hybridization flexible heat integration and real-time optimization. Chemical Engineering and Processing: Process Intensification 139, 155–171.
Regina, M., Schlomann, B., Krail, M., Rohde, C., 2016. Towards a green economy in Germany? The role of energy efficiency policies. Appl. Energy 179, 1293–1303.
Roe, M.J., 2006. Legal origins, politics, and modern stock markets. Harv. Law Rev. 120, 462–527.
Rubin, D.B., 1987. Multiple Imputation for Nonresponse in Surveys. Wiley.
Steven, P., 2020. Gasoline demand in middle-income countries. Econ. Bull. 40 (4), 1–13.
Suri, G.V., Chapman, D., 1998. Economic growth, trade and energy: implications for the environmental Kuznets curve. Ecol. Econ. 25 (2), 195–208.
Swain, R.B., Kambhampati, U.S., Karimu, A., 2020. Regulation, governance and the role of informal sector in influencing environmental quality. Ecol. Econ. 173, 106649.
Syrquin, M., Chenery, H.B., 1988. Patterns of development, 1950–1983. World Bank Discussion Papers no. 41, Washington, DC.
Tandon, N., 2012. Empowerment of Women in a Green Economy in the Context of Sustainable Development and Poverty Eradication. Networked Intelligence for Develop- ment. http://www.networkedintelligence.com/wp/wp-content/uploads/2012/04/Empower ment-of-Women-in-a-Green-Economy-Tandon.pdf.
Tirkaso, W.T., Gren, I.-M., 2020. Road fuel demand and regional effects of carbon taxes in Sweden. Energy Policy 144, 11648.
Tran, K.C., Tsionas, E.G., 2013. GMM estimation of stochastic frontier model with endogenous regressors. Econ. Lett. 118, 233–236.
Tran, K.C., Tsionas, E.G., 2015. Endogeneity in stochastic frontier models: copula approach without external instruments. Econ. Lett. 133, 85–88.
USAID, 2005. Removing Barriers to Formalization: The case for reform and emerging best practice, United States Agency for International Development http:// www.oecd.org/dac/povertyreduction/38452590.pdf.
Wang, H.J., Ho, C.W., 2010. Estimating fixed-effect panel stochastic frontier models by model transformation. J. Econ. 157, 286–296.
Wenz, L., Levermann, A., Auffhammer, M., 2017. North–south polarization of European electricity consumption under future warming. Proc. Natl. Acad. Sci. 114 (38), E7910–E7918.
WIEGO. Urban informal workers and the green economy. Retrieved on 01/012020 from http://www.wiego.org/sites/default/files/resources/files/WIEGO_Urban_Informal_Workers_Green_Economy.pdf.
Wilcox, R.R., 1998. The goals and strategies of robust methods. Br. J. Math. Stat. Psychol. 51, 1–39.
Wilcox, R., Keselman, H., 2003. Modern robust data analysis methods: Measures of central tendency. Psychological Methods 8, 254–274.
Xu, G., 2011. The role of law in economic growth: a literature review. J. Econ. Surv. 25 (5), 833–871.
Zhang, L., 2017. Correcting the uneven burden sharing of emission reductions across provinces in China. Energy Econ. 64, 335–345.
Zhang, L., Adom, P.K., 2018. Energy efficiency in China: how persistent are the movements to/from the frontier. Energy J. 39 (6), 147–169.
Zhang, H., Mu, J.E., McCarl, B.A., Yu, J., 2022. The impact of climate change on global energy use. Mitig. Adapt. Strateg. Glob. Chang. 27, 9.