Makananisa, Mangalani Peter; Koloane, Cathrine Thato; Schneider, Friedrich

Article

Modelling the shadow economy of South Africa : using the currency demand and MIMIC approach

Journal of economics and political economy

Provided in Cooperation with:
KSP Journals, Istanbul

Reference: Makananisa, Mangalani Peter/Koloane, Cathrine Thato et. al. (2020). Modelling the shadow economy of South Africa : using the currency demand and MIMIC approach. In: Journal of economics and political economy 7 (1), S. 27 - 46.
doi:10.1453/jepe.v7i1.2025.

Journal of
Economics and Political Economy
www.kspjournals.org
Volume 7 March 2020 Issue 1

Modelling the shadow economy of South Africa: Using the currency demand and MIMIC approach

By Mangalani P. MAKANANISA [a†]
Cathrine T. KOLOANE [b]
& Friedrich SCHNEIDER [c]

Abstract. This paper estimates the size and development of the South African shadow economy (SE) using two indirect approaches namely, the Multivariate Indicator Multivariate Causes (MIMIC) model and the Currency Demand Approach (CDA). The study uses time series from 2000 to 2019 (using quarterly data) to estimate the SE of South Africa for the period 2004 to 2018. The average estimated size of the SE from the CDA and MIMIC model are 22.47% and 25.45% respectively. Overall, the MIMIC and CDA models are both showing a slight decreasing trend for the same period. The study recommends further analysis to be conducted on economic segments in order to explore the SE activity distribution between different economic sectors; resulting in an easier way to identify, locate and monitor unrecorded businesses and also increase revenue collections and minimise non-compliance for different sectors.
Keywords. Shadow Economy of South Africa, GDP, CDA, MIMIC.
JEL. C32, H26, I2, O17, P48.

1. Introduction

South Africa is the second largest economy in Africa, second to Nigeria. Among the key sectors that keep the economy running are the manufacturing, wholesale and retail trade, financial services, transport, mining, agriculture and tourism. The economy is however dominated by finance and business services (19%), government (18%), trade (15%) and manufacturing (13%). The South African economic growth has been stagnant in recent years and the current load shedding has aggravated the situation with no real growth prospects expected in the future. The International Monetary Fund (IMF) has cut South Africa's growth prospects for 2020, expecting only a measly 0.8% growth rate (Business Tech, 2020).

[†a] South African Revenue Service, Operational Research, Pretoria, South Africa.
☎. +2782 4564669 ✉. manginduvho@gmail.com
[b] South African Revenue Service, Operational Research, Pretoria, South Africa.
☎. +2782 906 0090 ✉. Tkoloane09@gmail.com
[c] Johannes Kepler University Linz, Research Institute for Banking & Finance, Linz, Austria.
☎. +43 732 2468 7340 ✉. Friederich.schneider@jku.at

Traditionally, South Africa's economy has been mainly dependent on the primary sector and mining was a key output, particularly gold mining. Mining has been the foundation and the driving force behind South Africa's economic and industrial development that carried South Africa onto the world stage. However, over the past 15 years, the overall mining production has declined marginally mainly due to the decline in gold production (SA Mine, 2019). Since the early 1990s, South Africa has seen a structural shift in output with the tertiary sector, which includes wholesale and retail trade, tourism and communications, now taking centre stage (Brand South Africa, 2018).

According to the Mineral Council South Africa (2020), illegal mining is on the rise in South Africa. It is now taking place on a large scale nationally. Furthermore, the UNCTAD report (2016) refers to fraudulent misinvoicing of precious metals by traders in order to evade and avoid payment of taxes.

South Africa is used as a strategic gateway to Europe by most African countries. Its great infrastructure and accessible modes of transport has made it even more attractive for foreign investments. As a result, South Africa has seen an influx of foreigners establishing businesses in the country. Some of these foreigners are en- route to Europe and whilst in the country establish small businesses such as retail shops, salons, internet cafes and hardwares to earn a living. They usually prefer payment in cash in order to avoid payment of any taxes. South Africa has been one of the continent's top destination for Chinese investment for years. China malls and other less formal business have mushroomed everywhere in the cities and even as far as remote villages. Some of these businesses also prefer payment by means of cash.

The Global Corruption Barometer Africa (2019) reveals that 64% of South Africans believe that corruption has increased in 2018 and 70% thinks that the government is not doing enough to tackle corruption. State capture has become the buzzword in most citizens' lips. The perceived high levels of corruption is discouraging citizens to pay their taxes (New York Times, 2018).These and other factors serve as a fertile ground for SE activities.

The inability of the South African tax authority to meet its yearly tax revenue has further fuelled speculation that the formal economy might be losing out to the informal economy. Tremendous pressure has been mounting on the tax authority to quantify the size of the SE to enable targeted enforcement actions (Lol, 2019).

It is against this background that this study was initiated. This paper aims to contribute in understanding the development and size of the SE in South Africa for the period 2004 to 2018.The paper will only focus on the legal activities that would ordinarily contribute to the country's GDP. To our knowledge, this is one of the first comprehensive study to combine the Currency Demand Approach and the MIMIC model to estimate the size of the SE in South Africa.

The rest of this paper is organised as follows. Section 2 presents the literature review. Section 3 shows the theoretical models. Section 4 provides the model results and analysis. Section 5 shows the discussion. A conclusion is shown in section 6. Finally, appendices are provided in section 7.

2. Literature review

Many developing countries are struggling to measure the SE mainly because of the ambiguity in its definition and the fact that it is so closely intertwined with the formal economy. SE is sometimes referred to as the underground economy, illicit economy, hidden economy, grey economy, black economy, unrecorded economy, cash economy and informal economy. It may refer to either illegal activities or legal activities, which were not subjected to taxes or the required licenses (Hall, 2019).

Medina & Schneider (2019) defines the shadow or informal economy as all economic activities hidden from the official authorities for monetary, regulatory and institutional reasons. Monetary reasons could be avoiding paying taxes and all social security contributions, regulatory reasons could be the regulatory framework burden or government bureaucracy and institutional reasons include the weak rule of law, corruption law or the quality of political institutions.

SE activities can be either legal or illegal. Sharapenko (2009) refers to a definition proposed by Popov (1999) which defines the SE as the aggregate of illegal economic activities that feed felonies of different degrees. The author concentrates on the illegal activities of the economy, which would not necessarily be subjected to taxes and are therefore not the focus of this research paper.

There are various methods used to measure the SE but can generally be divided into two categories namely: direct and indirect approaches.

2.1. Direct approaches

These microeconomic approaches rely on surveys and tax auditing. The main advantage of surveys is the wealth of information that can be obtained on the structure of the SE. However, surveys are dependent on the honesty of the respondents and their willingness to disclose any illicit activities. The results from surveys are sensitive to the way the questionnaires were formulated which also affects the comparability of the results across countries. Tax compliance data might be biased as the selection of taxpayers is not random but based on the likelihood of tax fraud and the estimates will only include the portion of SE discovered by the authorities and not all hidden activities (Medina & Schneider 2018; Schneider & Buehn, 2016).

2.2. Indirect approaches

These are macroeconomic approaches that use various economic and other indicators that contain information about the development of the SE

over time (Schneider & Buehn, 2016; Schneider & Enste 2000; Medina & Schneider 2018). The six main indirect approaches are:

- The discrepancy between national expenditure and income statistics

In national accounting, the income measure of Gross National Product (GNP) should equal the expenditure measure of GNP. The discrepancy between the expenditure measure and the income measure can be used as an indicator of the extent of the SE.

- The discrepancy between official and actual labour

A decrease in the labour force of the formal economy can indicate increased activity in the SE. However, these estimates are perceived as weak indicators of the size and development of the SE as individuals can simultaneously participate in the shadow and formal economies (Schneider & Buehn, 2016).

- The transactions approach

This approach is based on the assumption that there is a constant relation over time between the volume of transactions and official GNP. The main disadvantage of this method is that precise figures of the volume of transactions is not easily available especially for cash transactions. Furthermore, an assumption is made that the gap between the volume of transactions and GNP is due to SE which might not necessarily be the case (Schneider & Buehn, 2016).

- The physical input method (electricity consumption)

This method involves two approaches namely: The Kaufmann – Kaliberda (1996) Method and the Lackó method. In the Kaufmann - Kaliberda method, electric power consumption is used as an indicator of the overall economic activity of a country. Lackó (1999, 2000a, b) assumes that a certain part of the SE is associated with the household consumption (HHC) of electricity.

- The Currency Demand Approach

Originally formulated by Cagan (1958), this approach assumes that SE activities are conducted using cash. An increase in the size of the SE will therefore increase the demand for currency (Schneider & Buehn, 2016). Although widely used, this approach has the following disadvantages:

- ➢ This approach may underestimate the size of the SE as not all transactions in the SE are conducted in cash.
- ➢ The absence of reliable data such as tax morality is a challenge.

- The MIMIC Approach

The MIMIC model is a special type of structural equation modelling (SEM) based on the statistical theory of unobserved variables (Hassan & Schneider, 2016). This model is confirmatory as it confirms the influence of a set of exogenous causal variables on the latent variable (SE), and the effect of the SE on macroeconomic indicator variables (Farzanegan, 2009).The MIMIC model has the following advantages:

- ➢ Considers multiple indicator and causal variables at the same time.
- ➢ Flexible - varies causal and indicator variables depending on the particular features of the SE activity.

- Uses maximum likelihood (ML) estimation procedures, which are well known and are generally optimal if the sample size is large enough.
- Does not need any restrictive assumptions to operate.
- Can be applied to other informal economic activities not only the SE activities.

The main disadvantages of this model is in its confirmatory nature i.e. it confirms a given model and does not find a suitable model. The benchmarking procedure requires experimentation, and a comparison of the calibrated values in a wide academic debate. No consensus exists on the most reliable benchmarking procedure.

3. The two used theoretical models

3.1. MIMIC model

The multivariate indicators multivariate causes (MIMIC) model refers to the type of models that involve a set of indicators and a set of causals variables. This type of regression involves multiple equations and can also include some unknown variable(s) to be solved or estimated from the observed variables provided there is a theoretical relationship between the variable(s). Thus, one needs to consider a theoretical relationship or a hypothetical relationship between the unobserved, causals and the indicators to be included in the model

The MIMIC model with unobserved (latent) variables are special type of models based on the covariance among variances to estimate the latent variable. Hence this models are sensitive to the data skewness and outliers (Gana & Broc, 2019). The transformation of data to normality becomes crucial in order to obtain reliable model results. The theoretical hypothesis concerning the variables included and the direction of the relationship should be the foundation as several models could best fit the latent variable hence the relationship is spurious.

Nevertheless, the aim of the MIMIC model with latent variable is to obtain a model whose covariance matrix approaches or mimic the observed sample covariance matrix (Hassan & Schneider, 2016). The model is divided into two parts, which are the Measurement and the Structural equation model. The measurement part of the model includes a set of indicators and the structural equation consists of causal variables.

The mathematical representation of the MIMIC model with latent variables is shown in equation 1 and 2.

$$Y_i = \lambda_i \eta + \varepsilon_i, \ i = 1, \dots, n \quad (1)$$

Equation 1 above is the measurement model where:

Y_i represents a vector of indicators,
λ_i represents the regression coefficients,
ε_i represents a vector of white noise errors, and
η represents the latent variable

$$\eta_t = \alpha_1 X_{1t} + \alpha_2 X_{2t} + \cdots + \alpha_p X_{pt} + \varepsilon_t \quad (2)$$

Equation 2 above is the structural equation model (SEM) where:
$X_1, X_2, \ldots, X_p$ represents exogenous causal repressors,
$\alpha_1, \alpha_2, \ldots, \alpha_p$ represents the model coefficients,
η_t represents the unobserved or latent variable, and
ε_t represents the model error term

From equation 1 and 2, the general structure of the MIMIC model can be represented by Figure 1 below;

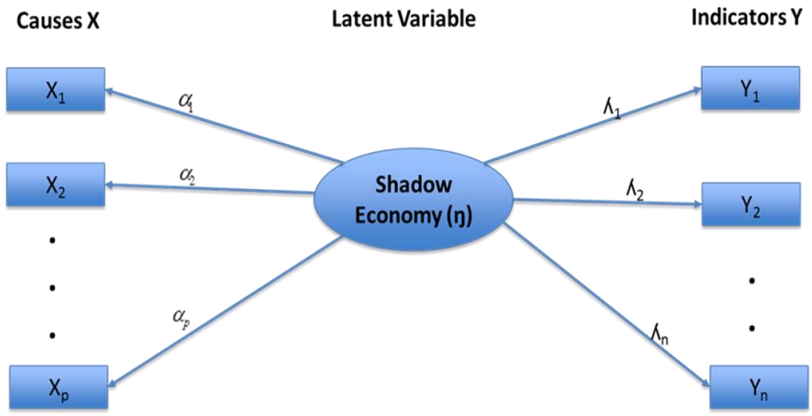


Figure 1. *General structure of a MIMIC model*

However, the MIMIC model only generates the fitted indices and requires prior estimation of the size of the SE to be available. Thus, to obtain the series of estimates for the latent variable a benchmark procedure shown in equation 3 is used.

$$\hat{\eta}_t = \frac{\tilde{\eta}_t}{\eta_{base\_year}} \eta * base\_year \quad (3)$$

where:
η_{base_year} represents the value of the MIMIC index in thebase year,
$\eta * base\_year$ represents prior estimation of the size of the shadow economy in South Africa in the base year
$\tilde{\eta}_t$ represents the value of the MIMIC index at time t,
$\hat{\eta}_t$ represents the SE at time t

The causal variables considered in the case of South Africa are discussed in the following section.

3.1.1. Causal variables

a. Tax burden

This is the most important and widely used variable affecting the size of the shadow economy. A considerable amount of studies confirms a highly

significant positive effect of tax burden on the shadow economy (Amoh & Adafula, 2019; Ariyo & Bekoe, 2012; Klaric, 2010).

Taxes increase the production costs of goods and services, which translates into a higher selling price in the formal market. As a way of increasing one's wealth, individuals might be tempted to evade tax. Therefore it is reasonable to assume that the greater the tax burden, the greater the willingness to evade it and underground and informal production is more likely to occur.

Hypothesis: The higher the tax burden, the larger the size of the shadow economy, ceteris paribus.

b. Business regulations

A highly regulated economy may reduce choices available to individuals and might lead them into the informal economy. Regulations such as barriers to entry and certain policies can drive businesses to consider trading in the informal economy (McMillan, 2006). In the South African context, most regulations such as the Black Economic Empowerment (BEE) were introduced to redress the imbalance of the past brought about by the apartheid government. According to Frontier Economics (2012), regulations introduced to address equality and social cohesion are most likely to have a negative impact on the formal economy. Empirical studies such as Schneider (2005), Enste (2005) and Schneider & Hametner (2007) suggest that increased regulation leads to a growing SE.

The South African economy is somewhat oligopolistic in that there are only a few players in different economic sectors because of the high level of regulation in some sectors. For example, the South African banking sector is mainly dominated by five or so commercial banks while, for instance, the Kenyan banking industry comprises of about 43 commercial banks. Currently, Kenya has one of the fastest growing economy in Africa as a result. Regulation is one of the factors that determines the ease with which businesses operates.

Hypothesis: The higher the regulations in an economy, the higher the SE activities, ceteris paribus.

c. Unemployment

Unemployment has an ambiguous effect on SE. Schneider & Hametner (2007) and Dell'Anno *et al.*, (2007) established that unemployment can influence individuals to operate in the SE to find jobs. If individuals cannot find work in the formal economy, they will look for it in the informal one. In such cases, SE can offer them relatively higher income as no taxes are paid.

On the other hand, Hassan & Schneider (2016) showed that in the Egyptian economy unemployment does not affect the development of the SE over time as the availability of jobs in both informal and formal

economy is limited due to the continuous contraction of the overall economy.

In the case of South Africa, the persistent rise in unemployment can directly contribute to the increase in SE as people who get unemployed or laid-off from the formal economy seek alternative means of survival in the SE.

Hypothesis: The higher the unemployment, the larger the size of the shadow economy, ceteris paribus.

d. Self-Employment

The rate of self-employment as a percentage of labour force in the formal economy is regarded as one of the factors influencing the shadow economy (Hassan & Schneider, 2016). A rise in self-employment increases the potential number of opportunities to conceal income from the authorities (Dell'Anno *et al.*, 2007). Currently, over 10% of South Africa's workforce is self-employed. Due to the difficulties faced in the South African job market, another dominant cause of the rise in self-employment is the lack of finding jobs. Other reasons include unhappiness at traditional jobs, desire for greater flexibility, choice of own workspace or constant conflict with managers or colleagues. Over the last few years, there has been a significant mind-set shift and with it has emerged a workforce, which values flexibility over stability. It has led to an increased entrepreneurial activity in the South Africa (Koekemoer, 2018).

Hypothesis: The higher the self-employment rate, the larger the size of the shadow economy, ceteris paribus.

e. Household debt

If individuals become indebted, they are more likely going to avoid paying taxes as tax is not amongst the priority items on which household income would be spent. South Africa has a higher rate of debt to disposable income ratio, this will most likely tempt heavily indebted households to look for other sources of income to supplement the main income, and this additional income will often not be declared.

Hypothesis: The higher the household debt, the increase in the SE activities, ceteris paribus.

f. Government employment

The expected sign for this indicator is ambiguous. Some authors find a negative relation arguing that in some sectors the presence of the state could disincentive people to incorporate in the SE. In South Africa, the government wage bill is considered very high and the current government is considering laying off some of the aging and unproductive workforce.

Retrenchments and early retirement packages will in most likelihood result in laid off workers engaging in SE activities.

Hypothesis: The lower the government sector, the larger the SE activities.

However, other papers find a positive relation arguing that a rise in the size of the public sector gives relevant incentive to enter in the informal sector. Dell'Anno (2007) states that most researchers support a decreasing role of the public sector in the economy. An increased public sector means that government officials have more power over decisions and will result in more corruption. Schneider & Enste (2002) agrees by stating that bribery and dishonesty of civil servants is another determinant of the SE.

Hypothesis: The larger the government sector, the larger the SE activities, ceteris paribus.

g. Size of mining sector

The mining sector plays a significant role in the South African economy. According to the UNCTAD report (2016), South African miners of silver, platinum group metals, gold and iron ore had systematically and fraudulently indulged in mis-invoicing in order to evade taxes and other legal obligations.

The emergence and growth of illicit mining activities in South Africa contributes to the growth in the SE as most if not all the output is sold in the black market as the miners and people they supply are most likely not licensed to trade in the commodities extracted.

Hypothesis: The lower the formal mining sector activities, the higher the SE activities, ceteris paribus.

3.1.2. Indicator variables

a. GDP

Existing literature indicates that there is an ambiguous relationship between formal economy and the shadow economy (Hassan & Schneider, 2016). An increase in the size of the SE leads to a decrease in the official economy because productive resources and factors are absorbed by the SE creating a depressing effect on the growth of the official economy. The lower the GDP, the most likely it is that the people will look for opportunities in the SE (Buehn & Farzanegan, 2013; Schneider & Enste, 2013; Dell'Anno *et al.*, 2007).

Hypothesis: The larger the size of the SE, the lower the GDP, ceteris paribus.

Other authors such as Schneider *et al.*, (2003) find a positive relationship between formal and informal economy. The SE might allow poor people to find ways to produce and sell cheap products as a way of generating

income. The increased demand in the informal economy spills over into the formal economy.

Hypothesis: The larger the size of the SE, the larger the GDP, ceteris paribus.

Schneider (2005) argues that the relationship is negative for developing countries and positive for the developed and transition countries.

b. Labour force participation rate

Labour force participation rate has an ambiguous effect on the SE. On the one hand, SE absorbs resources from the formal economy as human capital shifts to the SE. Therefore, there is a negative relationship between the labour force and the SE.

Hypothesis: The larger the size of the SE, the lower the labour force participation rate.

On the other hand, Dell'Anno (2007) found a positive significant relationship between the SE and labour force participation in Portugal. Registered labour force might conduct informal activities during holidays, after working hours or on weekends.

Hypothesis: The larger the size of the SE, the larger the labour force participation rate, ceteris paribus.

c. Currency in circulation outside the banks

The shadow or hidden transactions are mostly conducted in cash rather than with credit/debit cards, cheques or bank transactions in order to avoid detection by authorities. An increase in the size of the SE will therefore increase the demand for currency (Schneider & Buehn, 2016).

In South Africa, there has been an influx of foreigners who set up their businesses and prefer payment in cash. Some of traders in the growing number of China malls also prefer payment in cash. Therefore one can expect an increase in the shadow economy.

Hypothesis: The larger the size of the SE, the larger the currency in circulation, ceteris paribus.

3.2. Currency demand approach

A currency demand function can be written as:

$$C_0 = A(1+\theta)^{\alpha} Y_0^{\beta} \exp(-\gamma i) \quad (4)$$

Where C_0 stands forobserved cash and θ represents the incentive variable that motivates individuals to make hidden transactions. This is a key variable in the CDA and can be approximated by the tax burden or the

intensity of government regulation. Y_0 is the official real GDP and $i$denotesthe interest rate or inflation rate representing the opportunity cost of holding cash. A, α, β, γ represents positive parameters (Cagan, 1958).

Data on the following variables was collected for possible inclusion in the CDA model: currency in circulation, nominal GDP per capita, tax burden, deposit interest rate, unemployment, self-employment and government employment per labour force (representing a regulatory indicator). Unemployment had very high correlation with GDP and government employment (approximately 90%), hence it was excluded from the model and deemed to be explained by the two variables. Government employment and self-employment were highly correlated (greater than 80%).

In order to capture the long-run relationships of the explanatory variables on currency demand, the following model was constructed:

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 TAX_t + \beta_3 REG_t + \beta_4 SELF_t + \beta_5 R_t + \varepsilon_t \quad (5)$$

with $\beta_1 > 0$, $\beta_2 > 0$, $\beta_3 > 0$, $\beta_4 > 0$ and $\beta_5 < 0$

where C_t represents the natural logarithm of currency in circulation normalised by GDP

Y_t represents the natural logarithm of nominal GDP per capita

TAX_t represents the natural logarithm of total of tax revenues normalised by nominal GDP

REG_t approximated by public employment in relation to total labour force

$SELF_t$ represents the natural logarithm of the self-employment per capita.

R_t represents the natural logarithm of the deposit interest rate

ε_t represents the error term

$\hat{C}$, which is the amount of currency demand in both the formal and shadow economies, can be estimatedfrom equation (5). Setting the incentive variable (θ) to the minimum and leaving all the other variables unchanged, yields $\tilde{C}$. The difference between $\hat{C}$ and $\tilde{C}$ is the extra currency in the economy, referred to as EC. This is the illegal money used in the SE to conduct transactions. Assuming equal income velocity of currency (v) for both the shadow and the formal economy, the size of the SE is estimated by multiplying EC by the income velocity of currency gives the estimate of the SE v i.e.

$$Y_{informal} = EC * V \quad (6)$$

However, equal income velocity of currency only holds if $\beta_1 = 1$. If $\beta_1 \neq 1$, a correction method needs to be applied to the results to obtain accurate estimates of the SE. The following proposed method by Ahumada *et al.*, (2007) was applied:

$$\frac{Y_{informal}}{Y_{formal}} = \left[\frac{C_{informal}}{C_{formal}}\right]^{\frac{1}{\beta}} = \left[\frac{\hat{Y}_{informal}}{\hat{Y}_{formal}}\right]^{\frac{1}{\beta}} \quad (7)$$

where$Y$ is GDP ,$C$is currency and $\beta$is the income elasticity.

4. Econometric/empirical results and analysis

This section discusses the econometric/empirical results from both the CDA and MIMIC model. The main sources of data for the models were Statistics South Africa (STATSSA) and South African Reserve Bank (SARB). Quarterly data series from the period 2000 to 2019 was used for the analysis to derive the annual shadow economy estimates.

The analysis of the results includes the best-fit models, evaluation of the results or model diagnostic and the model estimates for the period of interest.

4.1. CDA model

Several variables such as unemployment rate, government employment, self-employment, HHC, household debt (HHD) and inflation rate were considered for inclusion in the model. However due to multi collinearity and non-stationarity of the variables, only tax burden, self-employment, government employment, deposit interest rate and nominal GDP were included in the model. The variables were tested for the presence of unit root using the Augmented Dicky-Fuller (ADF) test. Based on the results of the unit root test, the time series were non-stationary at level but after taking the first differences, the time series became stationary. Since the variables are all integrated of the same order I(1), Johansen cointegration test was used to test for cointegration. For this test, the optimal lag length was determined using a VAR model. The optimal lag length is 4 according to Hannan-Quinn (HQ) information criterion. Using a lag length of 4, the Johansen cointegrating equation test indicated a rank of 3 for both the trace statistic and the max-eigen value test at 5% critical value. It can then be concluded that at least3 cointegrating relationships exists between the variables in the long run.

The Vector Error Correction Model (VECM) can now be used as the stationarity and the cointegration requirements have been satisfied.

4.1.1. Fitted CDA model and model evaluation

Table 1 below shows the model results. All the variables are significant at 5% level. As expected, the VECM results from Table 2 above shows positive coefficient for tax burden, regulator indicator (government employment), self-employment. Contrary to expectation, a positive coefficientis observed for deposit interest rate. A negative coefficient is observed for GDP.

Table 1. *The results of the CDA model from 2001 to 2018*

Variables	Statistics
C	1.00
	-0.937
Y	(0.437)
	[2.14]
	0.190
TAX	(0.082)
	[2.32]
	0.659
REG	(0.323)
	[2.04]
	0.092
R	(0.044)
	[2.08]
	0.339
SELF	(0.148)
	[2.29]
Constant	-0.005
	38.858
Autocorrelation LM test	p-value 0.342
Jarque-Bera (Normality test)	1.329
	p-value 0.515

Notes: All variables are in natural logarithms. Standard errors are in parentheses () and T-statistics in [].
Source: Authors computation

There is no serial correlation of residuals as p-value of the LM test is greater than 0.05. The Jarque-Bera test indicates that the residuals are normally distributed.

Table 2 below shows the results for the cointegrating equations. The third error correction term indicates long-term causality between the variables.

Table 2. *Cointegrating equations*

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
D_lncur_cgdp						
_ce1						
L1.	-.0176491	.2432886	-0.07	0.942	-.494486	.4591878
_ce2						
L1.	.1782368	.0549109	3.25	0.001	.0706134	.2858602
_ce3						
L1.	-.6609272	.2297285	-2.88	0.004	-1.111187	-.2106676

Source: Authors computation

4.1.2. CDA shadow economy estimates

Based on the results above, the estimates of the size of the SE are shown by Figure 2 below.

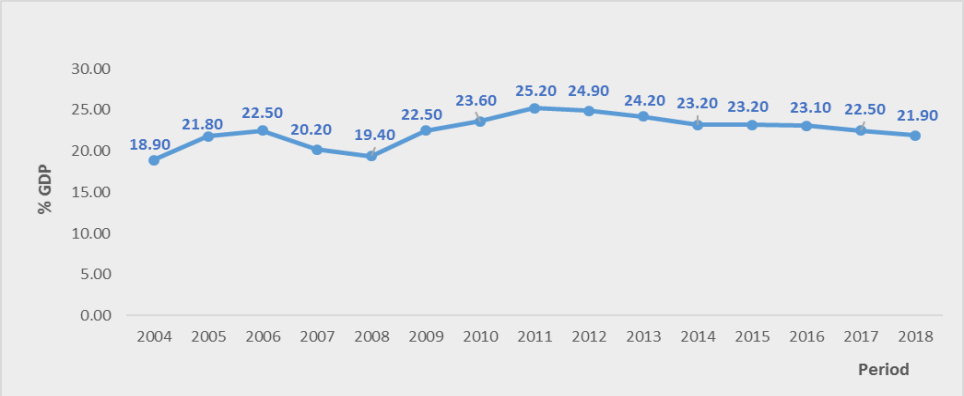


Figure 2. Size and development of the South African SE as % of GDP using CDA
Source: Authors computation

The CDA model estimates the SE to be 22.47% on average for the period 2004 to 2018, with the minimum value of 18.90% in 2004 and the maximum value of 25.20% in 2011. The estimates have been marginally decreasing from 2012 and the figures are stable around 22%.

4.2. MIMIC model

The three indicators considered to explain the shadow economy activities were nominal GDP, labour force participation rate (LFR) and currency in circulation (CURRENCY). Five causal variables were found to be significant in explaining the SE activities namely; tax burden (TAXB), household income (HHI), HHD & Unemployment (UNEMP) interaction, HHC and the sector mining GDP. These were all expressed as a percentage of total GDP at current prices. The variables used in the MIMIC model were transformed to be normally distributed and stationary for better model fitting.

The initial data set included other causal variables such as Consumer price index (CPI), social benefits, total government subsidies, government employment, finance GDP, agriculture GDP and interest rate. However, at 5% level of significance, the model considered those variables insignificant in explaining the SE activities.

4.2.1. Fitted MIMIC Model

Table 3 below shows the accepted MIMIC model with three indicators namely, nominal GDP, Currency and LFR and five causal variables namely, TAXB, HHI, HHD,UNEMP, HHC and the sector mining GDP, all expressed as a percentage of nominal GDP.

The indicators GDP, Currency and LFR from the measurement model are highly significant i.e. less than 1% level of significance. Both Currency and LFR have positive coefficients, indicating a direct relationship with the SE. However, GDP shows an opposite relationship with the SE.

The structural equation model, which links the causal variables to the SE, indicates that TAXB, HHI, HHC and the interaction of Unemployment & HHD (UNEMP*HHD) are positively related to the SE activities, implying

that an increase in any of those causal variables will results in an increase in the SE activities.

The mining sector was highly significant in explaining some movements in the SE activities. When the mining sector GDP decreases, the SE activities increases (an indirect relationship). This shows the uneven share of resources from mining sector and the informal sector in South Africa.

Table 3. *MIMIC model results*

Measurement Model				
	Estimate	Std.Err	z-value	P(>|z|)
n =~				
LFr	0.201	0.060	3.316	0.001
C_CGDP	0.489	0.075	6.556	0.000
CGDP	-0.062	0.010	-6.357	0.000
Structural Equation Model				
	Estimate	Std.Err	z-value	P(>|z|)
n ~				
TAXB	0.405	0.123	3.297	0.001
HHI_CGDP	0.160	0.079	2.029	0.042
HHD*UNEMP	0.034	0.010	3.564	0.000
HHC_CGDP	0.333	0.083	4.022	0.000
MINGDP_CGDP	-0.916	0.287	-3.195	0.001
Statistical Test				
Root Mean Square Error of Approximation (RMSEA)				0.044
Standardized Root Mean Square Residual (SRMR)				0.022
Comparative Fit Index (CFI)				0.995
Tucker-Lewis Index (TLI)				0.986

Notes: n: Latent variable / Shadow economy. LFR: Labour force participation rate. CURRENCY: Currency in circulation per current GDP. CGDP: Nominal gross domestic product. MINGDP: Mining GDP. TAXB: Tax per GDP (Tax burden). UNEMP: Unemployment per labour force. Household Income (HHI), Household debt (HHD) and Household consumption (HHC) are as % of GDP

Source: Authors computation

4.2.2. MIMIC model evaluation

Several statistical tests exist for selecting the best fit model(s). The following statistics were used: Root Mean Squared Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) statistics to select the best fitting MIMIC model as presented in table 4.3.

The value of the RMSEA in Table 4.3 is 0.044 with a p-value of 0.449 ($p > 0.05$) signifying that the model of choice has a closer fit. The SRMR of 0.022 was observed, this statistic works similar to RMSEA but on the standardized data. CFI computed by the model was 0.995 greater than the acceptable value of 0.9. Furthermore, the TLI of 0.986, which is greater than the acceptable value of 0.9, was observed from the model fitted. The TLI statistic is more restrictive than the CFI. Based on the fitted MIMIC model, the SE estimates where derived as a percentage of nominal GDP and are presented in the next section.

4.2.3. MIMIC shadow economy estimates

The MIMIC model depends on some pre-determined SE estimate for the base year from another model. The base year of 2010 was chosen to align with the base year of the observed causal variables from STATSSA. This paper uses the predetermined own computation estimate of 23.6% for the base year from the CDA in section 4.1 above. The South African SE estimates are shown in Figure 3 below.

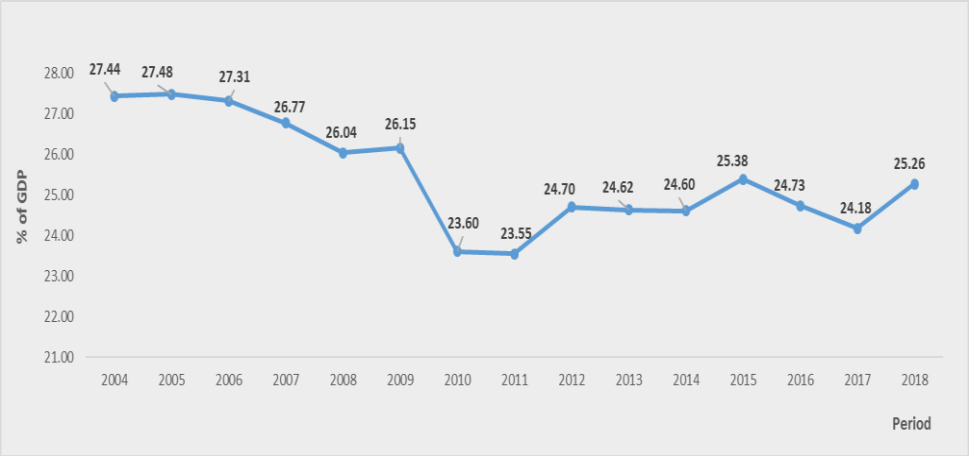


Figure 3. *The size of the SE in RSA as % of GDP – MIMIC Model*
Source: Authors computation

Using the 2010 base estimate of 23.6% from CDA model, the MIMIC model estimates the South African SE to be 25.45% on average for the period 2004 to 2018, with the minimum value of 23.55% in 2011 and maximum value of 27.48% in 2005. A gradual overall decreasing trend in SE is observed from the fitted MIMIC model.

5. Discussion

This study uses the scientific "indirect" methods, which are CDA and the MIMIC model to estimate the size of the SE in South Africa. The two models estimate the SE using economic indicators at macro-level and are viewed to be superior when compared to direct methods such as surveys and tax auditing (Giles & Tedds, 2002; Hassan & Schneider, 2016).

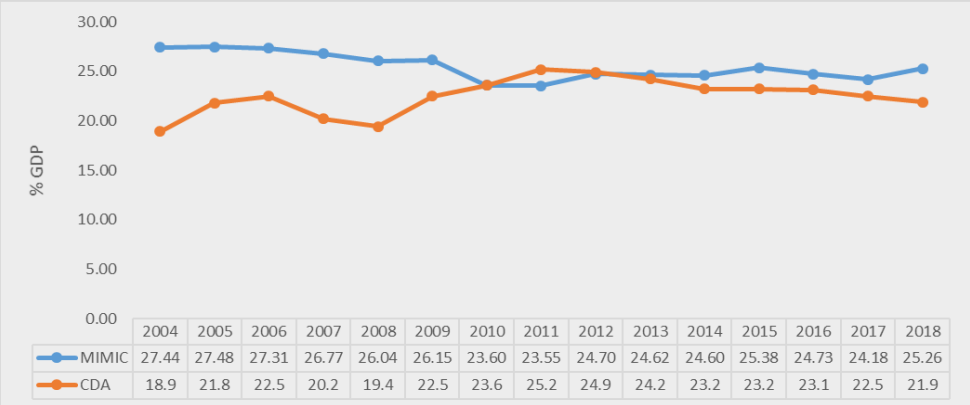
The sample data from 2000Q01 to 2019Q03 was initially used for fitting the two models. However, due to lack of complete data in other variables and the non-normal nature of some variables (even after transformation), the estimates were only generated from 2004 to 2018.

The MIMIC model was evaluated using the commonly used statistics, Root Mean Squared Error and the Comparative Fit Index CFI (with the benchmark value of 0.9). The value of the RMSEA was 0.044 significant at 5% level of significance. The CFI computed by the model was 0.995, greater than the acceptable value of 0.9. Based on the fitted MIMIC model, the SE estimates where derived as a percentage of nominal GDP.

Similarly, the CDA model residuals were evaluated to check for autocorrelation and normality. The probability values indicates the absence

of autocorrelation and normal distribution of the residuals at 5% level of significance.

Figure 4 below compares the size of SE estimated from MIMIC and CDA.



	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
MIMIC	27.44	27.48	27.31	26.77	26.04	26.15	23.60	23.55	24.70	24.62	24.60	25.38	24.73	24.18	25.26
CDA	18.9	21.8	22.5	20.2	19.4	22.5	23.6	25.2	24.9	24.2	23.2	23.2	23.1	22.5	21.9

Figure 4. *MIMIC vs. CDA estimates comparisons*
Source: Authors computation

The average estimated size of SE from the CDA and MIMIC model for the period 2004 to 2018 are 22.47% and 25.45% respectively. This is a difference of 2.98% between the two models, thus the models discrepancies are minimal on average. The estimated range was 23.55% to 27.48% for MIMIC and for the CDA estimates were between 18.90% and 25.2%. Therefore, the overall estimates for the CDA models were a bit lower than those from the MIMIC model.

5. Conclusion

The CDA and the MIMIC models were derived and evaluated for better model fitting to estimate the South African shadow economy. According to the CDA, on average the SE accounts for 22.47% of the formal economy and according to the MIMIC model, the SE accounts for 25.45% on average. The discrepancy between the MIMIC and CDA model was around 2.98% on average for the period 2004 to 2018. Overall, both the models show a slight decreasing trend for the period.

As can be observed from Table 4 below, all the hypotheses were confirmed by both the CDA and the MIMIC model with the exception of the deposit interest rate. In South Africa's case, CDA model suggests that individuals will transact in cash despite the increase in deposit interest rate. Therefore the increase in deposit interest rate does not serve as a motivation to stop engaging in the SE activities.

Regulation could result in either a positive or a negative relationship with shadow economy. In our case, the model suggests that the growth in the government sector will result in more activities in the shadow economy due to government employees being susceptible to bribery and corruption. GDP could also have a positive or negative relationship with the shadow economy. The model results confirms Schneider (2005)'s assertion that the

relationship is negative for developing countries. Similar results were obtained for mining sector GDP. The relationship between unemployment and shadow economy was insignificant. The relationship became positive and significant, only when unemployment interact with HHD.

Table 4. *Empirical confirmation of the hypotheses*

Variables (Hypothesized sign)	Method	Result
1. Total tax burden (+)	CDA MIMIC	Both Confirmed
2. GDP (-)	CDA MIMIC	Both Confirmed
3. Self-Employment (+)	CDA	Confirmed
4. Regulation (+)	CDA	Confirmed
5. Mining sector GDP (-)	MIMIC	Confirmed
6. Labour force participation rate (+)	MIMIC	Confirmed
7. Currency (+)	MIMIC	Confirmed
8. Household debt (+)	MIMIC	Confirmed
9. Unemployment (+)	MIMIC	Confirmed

The results from this study are just the first step of understanding the overall SE activities and could be used by the government authorities of South Africa for decision making on a high level. However, there is a need for further analysis to be done to explore the SE activity distribution between different economic sectors in order to influence future enforcement plans and undertakings by government authorities; resulting in an easier way to identify, locate and monitor unrecorded businesses and also maximise revenue collections and minimise non-compliance for different sectors.

References

Ahumada, H., Alvaredo, F., & CanaveseA. (2007). The monetary method and the size of the shadow economy: A critical assessment. *Review of Income and Wealth*, 53(2), 363-371. doi. 10.1111/j.1475-4991.2007.00234.x

Amoh, J., & Adafula, B. (2019). An estimation of the underground economy and tax evasion: Empirical analysis from an emerging economy. *Journal of Money Laundering Control*. doi. 10.1108/JMLC-01-2019-0002

Ariyo, A., & Beloe, W. (2012). Currency demand, the underground economy and tax evasion: The case of Nigeria. *Journal of Monetary and Economic Integration*, 11(2), 129-157.

Brand South Africa. (2018). Accessed, 12 February 2020. [Retrieved from].

Buehn, A., & Farzanegan, M.R. (2013). Impact of education on the shadow economy: Institutions matter. *Economic Bulletin*, 33(3), 2052-2063.

BusinessTech. (2020). Accessed 12 February 2020. Available from. [Retrieved from].

Cagan, P. (1958). The demand for currency relative to the total money supply. *Journal of Political Economy*, 66(4), 302-328. doi. 10.1086/258056

Dell'Anno, R., Gomez-Antonio, M., & Pardo, A. (2007). The shadow economy in three Mediterranean countries: France, Spain and Greece. A mimic approach. *Empirical Economics*, 33, 51-84. doi. 10.1007/s00181-007-0138-1

Enste, D.H. (2005). The shadow economy in OECD and EU accession countries – Empirical evidence for the influence of institutions, liberalization, taxation and regulation. *in* C. Bajada, F. Schneider (Eds.), *Size, Causes and Consequences of the Underground Economy – An International Perspective*, Burlington: Ashgate Publishing Company.

Frontier Economics. (2012). *The Impact of Regulation on Growth. A Report Prepared for the Department of Business, Innovation and Skills*. Frontier Economics Ltd, London.

Farzanegan, M.R. (2009). Illegal trade in the Iranian economy: Evidence from a structural model. *European Journal of Political Economy*, 25(4), 489-507. doi. 10.1016/j.ejpoleco.2009.02.008

Gana, K., & Broc, G. (2019). *Structural Equation Modeling with Lavaan*. London: Wiley.

Giles, D.E.A., & Tedds, L.M. (2002). Taxes and the Canadian underground economy. *Canadian Tax Paper*, No.106.

Hall, A. (2019). Underground economy. Chicago: Encyclopaedia Britannica. [Retrieved from].

Hassan, M., & Schneider, F. (2016). Modelling the Egyptian shadow economy: A mimic model and a currency demand approach. *Journal of Economics and Political Economy*, 3(2), 310-339. doi. 10.1453/jepe.v3i2.788

Iol. (2019). SARS probes illicit economy. Accessed 03 February 2020. [Retrieved from].

Kaufmann, D., & Kaliberda A. (1996). Integrating the unofficial economy into the dynamics of post socialist economies: a framework of analyses and evidence, *in* B. Kaminski, (Ed.). *Economic Transition in Russia and the New States of Eurasia*. M.E. Sharpe, London.

Klaric, V. (2010). Estimating the size of non-observed economy in Croatia using the MIMIC approach. *Financial Theory and Practice*, 35(1), 59-90.

Koekemoer, P. (2018). The trend of self-employment. Accessed 14 February 2020. [Retrieved from].

Lackó, M. (1999). Electricity intensity and the unrecorded economy in post-socialist countries. *in* E. Feige, & K. Ott (Eds.). *Underground Economies in Transition*, Ashgate Publishing Company.

Lackó, M. (2000a). Do power consumption data tell the story? Electricity intensity and hidden economy in post-socialist countries, *in* E. Maskin, & A. Simonovits (Eds.). *Planning, Shortage and Transformation: Essays in Honor of Janos Kornai*, Cambridge: The MIT Press.

Lackó, M. (2000b). Hidden economy – An unknown quantity? Comparative analysis of hidden economies in transition economies, *Economics of Transition*, 8(1), 117–149. doi. 10.1111/1468-0351.00038

McMillan, J. (2006). Below the radar: Underground markets for the poor. *Harvard International Review*, 27(4), 46-50.

Medina, L., & Schneider, F. (2019). Shedding light on the shadow economy: A global database and the interaction with the official one. *CESifo Working Paper*, No.7981. [Retrieved from].

Medina, L., & Schneider, F. (2018). Shadow economies around the World: What did we learn over the last 20 years? *IMF Working Papers*, No.WP/18/17. [Retrieved from].

Minerals Council South Africa. (2020). Accessed: 12 February 2020. [Retrieved from].

SA Mine. (2019). Accessed 10 February 2020. [Retrieved from].

Sharapenko, D. (2009). Estimation of the shadow economy in Russia, A study conducted in partial fulfilment of the requirements for the degree of Masters of Arts. Central European University. Budapest, Hungary.

Schneider, F. (2005). Shadow economies around the world: What do we really know? *European Journal of Political Economy*, 21(30), 598-642. doi.10.1016/j.ejpoleco.2004.10.002

Schneider, F., & Buehn, A. (2016). Estimating the size of the shadow economy: methods, problems and open questions. *IZA Discussion Paper*, No.9820. [Retrieved from].

Schneider, F., & Enste, D.H. (2013). *The Shadow Economy. An International Survey: 2nd Edition.* Cambridge University Press.

Schneider, F., & Hametner, B. (2007). The shadow economy in Colombia: Size and effects on economic growth. Johannes KeplerUniversity of Linz, Department of Economics *Working Paper*, No.0703. [Retrieved from].

Schneider, F., Chaudhuri, K., & Chatterjee, S. (2003). The size and development of the Indian shadow economy and a comparison with other 18 Asian countries: An empirical investigation, Johannes Kepler U., Department of Economics, *Working Paper*, No.0302. [Retrieved from].

Schneider, F., Buehn, A., & Montenegro, C.E. (2010). New estimates for the shadow economies all over the world. *International Economic Journal*, 24(4), 443-461. doi. 10.1080/10168737.2010.525974

Schneider, F., & Enste, D.H. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, 38(1), 77-114. doi. 10.1257/jel.38.1.77

The Global Corruption Barometer Africa. (2019). Accessed 10 February 2020. [Retrieved from].

The New York Times. (2018). Accessed: 05 February 2020. [Retrieved from].

UNCTAD. (2016). Accessed: 15 December 2019. [Retrieved from].