# Modeling the Shadow Economy and Its Dynamics
# In Case of Pakistan

**Submitted by**

**Sumeet Ashok**

**Co-Author:** **Dr. Ahsan Ul Haq**

**Dr. Khalid Mehmood**

# ABSTRACT

In this study we discuss the detailed empirical and theoretical concepts which drive the shadow activities in the official economy. The thesis contains the joint modified approach used for the modelling of shadow economy of Pakistan, we have employed monetary approach and Multiple Causes and Multiple Indicators (MIMIC) model estimated by Structural equation modeling technique to estimate the size and dynamics of the underground activities. Currency demand model is estimated via ARDL approach; this method provides the point estimator of the underground economy in a base year then this point estimator will be incorporated in MIMIC model to get proper scale and for calibrations and benchmarking of shadow economy which prevails in the official economy and hidden from the authorities.

This study contributed to the existing literature which incorporated time series analysis of MIMIC model, as some causes and indicators don't fulfill the property of stationarities. We have estimated the long-run MIMIC model and Short-run MIMIC model including the Error Correction terms (ECM) after checking the cointegration relationship by Engle and Granger approach for the first time in Pakistan. The magnitude of the shadow economy is calculated by the best available econometrical tools which are available. The dimension of shadow economy varies from 50% in 1974 and 28% in 2015. The unusual decreasing trend in shadow economy can be witnessed in 1981-1985 and 1991-1995 due to declining in demand for cash, unemployment rate, and tax burden at the same time banking or financial sector development can be observed.



**Key Words:** Shadow Economy, Structural Equation Modelling, ARDL, MIMIC Model

## Introduction

This paper is an attempt to improve the econometrical methodology, multiple indicators multiple causes (MIMIC) model with the aim to know the dynamics, the size and the theoretical development of the shadow economy by incorporating the dynamic properties of the causes and indicators. So, this study will contribute to the existing literature by analyzing the cointegration relationship between the causes and indicators of the shadow economy considering their long and short run relationships for the first time in case of Pakistan.

The first objective of the study is to discuss the theoretical and empirical aspect of the shadow economy and highlight the contemporaneous methodologies that are used for the purpose. The second objective of the study is to capture the potential causes and indicators of the underground economy for Pakistan and then estimate the index of shadow economy through Multiple Causes and Multiple Indicator (MIMIC) model. We also analyze the short run and long run effect of the causes and indicators on shadow economy. In order to estimate the point estimator of the shadow economy monetary approach is employed that will be used for the benchmarking in the former model. The third objective of the study is to measure the size and dynamics of shadow economy for the time period (1972-2015) in a case study of Pakistan.

Shadow Economy (SE) refers to those activities which cannot be directly observable so its magnitude should be estimated. This phenomenon can be attributed with many names: "hidden, underground, informal, irregular, unofficial, parallel, invisible, second, subterranean and unrecorded economy" (Frey and Schneider, 2000, pp. 1-2). In this study, the term 'shadow' and 'underground' will be imparting the same meaning unless stated otherwise.

A national economy cannot be effectively manageable without knowledge of the magnitude of the economic activities that are running beside the official economy. Thus, shadow economy can be considered as the real parameter of the national economy. Shadow economy's existence influences the socio-economic perspective of the society including formation and redistribution of income, investment, trade, inflation, tax system, economic growth in general and most of our macroeconomics indicator. This issue is very important in developed and developing countries. Shadow Economy is considered a relevant indicator of conflict between individual and state (Schneider, 2005).

Although historically empirical estimation of shadow economy seems to be a problem because of varied enormously in terms of the methodology employed. For the last two decades' economists, have dedicated themselves to address this econometrics problem. In this regard economists, have developed the latest and robust techniques to get a relevant estimation of shadow economy for several countries, two points are to be noted here. Firstly, until now no unified approach has been given for the estimation of the shadow economy, every approach has some strong and weak points. Discussion and critiques are still going on. Secondly, estimated results of shadow economy for same countries and for same time periods often do not show consistency, so there is a prevalent confusion in understanding the dynamics of the shadow economy.

The study of Frey and Schneider (2000) elaborates some issues that can come up as a result of simple unawareness and incorrect estimation of the shadow economy. The first problem for the underestimation of the shadow economy is that it undermines the actual economic growth which leads the government to intervene and stimulate the economic growth by increasing government expenditure and monetary aggregates; when there is no need to take such measures. Such measures give rise to inflation which can be dangerous for other macroeconomic indicators as we have witnessed in mid-90's in European countries (Schneider,2000). Secondly, erroneous estimation of the underground economy may accentuate the other economic problem which is unemployment. As most of the labor force is enrolled in the shadow economy. Government intervenes by increasing their expenditure in order to create workplaces for the unemployed workers leading to excessive and inefficient social policies. Thirdly, incorrect estimation of the underground economy leads to underrate the GDP for not counting the goods and services in the shadow sector, this will mislead monetary policy of a country. The fourth problem that can arise is revenue of government is lost due to tax evasion, causing errors in budget accounting. In short economic condition of a country as whole are evaluated in a biased way.

The shadow economy up to now is a controversial issue. Economist faces the problem to estimate and defining the shadow economy (Schneider and Enste (2000)). How shadow economy can define? Three ways to define the shadow economy, (1) those economic activities that contradict to the contemporaneous legislation, this include the aggregate of the illegal activities, that feeds the crimes at a different level (Popov, 1999). (2) The system of production and redistribution which must be incorporated in the national product but are not taken in into account by national statistics officers and

are uncontrolled by the societal system (Frey and Schneider, (2000). Due to quantitative in nature make it best for economic purpose, moreover, these activities are not part of GNP like household activities. Third explanation is (3) all activities which are formed to satisfy the extraneous human needs and nurturing immoralities in human (Popov, 1999).

These all above definitions are found in literature and they cover the different aspect of shadow economy which prevails in the national economy, thus dividing SE into three large blocks, where the first one covers the unofficial economy. It comprises of the legal economic activity which is not considered by official statistics or hidden from the tax system of a country. The second block incorporates the fictitious economic activities, this includes the modified records, bribery, speculative transactions, theft and different frauds for giving and taking money. And the third block comprises of all activities which are prohibited by law basically these are illegal economic activities.

## Review of Literature

In daily routine, many individuals around this global world involve themselves in underground economic activities. The involvement of these people in underground economy is due to the weak regulation system of the government and in order to avoid taxes that is levied on them and save money as services and products are cheaper. The historical literature demonstrates that the rise in shadow economy was in a peak in 1970's. When a neoclassical school of thought immerged, in which government intervention is justified only through taxes in order to finance the public spending programs (Tanzi and Schuknecht, 1999). In the 1980's problems of shadow economy were in boom around the many countries, especially in the OECD, however in 1990's shadow economy got great attention due to dramatically increase in unemployment in the European Union. The discussion on the shadow economy does not go into detail in the scientific media and in newsprint, however, it is mostly discussed on the judgmental basis. Therefore, estimation of shadow economy become necessary to provide good policy for the country.

In recent studies, economists have shown keen interest to measure the gap between the observable and actual. Estimation of shadow economy get great attention in this regards, with the development of new econometric software different estimation techniques have been introduced to estimate the shadow economy. The direct

approaches are microeconomic approaches based on surveys and tax auditing. However, direct approaches can be appropriate as it provides detailed and comprehensive information about the structure or base of shadow economy but due to the difficulty in conducting surveys and sensitivity of questionnaire it is not possible. However, interviewees will hesitate to confess for their involvement in illicit work (Schneider and Enste, 2000). Due to the limitation of the direct approaches or microeconomic approaches, macroeconomic approaches evolved. Macroeconomic approaches are also known as indirect approaches, there are four indirect indicators to estimate the shadow economy: a) The difference between national income and nation expenditure, b) The difference between official labor force and actual labor force, c) Monetary approach and d) The physical input method. These all macroeconomic approaches are designed to estimate the shadow economy but with certain limitation, in literature, all above approaches are criticized on the basis of assumptions that we consider before estimating the shadow economy. Mostly researcher argues the result we get are due to these assumptions and another critique of this indirect approach is that they consider only one indicator for shadow economy instead of criticism still all these approaches are used to estimate the shadow economy in developed and developing countries.

After these, all criticism new technique has been developed to estimate shadow economy. Structural equation modeling (SEM) approach covers the above limitations, multiple indicators and multiple causes (MIMIC) model mostly used to observe the unofficial economy. SEM approach first employed in 1984 to estimate hidden economy for OECD countries by Frey and Week-Hannemann. In the model approach, we explicitly take multiple causes that can determine the existence and growth of shadow economy and its multiple effects on the overall economy. The empirical method used in MIMIC model is far different from the methods discussed above. In this latent variable cannot be measured directly but we analyze the covariance structure between observable variables (causes and indicator) later it provides the evidence about the relationship between causes and indicators and a latent variable.

SEM approach or MIMIC model replacing existing conventional micro (direct) and (macro) indirect approaches, its main advantage is the separation of causes and indicators and it also covers the limitation of indirect approach in which it was argued that this approach has one cause. However most notable while estimating shadow economy through SEM approach we consider its various causes. SEM approach also

have been criticized by (T. Breusch, 2005) up to extreme level that it is not suitable for estimating the shadow economy, he argues that MIMIC model approach is not suitable for the purpose and it is criticized on the ground that estimated results are dependent on data transformation, unit of measurement and the sample used. Immediate response has been given to Breusch in 2006 by Dell'Anno and Schneider, they empirically prove that still, MIMIC model is a most suitable technique to estimate the shadow economy. In most recent studies MIMIC modeling is employed for the purpose due to its thorough elaboration many causes and effects of the shadow economy. The table given below shows the pros and cons of different estimation methods used for the estimation of shadow economy.

**Table 1. 1: Estimation Approaches**

| Approaches and Comments | Method | Strong and Weak Points |
|---|---|---|
| **Direct**<br>- Only point estimates<br>- downward Biased<br>+ very detailed | Survey | Costly |
| | Actual income Minus income for Tax | - Only Taxable Activities |
| **Indirect or Discrepancy**<br>- imprecise due to mixed nature of factors<br>- Need a base year<br>- Under Mined bias | Income - Expenditures | + Various Levels<br>- Measurement error |
| | Labor Market | + Not consider employed in both sectors |
| | Transactions | - Having to assume constant ratio of transaction to GNP |
| | Currency - Demanded | + Elegant, data present<br>- only cash |
| | Cash Deposit ratio | - very simple<br>- not technical progress |
| | Physical inputs (Electricity use) | + Easy Data<br>- strong assumptions<br>- not technical progress |
| **Model**<br>+ Precise | MIMIC | + Comprehensive<br>- requires lots of data |

For the first time MIMIC (multiple indicator and multiple causes) that is model approach used as tool to estimate shadow economy by Frey and Weck-Hennemann (1984), he estimated the shadow economy for seventeen OECD countries for time span of 1960 to 1978. As this estimation technique is ingenious it incorporates more than one causes. In order to get the index of latent variable put the coefficient of causes and indicator variables in the estimated equation. They interpreted the results as a percentage of GNP and also explored high growth rate in shadow economy among these countries. Helberger and Knepel (1988) criticized the results of above study. They emphasis that structural modeling approach is one of the best approach to obtain the index of latent variable like shadow economy. However, the problem with results of Frey and Weck-Hennemann (1983) study is due to the variables which they incorporated as causal variables for shadow economy. Further the idea of pioneer study was extended by Schneider et al (2002) with econometrics modification to incorporate time series properties of data and allow lag adjustment in dynamic MIMIC model and estimated the shadow economy for United States.

The three studies got great attention in the robust field of study to estimate the shadow economy. i) Giles (1999) who estimated shadow economy for New Zealand, ii) Dell'Anno and Schneider (2003) estimated the shadow economy for Italy and iii) Bajada and Schneider (2005) they estimate the shadow economy for Australia. These all above studies used structural equation modeling approach to estimate the shadow economy. The study one further modified the MIMIC model approach and take into account the time series property of data, applied unit root and cointegration analysis for the estimation of shadow economy. The second study discusses the advantage and disadvantage of the SEM approach. Following are the main advantages a) show relationship between observable/manifest variable and unobserved/latent variable b) takes in to account the nonlinear properties of the data c) give detail symmetry and asymmetry information regarding the data and d) it can also consider the time series properties which makes SEM approach best statistical tool for economic research. However, the main disadvantage of this approach we are unable to give the proper meaning to unobserved variable because MIMIC model follows the confirmatory analysis rather than exploratory analysis. Last study gives the detailed estimation of shadow economy. All above studies are criticized by T.Breusch (2005) on the basis of divergence in studies. He explored that there is problem in interpretation of latent variable and approaches to calibration. Breusch shows the results were very sensitive

to the transformation of data on the basis of these three problem, he argue that MIMIC modeling is not suitable for the estimation of shadow economy. Quick response was given to Breusch after a year, Dell'Anno and Schneider (2006) show that still MIMIC modeling is best for the purpose.

Dell'Anno (2003) estimated the shadow economy for Italy. In order to carry out the econometrics analysis he used the multiple indicator and multiple causes (MIMIC) model approach. While estimating the shadow economy, determinant or causes of SE are different from the study of Portugal shadow economy. In this study he explored the government consumption/GDP, index of efficient justice and index of illegality are determinant of shadow economy. Indicators consist of, real GDP and currency outside the bank while estimating the Portugal SE he used variables social benefits paid by government. Government employment in labor force and subsidies as causes. The results of this study shows variation is shadow economy of Italy is 1 to 34 percent. He shows that model approach usually has the problem of indefinite matrix problem or non-positive definite matrix problem. In order to tackle this problem, he demonstrates that Monte Carlo simulation is used. This study also emphasis that MIMIC model approach have some weaknesses but still it best method to estimate the shadow economy in the field.

Dell'Anno (2007) finds out the statistical relationship between shadow economy and other macroeconomics. He estimated the shadow economy for Portugal using time series data from 1977 to 2004, for econometric analysis, they used MIMIC model approach (multiple indicator and multiple causes). The main causes of SE are as a) government employment in labor force, this variable represents the both economic freedom and an index of over burden of the public sector in the economy. B) Tax burden it is the most important determinant of SE economy; it is hypothesized more the tax burden there will be more incentive to work in the shadow economy. C) Subsidies are the payment made by government to protect the domestic industries in order to protect them. D) Social benefits paid by government in this variable current transfers are included which are received by the households for certain events and circumstances like unemployment, retirement and sickness. Social benefits have conflicting relationship with shadow economy likewise subsidies. E) Self-employment this variable considers to be main cause of shadow economy as rate of self-employment that is percentage of labor force is taken as determinant of shadow economy. The last variable F) Unemployment rate. In this study he considers the two

indicators of the shadow economy i) real gross domestic product index and ii) labor force participation rate. The result of this study shows magnitude of shadow economy ranges from 29.6 percent to 17.6 percent of the official GDP.

**Recent Studies on Shadow Economy in Pakistan**

The most recent empirical study on the estimation of shadow economy in case of Pakistan by Arby, Malik and Hanif in 2010 by using monetary approach, electric consumption approach and MIMIC (multiple indicator and multiple causes). They used time series data from 1966 to 2008 and calculated the magnitude of shadow economy. This is pioneer paper in case of Pakistan to know the magnitude and estimates the underground economy by using electric consumption approach and MIMIC approach. While using monetary approach they have addressed the stationarity and used the ARDL. MIMIC model consists of three main causes of shadow economy it includes ratio tax to GDP, ratio of M2 to GDP and the regime durability and indicators of shadow economy are currency in circulation and electric consumption growth. The results show 30 percent of the total economy is informal economy which is considerably decreased to 20 percent in 2000s. The result of this study are close to the result of previous studies on shadow economy for Pakistan. Kemal (2007) measure the shadow economy by using K & Q approach, this is basically discrepancy approach basically they have calculated the total consumption in private sectors, from the house hold survey for the population then it is adjusted for net trade and calculate the true estimate of GDP, which is compared to the GDP of National Income Account. The difference between these two GDP is equal to Shadow economy. This study shows that magnitude of shadow economy is rising till 1990's. However, the study of Ahmed and Haider (2006) shows magnitude of shadow economy is decreasing trend.

Ahmad et al. (1995) estimate the underground economy of Pakistan through the currency or monetary approach for the period 1960-1990. They founded that black economy as a percent of GDP has shown fluctuating trend and tax evasion has increased over the number of years, but the black economy has registered as decline. In addition, they estimate the revenue less due to presence of black economy. They found that 40 Rs billion to 45 billion loss in 1989-90.

Zafar Iqbal et al., (1999) estimated the black economy and tax evasion in the different sectors of economy over the period 1973-96. They found that the size of overall underground economy is remarkably increased form 15 billion in 1973 to 115 billion

in 1996. The total tax evasion in 1973 was 1.5 billion, which has peak to 152 million in 1996. Furthermore, the various sector of underground economy shown that income from underground economic activities in the external sector is higher than the other domestic sector and non-tax payers sector. In addition, they concluded that the loss from revenues of taxes and expenditure on public services are the contributing factor of higher budget deficit, uncertainty cost of doing business.

M. Ali Kemal (2003) estimated the size of underground economy and tax evasion by using the method of Tanzi and Feige methodologies, which are based on the direct and indirect method. He observed that the magnitude of shadow economy in 1974 is 25.76, which is increased to 35.28 percent in 1990. In 1998 the size of underground economy is 70.92 percent and it deceased to 25.76 percent in 2002. Tax evasion was estimated 2.74 percent in 1973, 4.73 percent in 1990 and 9.40 percent in 1998. It declines 5.99 percent in 2002. Furthermore, he argued that the size of the underground economy changed with the change variables or benchmark and time period due to the change in the magnitude of the parameters. He concluded that the good governance system may help in reducing magnitude of the black economy in Pakistan.

## METHODOLOGY AND DATA

### Currency Demand Model and Data

The methodology which I have selected to estimate the magnitude of shadow economy is modified form Vito Tanzi (983). This approach consists of specifying a demand for currency equation to be used to derive the effect of a change in the tax level on the demand. Tanzi used the following currency demand function to know the dynamics of the SE:

$$\ln\left(\frac{CM}{M_2}\right)_t = \alpha_0 + \alpha_1 \ln(1+TW)_t + \alpha_2 \ln(\frac{WS}{Y})_t + \alpha_3 lnR_t + \alpha_4 \ln(\frac{Y}{N})_t + U_t$$

Where (ln) represents the natural logarithms,

$CM/M_2$ represents the ratio of current holdings to M2 and deposits accounts, TW stands for weighted tax rate, WS/Y is the ratio of wages to nation income, R is the interest rate use to capture the opportunity cost of holding currency, and Y/N is the per capita income.

The above model has two main econometrical problems, firstly, it is estimated using least square method (OLS) without taking into account the time series properties of the data. Secondly, dependent variable is in log form this has the problem of disaggregation, (Arby, Malik and Hanif 2010).

In this study we use the modified form currency demand approach that have been employed in the study of Arby at el. (2008). This address the problems with Tanzi's model, and allow us to use both stationary and non-stationary variables in one model. For this purpose, we employ autoregressive distributive lag (ARDL) model as suggested by Pesaran et al (2001). As ARDL estimators are super consistent and give valid inferences about long-run parameters. We apply the following model to establish the long run cointegration relationship between the currency and other related dependent variables.

$$\Delta CM_t = \alpha_0 + \beta_1 CM_{t-1} + \beta_2 T.GDP_{t-1} + \beta_3 R_{t-1} + \beta_4 INF_{t-1} + \beta_5 F_{t-1} + \beta_6 GGDP_{t-1} + \beta_7 GYD_{t-1} + \sigma_1 \Delta TGDP_t + \sigma_2 \Delta F_t + \sigma_3 \Delta INF_t + \sigma_4 \Delta R_t + \sigma_5 \Delta GGDP_{t-1} + \sigma_6 \Delta GYD_{t-1} + \epsilon_t \dots\dots (\mathbf{1})$$

Where $CM_t$ represents the dependent variable in our case currency in circulation to M2 ratio.

i) $T.GDP_t$ ratio of total taxes to GDP, represents the tax burden an increase in the tax burden raise the relative price of taxable versus nontaxable economic activities. So increase in taxes will increase the shadow or underground activities thus demand for currency increases this implies that tax to GDP ratio affect currency ratio positively, Tanzi (1983).

ii) The financial sector development $F_t$, improvement in the financial sector will lower the demand for currency for the transaction purposes. Indicator of financial development is represented by the ratio of total demand and total time liabilities to nominal GDP.

iii) Market discount rate $R_t$ is used as a proxy for the interest rate, as rise in interest rate will increase the opportunity cost holding the currency, demand for currency decreases and demand for deposits increases Tanzi (1983).

iv) Growth rate of GDP deflator is used as proxy for inflation rate. Demand for currency is highly effected by the inflation. As inflation rises people need more money to buy goods and services.

v) Growth rate of real per capita GDP is used as measure of economic development of a country. Theoretically negative relationship between economic development and currency demand.

**Table 1. 2: Description of Variables and Source of Data**

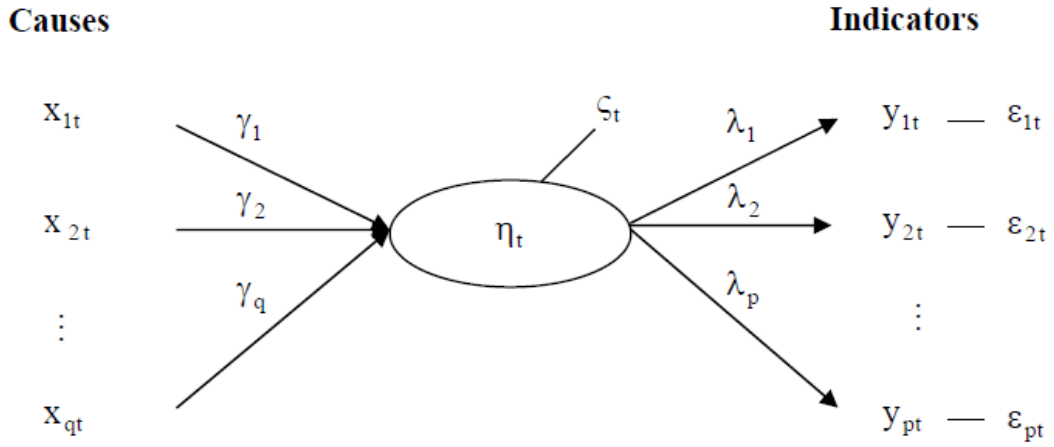
| Variable | Symbol | Description | Source | Annotation |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| cash holdings to currency and deposits accounts | $\boldsymbol{CM_t}$ | Ratio of currency in circulation to M2 | Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Tax | $(\boldsymbol{T.GDP})_t$ | ratio of total taxes to nominal GDP | Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Financial sector development | $(\boldsymbol{F})_t$ | Ratio of total demand and time liabilities to nominal GDP. | Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Interest rate | $\boldsymbol{R_t}$ | Discount rate | IFs CD-ROM/ SBP statistical Bulletins | Yearly Data (1972-2015) |
| Inflation rate | $\boldsymbol{INF_t}$ | GDP Deflator | Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Growth rate of real GDP | $(\boldsymbol{GGDP})_t$ | Ratio of real GDP to population in difference | Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Growth rate of disposable income | $(\boldsymbol{GYD})_t$ | Rate of change of disposable income | Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |

**DY (MIMIC) and EMIMIC Model**

Multiple Causes and Multiple Indicators (MIMIC) model consist of both observable and unobservable variables. In order to estimate the magnitude of unobservable variable structural equation modeling technique is used. To find out the magnitude of the coefficient it basically minimizes the discrepancies between the observed covariance matrix and sample covariance matrix. MIMIC model consist of two parts; the structural model equation and measurement equation model. As shown in below figure

**Figure 1. 1: MIMIC Model**




The structural equation model is represented by following equation:

$$\eta_t = \boldsymbol{\gamma' X_t} + \varsigma_t \ \ldots\ldots\ldots\ldots (2)$$

Where $X_t = (X_{1t}, X_{2t}, \ldots\ldots\ldots, X_{qt})$ is (1xq) vector of potential cause of latent variable $\eta_t$ overtime as indicated by subscript t. The coefficient of the structural equation part is represented by: $\gamma' = (\gamma_1, \gamma_2, \ldots\ldots\ldots, \gamma_q)$ a vector (1xq) that gives the description of the causal relationship between unobservable variable $\eta_t$ and its causes. Unexplained component of the structural model is represented by error term $\varsigma_t$. MIMIC model assumes $\mathbf{E}(\eta_t) = \mathbf{E}(X_t) = \mathbf{E}(\varsigma_t) = 0$, this implies that causes variables are measured as deviation from its mean and $\mathbf{E}(X_t\ \varsigma_t') = \mathbf{E}(\varsigma_t\ X_t') = 0$ , this means there is no correlation between error term and causal variables.

The measurement model represents the latent/unobservable variable in terms of observable variable and it is given by following equation:

$$y_t = \lambda' \eta_t + \varepsilon_t \ \ldots\ldots\ldots\ldots (3)$$

Where $y_t = (y_{1t}, y_{2t}, \ldots\ldots\ldots, y_{pt})$ is (1xp) vector of indicators of the latent variable $\eta_t$ overtime as indicated by subscript t. The regression coefficient of the measurement model is represented by: $\lambda' = (\lambda_1, \lambda_2, \ldots\ldots\ldots, \lambda_p)$ a vector (1xp) that gives the description of the magnitude of the expected change of the respective indicator for unit change in the latent/unobservable variable $\eta_t$. Where $\boldsymbol{\varepsilon_t} = (\boldsymbol{\varepsilon_{1t}}, \boldsymbol{\varepsilon_{2t}}, \ldots\ldots, \boldsymbol{\varepsilon_{pt}})$ is a (1xp) vector of disturbances each element in vector is white noise error term. Measurement model part assumes $\mathbf{E}(\mathbf{y_t}) = \mathbf{E}(\boldsymbol{\varepsilon_t}) = 0$ , this implies that indicator variables are measured as deviation from its mean. Second assumption is same as

above that disturbance term in measurement model is not correlated to either causal variable $X_t$ or to the unobservable variable $\eta_t$, hence, $\mathbf{E}(X_t\ \boldsymbol{\varepsilon_t}\ `) = \mathbf{E}(\boldsymbol{\varepsilon_t}\ X_t`) = 0$ and $\mathbf{E}(\eta_t\ \boldsymbol{\varepsilon_t}`) = \mathbf{E}(\boldsymbol{\varepsilon_t}\ \eta_t`) = 0$. Third and last assumption of measurement model $\mathbf{E}(\zeta_t\ \boldsymbol{\varepsilon_t}`) = \mathbf{E}(\boldsymbol{\varepsilon_t}\ \zeta_t`) = 0$, this implies the disturbance term or each white noise term don't correlate to the unexplained component of the model error term.

By using the structural equation model, equation (2) and measurement model, equation (3) we will derive the MIMIC covariance matrix $\sum$. Expressing equation (2) and (3) in terms of covariance:

$$\sum = \begin{bmatrix} \boldsymbol{Var(y_t)} & \boldsymbol{Cov(y_t X_t)} \\ \boldsymbol{Cov(X_t y_t)} & \boldsymbol{Var(X_t)} \end{bmatrix} \ldots\ldots(4) \quad = \mathbf{E}\left(\begin{bmatrix} \boldsymbol{y_t} \\ \boldsymbol{X_t} \end{bmatrix} \begin{bmatrix} \boldsymbol{y_t} \\ \boldsymbol{X_t} \end{bmatrix}'\right)$$

Before applying operation of multiplication, transpose and expectation remember the above assumptions of the MIMIC model. For complete derivation see appendix (A) so by using equation (a), (b), (c) and (d) we derive the following covariance matrix.

$$\sum = \begin{bmatrix} (\boldsymbol{\gamma'\Phi\gamma} + \boldsymbol{\psi})\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon} & \boldsymbol{\lambda\gamma`\Phi} \\ \boldsymbol{\Phi\gamma\lambda}` & \boldsymbol{\Phi} \end{bmatrix} \ldots\ldots\ldots(5)$$

Where, $\boldsymbol{\psi}$ represents the variance of error term $\boldsymbol{\zeta_t}$ or $Var(\zeta_t) = \boldsymbol{\psi}$, Covariance matrix of (qxq) causal variables $X_t$ is given by $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta_\varepsilon}$ represents the (pxp) covariance matrix of white noise error term $\boldsymbol{\varepsilon_t}$ in the measurement model. Now substituting equation a, b, c, and d and in equation (4) we get the covariance matrix of the MIMIC model.

The above matrix represents the relationship between observable variable i.e. causes and indicators while the decomposition this matrix will provide us the structure between observable variable and latent/unobservable variable in our case shadow economy. Now we elaborate the above model to consider the idea of cointegration.

**The MIMIC Model and Cointegration**

For modeling the shadow economy in most of the studies the time series properties of indicators and causes are not analyzed. Most importantly, variables are not stationary and order of integration is different from zero (I(d), d>0). If these properties of data are not fulfilled, then problem of spurious regression may arise. While in other studies, variables are of order I(1) were differenced to remove the randomness and trend component then Box and Jenkins method were employed to analyze the economic relationship. However, the drawback of this method is that long run

information lost. Later Granger and Weiss (1987) show two variables (dependent and independent) are integrated of order I(1) may have their linear combination of I(0). If this is the case then variables have cointegration among them. This theory can be extended the structural equation modeling MIMIC model (Andreas and Schneider, 2008).

To derive the error correction MIMIC (EMIMIC) substitute the equation (2) in equation (3) this yield the

$$\boldsymbol{y = \Pi X_t + Z_t} \ldots\ldots (6)$$

Where $\boldsymbol{\Pi = \lambda\gamma'}$ and the error term $\boldsymbol{Z_t = \lambda\zeta_t + \varepsilon_t}$. $Z_t$ in above equation (6) represents the vector of (px1) that represent the linear combination of error terms (white noise) $\boldsymbol{\zeta_t}$ and $\boldsymbol{\varepsilon_t}$ from the structural equation part and measurement equation model part, $\mathbf{Z_t \sim (0,\Omega)}$. $\boldsymbol{\Omega}$ represents the covariance matrix of $\mathbf{Z_t}$, $\boldsymbol{Cov(Z_t) = \lambda\lambda`\psi + \Theta_\varepsilon}$. Equation (5) is comparable to the simultaneous regression equation model where $y_t$ is (1xp) vector of endogenous variables that represent the indicators for latent variable and $X_t$ is (1xq) vector of exogenous variables that are the causes of the unobservable variable. Thus theory of cointegration is possible in MIMIC model (Andreas and Schneider, 2008).

Now we know that the linear combination $\boldsymbol{Z_{jt} = y_{jt} - \pi_j X_{it}}$ exists, where $Z_{jt}$ , j = 1,…..,p for time period subscript t is stationary white noise series, then variables will be consider to be cointegrated ( Engle and Granger ). Here $\boldsymbol{\pi_j}$ is $j^{th}$ (1xq) row vector of $\boldsymbol{\Pi}$ matrix thus cointegrating vector is $[1, \boldsymbol{\pi_j}]$. As $Z_{jt}$ , j = 1,….,p consist of q+1 variables, thus there can be more than one cointegrating vector (Greene). If p indicator variables are order of integration of I(1), the number of cointegrating vector is (p.q). One thing to be noted here not every variables is I(1) there may be macroeconomic variables that of order integrated I(0). We therefore include the causes in equation (6) that are I(0) so $v_t$ `= ($v_{1t}$, $v_{2t}$,….,$v_{rt}$) denote the vector of causes variables thus equation (6) yields the:

$$\boldsymbol{y_t = \Pi X_t + B V_t + Z_t} \ldots\ldots (7)$$

Where $\mathbf{B = \lambda\beta`}$ and $\boldsymbol{\tau}$`$\mathbf{= (\tau_1, \tau_2,\ldots,\tau_r)}$ represent the coefficient vector (1xr) of the order of integration I(0) causal variables in the structural equation relationship. As r of the causal variables are I(0), $\boldsymbol{\Pi = \lambda\gamma'}$ is the vector of order [1x(q-r)], $X_t$ is the vector of [(q-r)x1] and the dimension of B and $V_t$ is (pxr) and (rx1) respectively. If r of the causes variables are integrated of order I(0) then linear cointegrating vector for every white noise term $Z_{jt}$, j = 1,….,p in above equation will be (q-r). Similarly if s ≤ p are

the indicator variable are individually of I(0), then maximum number of linear cointegrating vector will decreases to (q-r)-s.

As we know that first difference of $\Delta y_t$ , $\Delta X_t$ and $Z_t$ are I(0), thus we can make the following equation:

$$\Delta \boldsymbol{y_t} = \boldsymbol{A\Delta X_t + BV_t + KZ_{t-1} + W_t} \ldots\ldots (8)$$

Where $W_t$ represents the white noise error term and all variable in above equation are integrated of order I(0), $\boldsymbol{Z_{t-1}}$ represents the one lagged error equilibrium of the cointegrated long run equation and represents the error correction term in the dynamics (Engle and Granger 1987). $\Delta \boldsymbol{y_t} = \boldsymbol{y_t - y_{t-1}}$, $\Delta \boldsymbol{X_t} = \boldsymbol{X_t - X_{t-1}}$, $\boldsymbol{Z_{t-1}} = \boldsymbol{y_{t-1} - \Pi X_{t-1}}$and coefficient matrices short run dynamics are represented by A, B and K in the model specification. Moreover $\mathbf{A} = \boldsymbol{\lambda\alpha`}$is the [px(q-r)] coefficient matrix of the I(1) causes in first difference form, $\mathbf{B} = \boldsymbol{\lambda\beta`}$ is the (pxr) matrix coefficient of I(0) causes and $\mathbf{K}= \boldsymbol{\lambda k`}$ is the (pxp) matrix coefficient for long run equilibrium error correction term. Thus equation (6) and (7) describe the EMIMIC model. Now we apply the fundamental rule of the structural equation modeling in which we minimize the discrepancy between observed covariance matrix and covariance matrix.

The covariance matrix for equation (7) can be represented by $\sum$.

$$\sum = \begin{bmatrix} \boldsymbol{Var(y_t)} & & \\ \boldsymbol{Cov(X_t, y_t)} & \boldsymbol{Var(y_t)} & \\ \boldsymbol{Cov(V_t, y_t)} & \boldsymbol{Cov(V_t, X_t)} & \boldsymbol{Var(V_t)} \end{bmatrix}$$

We derived the observed variables covariance matrix as a function of the model parameters. So we formulate this covariance matrix with MIMIC model assumption as discus above, (see Appendix) for complete derivation.

$$\sum = \begin{bmatrix} \boldsymbol{\lambda\Phi_2(\gamma'\Phi\gamma + 2\gamma`N\tau + \tau`\tau)\lambda` + \Theta_\varepsilon} & & \\ \boldsymbol{(\Phi_1\gamma + N\tau)\lambda`} & \boldsymbol{\Phi_1} & \\ \boldsymbol{(N`\gamma + \Phi_2\tau)\lambda`} & \boldsymbol{N`} \quad \boldsymbol{\Phi_2} & \end{bmatrix} \ldots\ldots\ldots (9)$$

Where N represents the **Cov ($\mathbf{V_t}$ , $\mathbf{X_t}$)**, covariance matrix for the I(1) and I(0) causes is denoted by $\boldsymbol{\Phi_1}$ and $\boldsymbol{\Phi_2}$ respectively. Other notations hold the same definition as discuss above.

Using the equation (8) the covariance matrix of short run error correction mechanism is:

$$\sum = \begin{bmatrix} \mathbf{Var}(\Delta \mathbf{y}_t) & & & \\ \mathbf{Cov}(\Delta \mathbf{X}_t , \Delta \mathbf{y}_t) & \mathbf{Var}(\Delta \mathbf{X}_t) & & \\ \mathbf{Cov}(\mathbf{V}_t , \Delta \mathbf{y}_t) & \mathbf{Cov}(\mathbf{V}_t , \Delta \mathbf{X}_t) & \mathbf{Var}(\mathbf{V}_t) & \\ \mathbf{Cov}(\mathbf{Z}_{t-1} , \Delta \mathbf{y}_t) & \mathbf{Cov}(\mathbf{Z}_{t-1} , \Delta \mathbf{X}_t) & \mathbf{Cov}(\mathbf{Z}_{t-1} , \mathbf{V}_t) & \mathbf{Var}(\mathbf{Z}_{t-1}) \end{bmatrix}$$

Taking into account the above assumptions, we derived the equation's short run covariance matrix in terms of the model parameters (see Appendix).

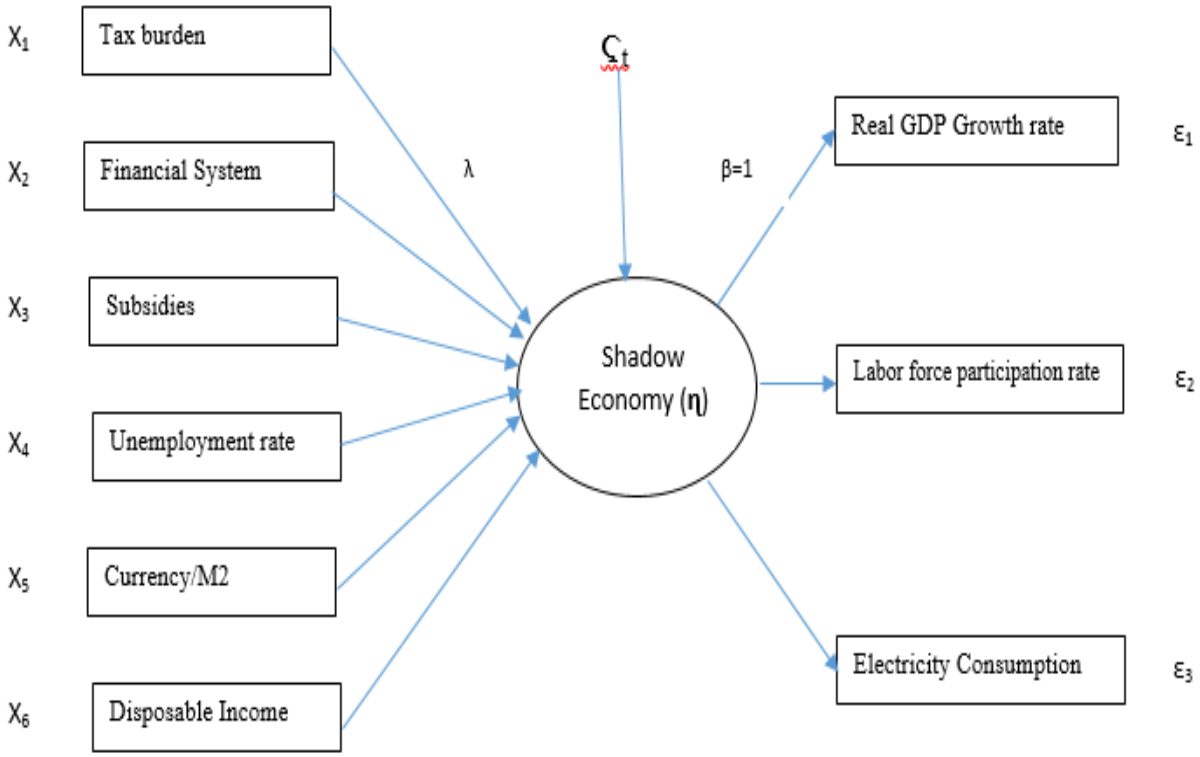
$$\sum = \begin{bmatrix} \boldsymbol{\lambda}(\boldsymbol{\alpha}'\boldsymbol{\Phi_3}\boldsymbol{\alpha} + 2\boldsymbol{\alpha}`\boldsymbol{M}\boldsymbol{\beta} + \boldsymbol{\beta}`\boldsymbol{\Phi_2}\boldsymbol{\beta} + \mathbf{k}`\Omega\mathbf{k}) & & & \\ (\boldsymbol{\Phi_3}\boldsymbol{\alpha} + \boldsymbol{M}`\boldsymbol{\beta})\boldsymbol{\lambda}` & \boldsymbol{\Phi_3} & & \\ (\boldsymbol{M}\boldsymbol{\alpha} + \boldsymbol{\Phi_2}\boldsymbol{\beta})\boldsymbol{\lambda}` & \mathbf{M} & \boldsymbol{\Phi_2} & \\ (\boldsymbol{\psi}\boldsymbol{\lambda}`\boldsymbol{k} + \boldsymbol{\psi})\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon} & \mathbf{0} & \mathbf{0} & \Omega \end{bmatrix} \ldots\ldots\ldots (10)$$

Where M = $\mathbf{Cov}(\mathbf{V_t}, \Delta\mathbf{X_t})$, now if we compare the equation (9) and (10) with equation (5) that represents MIMIC model covariance matrix, then the effect of cointegration can be seen. Thus above covariance matrix is adjusted by the long run equilibrium error term's covariance matrix $\boldsymbol{\Omega}$ and the error correction term's parameter vector k. Moreover$(\boldsymbol{\Phi_3}\boldsymbol{\alpha} + \boldsymbol{M}`\boldsymbol{\beta})\boldsymbol{\lambda}`$,$(\boldsymbol{M}\boldsymbol{\alpha} + \boldsymbol{\Phi_2}\boldsymbol{\beta})\boldsymbol{\lambda}`$, $\boldsymbol{\Phi_2}, \boldsymbol{\Phi_3}$ and **M** representing the sub matrix of the causal variables. However, it can be seen in equation (9) and (10) that $\sum$ is the function of the model's parameter $\boldsymbol{\alpha}$, **β**, **k**, and **λ** and of covariance, this implies that estimation of the covariance matrix of EMIMIC model is possible. In my theses this model will be employed to measure the development and magnitude of the Pakistan's shadow economy.

**Data and its Description**

We will use the annual data from 1972 to 2015 for the estimation of the shadow economy. For this the complete list of variables are subdivided into two groups, causes and indicators of the shadow economy. These causes and indicators of shadow economy are modelled using the path diagram. Where potential causes of the shadow economy are shown on the left side and indicators are shown on the right.

**Figure 1. 2: MIMIC Model for Pakistan**




**Causes of the Shadow Economy**

**Tax burden** ($X_1$) is the most popular factor behind the existence of the shadow economy in the literature. The hypotheses are that as tax burden increases then citizens have resilient incentive to work in the underground economy. Tax burden is measured by the total taxes including direct and indirect taxes as a proportion of gross domestic product (GDP). **Financial Sector development or Banking sector development** ($X_2$) is another factor that determine the intensity of underground activities. Total demand and time liabilities as a proportion of GDP is taken as financial sector development. Theoretically transaction through banking or financial sector cannot avoid taxes, therefore it is expected that development of financial sector have negative effect of the development of shadow economy. The reason behind to take financial development as a determinant of shadow activities in country like Pakistan small proportion of population is using banking system. We expect that this will significantly affect the shadow economy, as the less developed financial institution is considered to be an incentive to work in the shadow sector.

**Subsidies** ($X_3$) are the unrequired payments made by the government to protect the enterprises on the basis of value of goods and services which they produce, sell and import. The relationship between shadow economy and subsidies are conflicting, on

one subsidies decreases the underground activities because the cost of working in shadow market rises. On the other hand it distorted the competition by altering the tax burden of enterprises thus could encourage the firms to move in shadow sector. **Unemployment rate** ($X_4$). As labor force of the shadow economy is consist of very heterogeneous workers. Firstly classified as officially unemployed people but they are component of the official labor force. Secondly the workers of the underground economy are retired people, minors and housewives who are not the part of the official economy. **CM/M2 ratio** ($X_5$) is consider to be the other potential cause of the shadow economy as mostly transaction in underground are carried out using the cash. **Disposable personal income** ($X_6$) is another factor determine the dynamics of shadow economy. Difference of GDP and direct taxes has been taken as disposable income. It is hypothesized that individual income pattern may notable effect on underground activities.

**Indicators of Shadow Economy**

To mirror the underground activities in economy we use three indicator variables after analyzing the detailed literature. We use **growth rate of real gross domestic product** (GGDP) ($Y_1$), **labor force participation rate** ($Y_2$) and **growth in electricity consumption** ($Y_3$). These three variables are suitable for the purpose due to following reasons. Real GDP represent the measure of official economy; lower the measure of official GDP then there will be more incentive for citizens to work in the shadow economy. As shadow economy offering more money than official economy, thus we expect in short negative relationship between real gross domestic product and shadow economy. However, in long run there is positive relationship between official and unofficial economy or complementary relationship exist between them. The demand for work in the shadow economy increases as services and maintenance prices are increasing, as official economy is growing substantially as a large share of consumption in the economy (Andreas and Schneider, 2008). Based on these theoretical considerations we employed the above casual variable for shadow economy as shown above.

Labor force participation rate ($Y_2$) is calculated as the ratio of total labor force and population of working age group. Decline in the labor force participation rate in the official economy indicates the underground activities in the unofficial economy (Giles, 1998). Thus by incorporating this variable as an indicator, it is possible empirically if there is flow of resources from official economy to shadow economy.

Third variable that included to mirror the shadow economy is growth of electricity consumption ($Y_3$). Empirically study of Kaufmann and Kaliberda (1996) demonstrate that electricity consumption is the best indicator of the shadow economy. Moreover, Lacko (1996, 1997 a,b) showed that considerable part of the underground economy is linked with household consumption of electricity. As it is comprises of household production and unregistered production and services. Thus we expect positive relationship between the growth of electricity consumption and shadow economy as shown in the study of (Arby, Muhammad and Hanif 2010).

The table below shows the time period, data sources and brief description of the all causal and indicator variables that are employed in our study.

**Table 1. 3: Description of Variables and Source of Data**

| Variable | Symbol | Description | Source | Annotation |
|---|---|---|---|---|
| **Causes** | | | | |
| Tax burden | ($X_1$) | total share of direct and indirect taxes as percentage of gross domestic product (GDP) | Various issues-Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Financial development | ($X_2$) | Ratio of total demand and time liabilities to nominal GDP. | Various issues-Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Subsidies | ($X_3$) | Subsidies /GDP | Various issues-Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Unemployment rate | ($X_4$). | Officially unemployed people and retired persons and housewives that are not part of labor force. | Various issues-Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| CM/M2 | ($X_5$) | Ratio of currency in circulation to monetary aggregate of the banking system | Various issues-Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Disposable Income | ($X_6$) | GDP minus direct taxes. | Various issues-Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| **Indicators** | | | | |
| Real GDP growth rate | ($Y_1$) | Rate of change of GDP. | Various issues-Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |
| Labor force | ($Y_2$) | Ration of total labor | Various issues- | Yearly Data |

| | | | | |
|---|---|---|---|---|
| participation rate | | force and population of working age | Economic Survey of Pakistan/ WDI | (1972-2015) |
| Growth in Electricity Consumption | ($Y_3$) | Growth rate of total Electricity Consumption | Various issues- Economic Survey of Pakistan/ WDI | Yearly Data (1972-2015) |

## RESULTS AND DISCUSSION

**Unit Root Test on Variables of ARDL Model**

Before regression analysis, it is important to check the time series properties of all the variables of currency demand model. In order to check the order of integration we apply the augmented Dickey Fuller test (ADF). The results are presented in table given below.

**Augment Dickey Fuller test on Level of Series**

| | **Constant** | | **Constant, Linear Trend** | |
|---|---|---|---|---|
| Variables | ADF statistic | Prob. | ADF statistic | Prob. |
| CM | -0.917702 | 0.7728 | -2.234933 | 0.4587 |
| FD | -1.807004 | 0.3723 | -2.460818 | 0.3450 |
| TGDP | -1.057182 | 0.7239 | -2.765254 | 0.2174 |
| R | -2.947552 | 0.0486** | -2.608022 | 0.2789 |
| GGDP | -5.563289 | 0.0000*** | -5.525117 | 0.0002*** |
| GYD | -5.360362 | 0.0001*** | -5.342939 | 0.0004*** |
| INF | -4.724797 | 0.0004*** | -4.662748 | 0.0028*** |

*Note: *** shows significant level at 1 %** indicate significance level at 5%,*

The table shows some of the variables are stationary at the level such as interest rate (R), growth rate of GDP (GGDP), inflation rate (INF) and growth rate of disposable personal income (GYD). The other variables are stationary at first difference such as currency in circulation over M2 ratio (CM), financial development (FD), tax to GDP ratio (TGDP). As the order of integration of all variables is either I(0) or I(1) and none of the variable is I(2), so we applied ARDL model. In order to construct the model, we used the ARDL method and model is estimate by using the person (2001) ARDL bound testing approach for the short run and long run relationship. We estimated currency demand model using the equation (1) as shown below.

**Estimated Results of ARDL Model**

| **Variable** | **Coefficient** | **Variable** | **Coefficient** |
|---|---|---|---|

| | | | |
|---|---|---|---|
| $Const$ | 0.067 | $GYD_{t-1}$ | 0.660901 |
| | (0.067)*** | | (0.012)* |
| $CM_{t-1}$ | -0.439 | $\Delta TGDP_t$ | 1.331245 |
| | (0.001)* | | (0.000)* |
| $T.GDP_{t-1}$ | 1.260748 | $\Delta F_t$ | -0.379558 |
| | (0.001)* | | (0.000)* |
| $R_{t-1}$ | -0.003579 | $\Delta INF_t$ | -0.010396 |
| | (0.009)* | | (0.000)* |
| $INF_{t-1}$ | -0.006933 | $\Delta R_t$ | -0.002023 |
| | (0.030)** | | (0.097)*** |
| $F_{t-1}$ | -0.010690 | $\Delta GGDP_{t-1}$ | -0.003150 |
| | (0.8888) | | (0.000)* |
| $GGDP_{t-1}$ | -0.009026 | $\Delta GYD_{t-1}$ | 0.247619 |
| | (0.000)* | | (0.008)* |
| **Diagnostics** | | | |
| $R^2$ | 0.830094 | $Adjusted\ R^2$ | 0.751210 |
| $F-statistic$ | 10.52288 (0.000) | $DW\ stat$ | 1.73 |
| $LM\ test$ | 0.657 (0.424) | $SE$ | 0.008 |

*Note: ***, **,* indicate the statistical significance level at 10%, 5% and 1% respectively.*

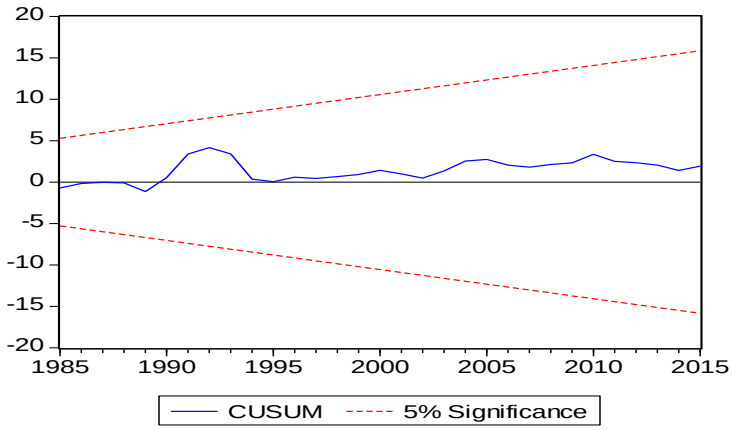
The validity test for the long-run relationship is tested by Wald coefficient restriction test. Null hypothesis is $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$. The calculated value of F-statistics is compared with the band of tabulated/critical values given by Pesaran et al (2001). The result of bound test is given in below table.

**Bound Test for Cointegration Analysis**

| | F- Statistics = 7.26 | | |
|---|---|---|---|
| **Critical Values** | **Lower Bond Values** | **Upper Bound Values** | **Decision** |
| 1% | 2.96 | 4.26 | Cointegration |
| 5% | 2.32 | 3.50 | Cointegration |
| 10% | 2.03 | 3.13 | Cointegration |

F-statistics is greater than upper bound value this implies that there is cointegration among the variables. The results of ARDL model is consistent with the theoretical frame work which have been discussed in previous chapter. The coefficients are statistical significant at 1%, 5% and 10% level of significance. CUSUM test is used to see the stability of the model, which is built on the collective sum of the recursive residuals. The figure 4.1 depicts the cumulative sum and the 5% critical lines. Breusch-Pagan LM test have been perform to check the autocorrelation, null hypothesis has been rejected this implies there is no autocorrelation.

**CUSUM Test of Stability**




In above table estimated results of currency demand model are given. All variables are significant at given level of significance and over all significance possess no problem. The numbers in the parentheses are probability values of the respective coefficient. Sign of coefficients are as per theoretical frame work, in this study it is hypothesized that taxes rises, people have incentive to involve themselves in tax avoiding activities. The use of currency facilitates the tax evasion, cash transaction accountability is difficult to manage, thus demand for cash rises and consequently, currency in circulation over money in ratio form is expected to rise. It can be confirmed from the above table that sign of coefficient of tax is positive in the long run as well as short run dynamics. The long run coefficient of interest rate is highly significant while the short run coefficient is significant at 10% level of significance. It possesses negative sign which confirms our hypotheses that, higher the interest rate may increase the opportunity cost of currency holdings and thus it leads to fall in demand for cash. Inflation is also incorporated in the above model in order to capture the opportunity cost of currency holding. When there is rise in inflation then people demand less money due to decrease in its value because hard assets get an increasing

return as inflation rise, thus sign of coefficient of inflation is negative which significant in long run as well as short run. Ratio of total demand and time liabilities to GDP is used as a proxy for financial development that signifies how much our banking sector are functional. According to theory that advancement in financial sectors especially in banking services may lower the demand for currency due technological prowess for transactional facilities. Thus the sing of financial development variable is negative which confirms the given hypotheses but the problem is that long run coefficient is insignificant implying that small proportion of population in Pakistan is using banking sector while other use the cash transactional systems. The variable disposable income per capita is constructed GDP minus direct taxes over population. Growth rate of this variable indicate growth of income after taxes, it is hypothesized that growth rate of disposable income is positively related to demand for currency, as demand of money is a result of the liquidity advantage of holding money.

After the approval of bound test and interpretations of the results following long run model is deduced from the estimated ARDL model.

$$\hat{}CM_t = \lambda_0 + \lambda_2 T.GDP_t + \lambda_3 R_t + \lambda_4 INF_t + \lambda_5 F_t + \lambda_6 GGDP_t + \lambda_7 GYD_t$$

$$\lambda_i = -\frac{\beta_i}{\beta_1} \text{ for } i = 0,2,3,4,5,6,7.$$

After the predicted calculated value of the currency demand model, the magnitude of underground economy and tax evasion can be determined as follows. For every year the estimated values of currency demand model with taxes $(\hat{}CM_t)_T$ and without taxes $(\hat{}CM_t)_{WT}$ are determined by employing above predicted regression equation. The change between $(\hat{}CM_t)_T - (\hat{}CM_t)_{WT}$ provides an overview of the level of currency holding that is due to taxes alternatively it means that the limit to which total taxes (direct +indirect) have influence people to amass higher number of currency. The scope of increased demand for cash highlights the size of tax evasion which is termed as illegal money. The mathematical expression for illegal money (IM) can be described as

$$IM = [(\hat{}CM_t)_T - (\hat{}CM_t)_{WT}] * M2$$

As Tanzi highlighted that legal money (LM) can be obtained by difference between M1 which is the sum of currency and demand deposits that is total money supply and estimated illegal money. The mathematical formula is as follows

$$LM = M1 - IM$$

After the determination of legal money, the income velocity of legal money can be calculated dividing the GNP by legal money which is shown as under.

$$IV = \frac{GNP}{LM}$$

The basic assumption of the model states that velocity of legal money and illegal money is the same, thus the magnitude of shadow economy is the result of the product of illegal money and income velocity of money. Mathematically it can be represented as

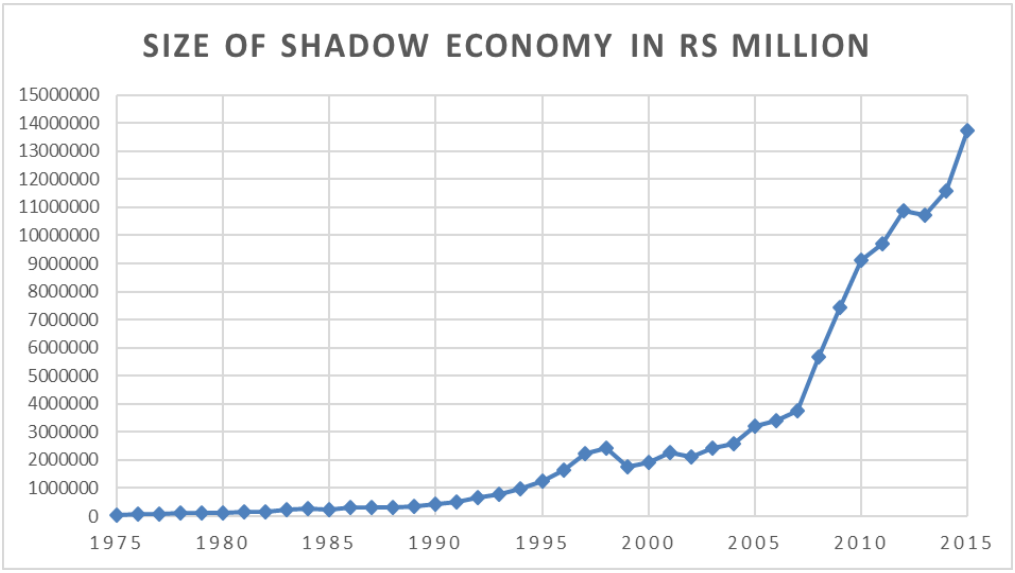
$$SE = IM * IV$$

Multiplication of estimated shadow economy with tax to GDP ratio will yield the total tax evasion in Pakistan. Mathematically represented by the formula

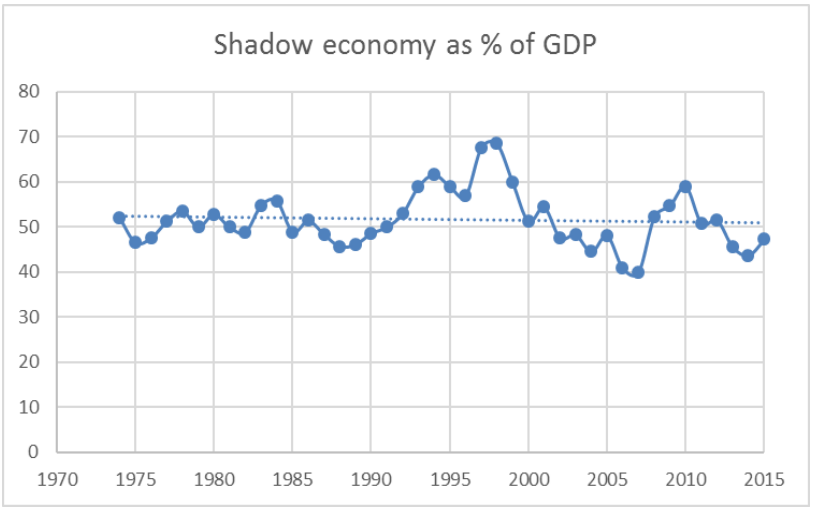
$$TE = SE * T.GDP$$

The dynamics of the underground economy can be observed in the below figure in estimated value of shadow economy is drawn against time.

**Size of Shadow Economy**

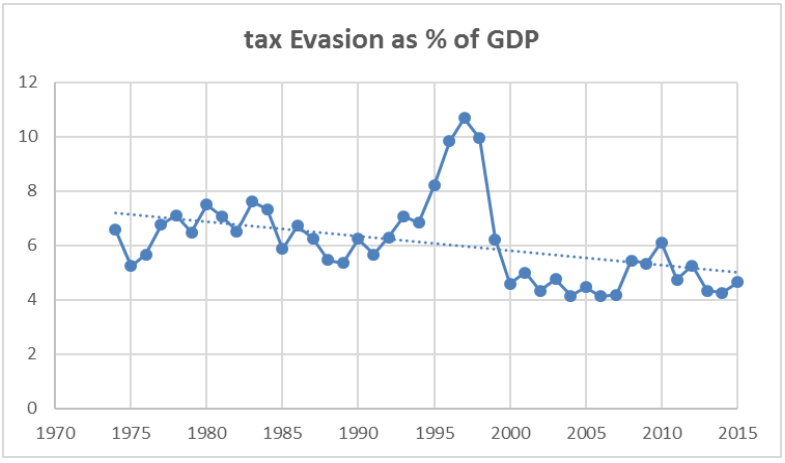



**Size of Shadow Economy as % of GDP**




The above figure depicts that a rise in shadow economy is observed steadily from 1990's onward. Now we calculate the shadow economy as a percent of the official economy the mathematical expression for this UE/GDP *100. The results for the estimation are presented in table in appendix in which detailed dynamics of shadow economy, income velocity of legal money, tax evasion as % of GDP and growth rate of underground economy is given. The dynamics of shadow economy as percentage of GDP is shown in above figure. It can be seen unusual growth of shadow economy in 1990's which is accompanied by growth rate of tax evasion is same period of time as shown in figure below. While in other years it is growing persistently with constant rate.

**Tax Evasion in Pakistan**




**Estimated Results of EMIMIC Model**

For valid and non-spurious relationship, it is imperative to analyze the order of integration of macroeconomic variables. To start our empirical analysis, the pre-testing of data which hypothesized the null hypothesis of unit root against the alternative of no unit root has been conducted. For this we employed Augmented Dicky Fuller (ADF) test in order to know the order of integration of the variables. The result of ADF test are shown in table 4.4.

As shown in the table, most of the variables like ratio of currency in circulation to M2 (CM), Financial sector development (FD), share of total taxes in GDP (TGDP), disposable income (YD) and labor force participation rate (LF) are of the order of integration I(1) using the standard unit root test. However some variables like subsidies (SUB), unemployment rate (UN), growth of real GDP (GGDP) and growth of electricity consumption (ELEC) are of the order of integration of I(0).

**Analysis of Stationarity**

| | **Constant** | | **Constant, Linear Trend** | |
|---|---|---|---|---|
| Variables | ADF statistic | Prob. | ADF statistic | Prob. |
| | | **CAUSES** | | |
| CM | -0.917702 | 0.7728 | -2.234933 | 0.4587 |
| FD | -1.807004 | 0.3723 | -2.460818 | 0.3450 |
| TGDP | -1.057182 | 0.7239 | -2.765254 | 0.2174 |
| SUB | -2.970125 | 0.046** | -2.920935 | 0.1665 |
| UN | -3.928693 | 0.00*** | -4.387129 | 0.00*** |
| YD | 1.602362 | 0.9991 | -1.702939 | 0.7312 |
| | | **INDICATORS** | | |
| GGDP | -5.563289 | 0.00*** | -5.525117 | 0.00*** |
| ELEC | -3.660310 | 0.00*** | -6.615544 | 0.00*** |
| LF | -2.520731 | 0.1179 | -2.257460 | 0.4469 |

*Note: ***, ** indicating the level of significance at 1% and 5% respectively.*

**Analysis of Cointegration in MIMIC Model**

In this section, we employ the two step cointegration approach of (Engle and Granger) to find if, the all seven causes are cointegrated with respective indicator. Here the specific indicator variables including growth of real GDP per capita, labor force participation rate and growth of electricity consumption represents the dependent variables while causes variables are independent variables. In order to find if the

seven causes are cointegrated with each indicators exhibiting a valid error correction term we estimate the least square regressions with variables in level as shown below.

$$GGDP = \alpha_0 + \alpha_1 T.GDP + \alpha_2 CM + \alpha_3 FD + \alpha_4 SUB + \alpha_5 UN + \alpha_6 YD + U_1$$

$$LF = \alpha_0 + \alpha_1 T.GDP + \alpha_2 CM + \alpha_3 FD + \alpha_4 SUB + \alpha_5 UN + \alpha_6 YD + U_2$$

$$ELEC = \alpha_0 + \alpha_1 T.GDP + \alpha_2 CM + \alpha_3 FD + \alpha_4 SUB + \alpha_5 UN + \alpha_6 YD + U_3$$

From these three equations, we obtain the residual $u_1$, $u_2$ and $u_3$respectively. Next we analyze the stationarity of these three residuals using the ADF test. If there exists cointegration between the causes and specific indicators then we guess ADF test to nullify the null hypothesis that says there is unit root against the alternative hypothesis, above stated error terms $u_1$, $u_2$ and $u_3$. The results of ADF test for these three error terms confirms that there is cointegration between the causes and specific indicators as shown in below table, we can reject the null hypothesis at given significance level.

**Table 4. 1 Stationarity Analysis of Error Terms**

| | **Constant** | | **Constant, Linear Trend** | |
|---|---|---|---|---|
| Variables | ADF statistic | Prob. | ADF statistic | Prob. |
| U1 | -4.874167 | 0.0003*** | -4.813242 | 0.0019*** |
| U2 | -3.034865 | 0.0397** | -2.365814 | 0.3911 |
| U3 | -6.367016 | 0.0000*** | -6.301333 | 0.0000*** |

*Note: ***, ** indicating the level of significance at 1% and 5% respectively.*

The cointegration correspondences allow us the estimation of a long run equilibrium MIMIC model for the magnitude and evolution of underground economy using the equation (2). The next step is the estimation of short run MIMIC model of equation using the equation (8), employing the difference of all causes and indicators which are of the order of integration I(1). While $u_1$, $u_2$ and $u_3$from the cointegration relationships are used as error correction terms in the estimation. The table 4.4 represents the coefficients of long run equilibrium and short run coefficient along with (ECM) error correction term here is represented by $u_1$, $u_2$ and $u_3$.

**Results and Estimation of MIMIC Model**

| Variables | Long-run MIMIC Model | | Short-run MIMIC Model | |
|---|---|---|---|---|
| **Causes** | | | | |
| Tax Burden | .261 | (2.132) ** | .019 | (.138) |
| Financial Sector | -.080 | (-1.735) * | -.086 | (-1.771) * |
| Subsidies | .290 | (2.167) ** | .166 | (1.478) |
| Disposable Income | -.198 | (-2.530) ** | -.284 | (-2.377) ** |
| Currency in circulation/M2 | .134 | (2.718) *** | -.007 | (-0.136) |
| Unemployment Rate | .144 | 1.475 | .238 | (1.919) * |
| U1 | -- | -- | -.060 | (-1.393) |
| U2 | -- | -- | -.199 | (-2.109) * |
| U3 | -- | -- | -.005 | (-0.203) |
| **Indicators** | | | | |
| GGDP | 1.000 | | 1.000 | |
| Labor force participation rate | -2.117 | (-2.930) * | -3.273 | (-2.458) ** |
| Growth of Elec. Consumption | 2.938 | (2.697) ** | 3.657 | (2.288) ** |
| **Diagnostics** | | | | |
| CMIN/DF | 3.43 | | 1.556 | |
| Root Mean Square | 0.036 | | 0.023 | |
| GFI | 0.890 | | 0.901 | |
| AGFI | 0.850 | | 0.860 | |

NOTE: *, **, *** indicating the level of significance at 1%, 5% and 10% respectively.

As discuss above currency demand model provide the index of shadow economy using only one factor that causes the underground activities. However, MIMIC model uses the multiple indicators and multiple causes to provide the detailed size and dynamics of the shadow economy. The results of currency demand model will be used in MIMIC model in order to estimate the size of underground economy. In case of Pakistan's shadow economy MIMIC model for the first time employed by Arby at el. in 2010. They used three causal variables tax/GDP ratio for tax burden, M2/GDP ratio and regime durability. While to mirror the shadow economy they used two indicators currency in circulation over M2 ratio and growth in electricity.

In this study we have employed the time series analysis of multiple indicator-multiple causes model, following the Andreas and Schneider (2008). we have employed six causal variables and three indicators. The selection of the causes and indicators of the shadow economy are based on the theoretical and empirical evidences found the in the literature. The above table presents the estimated results of MIMIC model, all coefficients possess the expected sign as per economic theory and significant at conventional level of significance. Tax burden (T.GDP) is consider to be most important determinant of the shadow activities. Theoretically it is hypothesized that rise in tax weight may encourages the workers and firms to work in the shadow economy. As taxes affect the pattern of labor-leisure and this drive the labors to work in underground or black economy. When there exists discrepancy between the total cost of labor in the sanctioned economy and the black economy or when income after taxes from work in the authorized economy is less than income of underground activities then people have strong reason to work in the shadow or black economy. This can be confirmed from the above results sign of tax burden coefficient is positive and possesses significance in the long run while in short run it has no effect on shadow economy. Financial sector development (FD) possess expected sign as per economic theory and significance at given level of significance, here it is hypothesized that when our banking sectors are developed then people demand less cash because of other financial instrument. This leads to lesser the magnitude of shadow economy. The coefficient of subsidies (SUB) is significant and possesses positive sign, as increase in subsidies will distort the competition as a result there will be significant effect in the net tax burden between the industries. This will provide incentive to industries to continue its underground activities, it can also be inferred from this that in Pakistan subsidies allocation is not targeting the market efficiency rather it is discriminating between the firms on the basis of different cartels, lobby and geographical location. The theoretical consideration behind the cause, disposable income is that if personal income (YD) is low in official economy due to tax burden, regulation and personal limitation, then individual has incentive to try his luck in underground economy. The coefficient of (YD) is negative and statically significant which confirms our above hypotheses. As most of the transaction in shadow economy is carried out by cash so it considered to be the most important cause of underground activities.

Ratio of currency in circulation to M2 (CM) is used as proxy for cash available in the market and it possesses the expected sign as per theoretical frame work and statistical significance in long run while in short run it does not possess no effect of shadow economy. The relationship between unemployment rate (UN) and shadow economy is ambiguous (Tanzi, 1990), here it means that composition of labor force for shadow economy is very heterogeneous. The group of people who are consider to be unemployed are the part of the official labor force while other group is composed of retired people, illegal immigrants, child labor and housewives who are not part of the official economy. Moreover, there are people who have jobs in official as well as unofficial economy in this scenario unemployment rate is weakly correlated with the shadow economy this can be depicted from our above estimated results where the coefficient of unemployment is insignificant in long run and possesses positive sign which shows that substitution effect is greater than income effect. In short run MIMIC model estimates the residual u2 from the cointegration relationship between the causes and labor force participation rate is statically significant and possesses expected theoretical sign. While other two residuals are not statically different from zero.

To indicate the shadow economy, we have employed the three indicators as discuss above. In order to estimate the parameters of relative size and on their level, researcher must fix the scale or it can be considered as normalization of the latent variable. The conventional way to know the comparative magnitude of the variable then one indicator's coefficient should be equated to non-zero that is equal to 1. So to estimate the MIMIC model here we fix the coefficient of GGDP. While two other indicators labor force participation rate (LF) and growth rate of electricity consumption (ELEC) possess expected sign as per theoretical frame work and significance at given level of significance. In literature labor force participation rate is considered to be one of the most important indicator as change in participation rate gives the information about the flow of resources between the official and the shadow economy. Shadow economy has negative and significant effect on labor force participation rate, as shadow economy increases labor force participation rate decreases as it reflects the movement of the work force from official economy to unofficial economy. Empirically study of Kaufmann and Kaliberda (1996) demonstrate that electricity consumption is the best indicator of the shadow economy. Moreover, Lacko (1996, 1997 a,b) showed that considerable part of the underground

economy is linked with household consumption of electricity. As it is comprising of household production and unregistered production and services. Our finding supports the Kaliberda's study coefficient of growth of electricity consumption is positive and statically significant, which demonstrate that rise in shadow economy increases the electricity consumption in commercial and household sectors.

After analyzing the estimated results, it is important to discuss the model diagnostics and different goodness-of-fit statistics that they are supporting our proposed model or not. Basically these goodness-of-fit statistics measures are based on fitting the model to the sample moments. So we have reported the four goodness of fit statistics in above table. The discrepancy function (CMIN) mathematical it is ratio of chi-square value and degree of freedom. CMIN is consider to be the most common fit test and gives the information about the least value of the discrepancy function that exist between the the estimated covariance matrix and sample covariance matrix for good model it must be less than 5. Goodness-of-fit index (GFI) and adjusted Goodness-of-fit index (AGFI) are also gives the information about the discrepancy between the estimated and observed covariances. GFI tells us how much percent of observed covariances are explained by the covariances implied by the model while AGFI are adjusted for the degree of freedom. Another fit test is root mean square error of approximation (RMSEA), conventionally if RMSEA is less than or equal to 0.05, then model considers to be good fit.

After obtaining the estimates of our EMIMIC model which include the long-run cointegration equilibrium relationship and short-run dynamics error correction terms of the MIMIC model, now we can make the index of Pakistan's shadow economy. Using the estimates of long-run we will evaluate the ordinal index then this can be converted in to cardinal scale using the average value of shadow economy obtained from the above currency demand approach. In the next step we calculate the short-run deviation from the equilibrium, finally taking these into account magnitude of shadow economy is calculated using the Bajada and Schneider's (2005) calibration methodology as shown below.

Now for benchmarking procedure we follow the Schneider et al. (2006). According to the identification rule or for normalization we set the coefficient of single indicator is equal to 1 and the index of the shadow economy as a percentage of GDP in the 1990 can be explained as follow

Measurement equation: $\frac{\boldsymbol{GDP_t - GDP_{t-1}}}{\boldsymbol{GDP_{1990}}} = \frac{\boldsymbol{\hat{\eta}_t - \hat{\eta}_{t-1}}}{\boldsymbol{GDP_{1990}}}$

To obtain an ordinal time series index for shadow economy we used the estimates of equation (1).

$$Structural\ equation: \frac{\boldsymbol{\hat{\eta}_t}}{\boldsymbol{GDP_{1990}}} = \boldsymbol{\gamma' X_{qt}} \ldots\ldots (11)$$

For further transformation of shadow economy as ratio of current GDP, the following operation have been employed.

$$\frac{\boldsymbol{\hat{\eta}_t}}{\boldsymbol{GDP_{1990}}}\left[\frac{\boldsymbol{\dot{\eta}_{1990}}}{\boldsymbol{GDP_{1990}}} * \frac{\boldsymbol{GDP_{1990}}}{\boldsymbol{\hat{\eta}_{1990}}}\right]\frac{\boldsymbol{GDP_{1990}}}{\boldsymbol{GDP_t}} = \frac{\boldsymbol{\eta_t}}{\boldsymbol{GDP_t}} \ldots\ldots (\mathbf{12})$$

Where

- $\frac{\boldsymbol{\hat{\eta}_t}}{\boldsymbol{GDP_{1990}}}$ is the ordinal index of shadow economy estimated by using equation.11
- $\frac{\boldsymbol{\dot{\eta}_{1990}}}{\boldsymbol{GDP_{1990}}}$ is the estimated value of shadow economy by using the above currency demand model in 1990.
- $\frac{\boldsymbol{\hat{\eta}_{1990}}}{\boldsymbol{GDP_{1990}}}$ is the estimated value of shadow economy in the year 1990 by using eq=(11)
- $\frac{\boldsymbol{GDP_{1990}}}{\boldsymbol{GDP_t}}$ will convert the index of shadow economy as changes respect to base year in a time series of SE/current GDP.
- $\frac{\boldsymbol{\eta_t}}{\boldsymbol{GDP_t}}$ is the estimated magnitude of shadow economy as percentage of official GDP.

Equation (12) can be simplified in to: $\frac{\boldsymbol{\hat{\eta}_t}}{\boldsymbol{\hat{\eta}_{1990}}} * \frac{\boldsymbol{\dot{\eta}_{1990}}}{\boldsymbol{GDP_t}} = \frac{\boldsymbol{\eta_t}}{\boldsymbol{GDP_t}}$

Now we can obtain the magnitude and development in shadow economy over the time, table below represent the time series value of shadow economy expressed as a percentage of official GDP for the sub-sample of five years, while the yearly estimates of shadow economy are given in appendix.

**Table 4. 2:Shadow Economy estimated by EMIMIC Model**

| Years | 1974-1980 | 1980-1985 | 1985-1990 | 1990-1995 | 1995-2000 | 2000-2005 | 2005-2010 | 2010-2015 |
|---|---|---|---|---|---|---|---|---|
| Shadow Economy | 49% | 46% | 44% | 39% | 35% | 35% | 32% | 30% |

In figure below represents the outcome of EMIMIC model by which we have estimated the shadow economy.

**Figure 4. 1: Time series plot of Shadow Economy as % GDP**

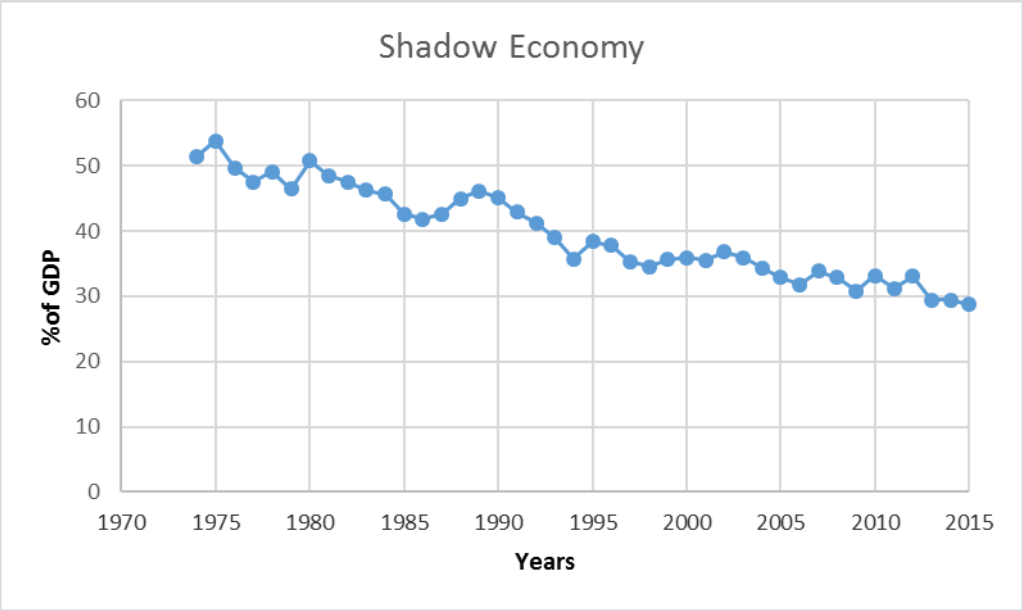



Figure 4.5 depicts the magnitude of shadow or underground economy as a proportion of gross domestic product (GDP) over time the size of shadow economy ranges from around 50% in 1974 to 28% in 2015. The unusual decrease in shadow economy is observed in 1980-1985 and 1990-1995 and the first twenty years' permanent change in the decreasing dynamics of the size of black economy. After 1998 shadow economy remains unfluctuating regardless of the initial values, it can be seen shadow economy as percent of official economy is stable in Pakistan from this it can be confirmed that shadow economy has shown development with almost the same rate of the official economy.

Empirical analysis of MIMIC model reveals that Taxes, financial sector development, currency/cash, unemployment, personal income and subsidies are the main causes of shadow economic dynamics. So we have calculated the five year averages of annual

growth rates of these causes and shadow economy to see reasons behind the development of shadow economy in Pakistan overtime.

**Table 4. 3:Average of annual growth rates**

| Years | T.GDP | SUB | UN | FD | YD | CM | SE |
|---|---|---|---|---|---|---|---|
| 1975-1980 | -0.30 | -3.56 | 5.30 | -0.95 | 14.72 | 1.56 | -0.24 |
| 1981-1985 | -3.25 | 6.65 | 0.28 | 0.13 | 10.69 | -4.43 | -3.49 |
| 1986-1990 | 1.95 | 2.36 | -1.14 | 1.86 | 8.70 | 1.88 | 1.11 |
| 1991-1995 | -0.25 | -10.69 | -4.72 | 3.79 | 12.71 | -5.09 | -3.22 |
| 1996-2000 | -5.83 | 9.28 | 7.52 | -4.91 | 11.90 | -0.58 | -1.32 |
| 2001-2005 | 0.03 | 11.82 | -4.64 | 4.66 | 8.75 | -2.43 | -1.78 |
| 2006-2010 | 1.86 | 11.20 | -0.82 | -2.72 | 14.54 | -0.06 | 0.16 |
| 2011-2015 | -0.45 | -20.42 | 1.77 | -0.18 | 11.63 | 0.83 | -2.85 |

From above table it can be seen that decreasing trend in the shadow economy from 1981-1985 and 1991-1995 was simultaneous with the large reduction in tax burden, unemployment rate and currency in circulation at the same time improvement in the banking sector can be observed. While the pattern of subsidies and personal income is somewhat ambiguous, there seems to be compensation among these causal variables.

## CONCLUSION AND POLICY RECOMMENDATION

In this thesis we attempt to model the shadow economy of Pakistan by using the joint modified currency demand approach and multiple causes and multiple indicator (MIMIC) approach. Detailed econometric analysis have been conducted to estimate the magnitude and dynamics of Pakistan's underground economy. We have employed the ARDL approach in currency demand model to get point estimator of shadow economy which is used in the MIMIC model in order to understand the dynamics of the shadow economy in long run.

Estimated magnitude of shadow economy is overestimated by currency demand model as it relies on only one indicator tax burden, it can be seen that magnitude of shadow economy is stable for sample time period around 50% of the official GDP. While in 1990's unusual growth of shadow economy approximately 70% of the official economy have been observed due to abundant tax evasion in this period. Due to limitation of monetary approach we have employed the detailed and comprehensive Multiple Causes and Multiple indicators (MIMIC) model estimated by structural equation modeling technique.

While estimating the MIMIC model we take into account time series properties of data (stationarity) and cointegration relationship which have been ignored in previous studies. So we have modeled the long run MIMIC model and short run MIMIC model which provide the detailed information about the dynamics of Pakistan's shadow economy and convergence have been witnessed in our study as sign of error correction term is negative and statically significant. The dimensions of shadow economy in Pakistan by using structural equation modeling (SEM) technique is between 50% to 28% of the official economy. In the first twenty years shadow economy is decreasing due to lower tax burden, unemployment rate, currency in circulation and at the same improvement the financial sectors. While in last thirteen years from 1998 to 2015 shadow economy is stable over time that 30% of the official economy indicating that shadow economy is growing at the same rate as official economy.

**Policy Recommendation**

Rules and regulation set by the government should be concise and clear regarding the tax pattern in the country, as it can be seen from the above results that magnitude of shadow economy is highly sensitive to taxes. Therefore, tax pattern should be simple and easy to follow otherwise it can drive the individual in the underground activities. Most of the transaction in underground economy is carried out through cash therefore autonomous monetary policy should be conducted for the stability of good economic system in a country. Financial sectors should be developed and government should conduct the surveys about the working age group so that composition of labor force should be more clear as in above result shadow economy was highly effected by unemployment rate in short run.

**Further Direction**

The combination of monetary approach and MIMIC model consider to be the best available option in literature to estimate the shadow economy. However, still further direction of research can be suggested, firstly government regulation should be taken into account as a causal variable of shadow economy. Secondly illegal activities like smuggling, bribery, robbery etc. should be taken into consideration while determining the dynamics of shadow economy and at last costly micro methods can provide more detailed and imperative information about development underground which can only be conducted by state offices.

# References


Ahmed, M., and Qazi Masood Ahmed (1995) Estimation of the Black Economy

Ahmed, Q. M., & Hussain, M. H. (2006). Estimating the black economy through monetary approach: a case study of Pakistan.

Arby, M. F., Malik, M. J., & Hanif, M. N. (2010). The size of informal economy in Pakistan.

Bajada, C., & Schneider, F. (2005). THE SHADOW ECONOMIES OF THE ASIA-PACIFIC. *Pacific Economic Review*, *10*(3), 379-401.

Bajada, C., & Schneider, F. (2005). *Size, Causes and Consequences of the Underground Economy: an international perspective*. Gower Publishing, Ltd.

Belev, B. (Ed.). (2003). The informal economy in the EU Accession Countries: size, scope, trends and challenges in the process of EU enlargement. Csd.

Breusch, T. (2005). *Estimating the Underground Economy, Using MIMIC Models*. Working Paper, National University of Australia, Canberra, Australia.

Buehn, A., & Schneider, F. (2008). MIMIC models, cointegration and error correction: An application to the French shadow economy.

Buehn, A., & Schneider, F. (2008). MIMIC models, cointegration and error correction: An application to the French shadow economy.

Buehn, A., & Schneider, F. (2012). Corruption and the shadow economy: like oil and vinegar, like water and fire?. *International Tax and Public Finance*,*19*(1), 172-194.

Cagan, P. (1958). The demand for currency relative to total money supply. In*The Demand for Currency Relative to Total Money Supply* (pp. 1-37). NBER.

Choi, J. P., & Thum, M. (2005). CORRUPTION AND THE SHADOW ECONOMY*. *International Economic Review*, *46*(3), 817-836.

Dell'Anno, R., & Schneider, F. G. (2006). *Estimating the underground economy by using MIMIC models: A Response to T. Breusch's critique* (No. 0607). Working Paper, Department of Economics, Johannes Kepler University of Linz.

Dell'Anno, R., Gómez-Antonio, M., & Pardo, A. (2007). The shadow economy in three Mediterranean countries: France, Spain and Greece. A MIMIC approach. *Empirical Economics*, *33*(1), 51-84.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, 251-276.

Feige, E. L. (1979). How big is the irregular economy?. *Challenge*, *22*(5), 5-13.

Frey, B. S., & Schneider, F. (2000). *Informal and underground economy* (No. 0004). Working Paper, Department of Economics, Johannes Kepler University of Linz.

Giles, D. E. (1999). Measuring the hidden economy: Implications for econometric modelling. *The Economic Journal*, *109*(456), 370-380.

Giles, D. E. (1999). The rise and fall of the New Zealand underground economy: are the responses symmetric?. *Applied Economics Letters*, *6*(3), 185-189.

Gutmann, P. M. (1977). The subterranean economy. *Financial Analysts Journal*, *33*(6), 26-27.

Helberger, C., & Knepel, H. (1988). How big is the shadow economy?: A re-analysis of the unobserved-variable approach of BS Frey and H. Weck-Hannemann. *European Economic Review*, *32*(4), 965-976.

Iqbal, Z., Qureshi, S. K., & Mahmood, R. (1998). *The Underground Economy and Tax Evasion in Pakistan. A Fresh Assessment* (No. 1998: 158). Pakistan Institute of Development Economics.

Kaufmann, D., & Kaliberda, A. (1996). Integrating the unofficial economy into the dynamics of post-socialist economies: A framework of analysis and evidence. *World Bank Policy Research Working Paper*, (1691).

Kemal, M. A. (2007). *A Fresh Assessment of the Underground Economy and Tax Evasion in Pakistan: Causes, Consequences and Linkages with the Formal Economy* (No. 2007: 13). Pakistan Institute of Development Economics.

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

O'Driscoll Jr, G. P., Feulner, J. E., & O'Grady, M. A. (2003). Index of Economic Freedom, The Heritage Foundation and Dow Jones & Company.

of Pakistan through the Monetary Approach. *The Pakistan Development*

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, *16*(3), 289-326.

Popov, O. S. (1999). *Metody numeryczne i optymalizacja*. Wydaw. Uczelniane PS.

*Quarterly Review*, 35, 427-53.

*Review* 34:4, 791–807.

Schneider, F., & Dell'Anno, R. (2003). The Shadow Economy of Italy and other OECD Countries: What do we know?. *Journal of public finance and public choice= Economia delle scelte pubbliche*, *21*(2), 97-120.

Schneider, F., & Enste, D. (2000). Shadow Economies around the World Size, Causes, and Consequences.

Schneider, F., & Enste, D. H. (2013). *The shadow economy: An international survey*. Cambridge University Press.

Schneider, F., Buehn, A., & Montenegro, C. E. (2010). Shadow Economies all over the World: New Estimates for 162 Countries from 1999 to 2007.*World Bank Policy Research Working Paper Series, Vol*.

Tanzi, V. (1983). The Underground Economy in the United States. *Banca Nazionale del Lavoro Quarterly*.

Tanzi, V. (1999). Uses and abuses of estimates of the underground economy. *The Economic Journal*, *109*(456), 338-347.

Tanzi, Vito (1980). "The Underground Economy in the United States: Estimates and Implications."

Tanzi, Vito (1983). "The Underground Economy in the United States: Annual Estimates, 1930-80." *IMF Staff Papers* 30:2, 283-305.

Torgler, B., & Schneider, F. (2007). Shadow economy, tax morale, governance and institutional quality: a panel analysis.

Weiss, L. (1987). Explaining the underground economy: state and social structure. *British Journal of Sociology*, 216-234.

# Appendix (A)

Before applying the expectations on matrix remember the following assumption of the MIMIC model.

1. $\mathbf{E}(\eta_t) = \mathbf{E}(X_t) = \mathbf{E}(\zeta_t) = \mathbf{E}(\mathbf{y_t}) = \mathbf{E}(\boldsymbol{\varepsilon_t}) = 0$ , this means variables are measured as deviation from its mean.
2. $\mathbf{E}(X_t\ \zeta_t`) = \mathbf{E}(\zeta_t\ X_t`) = \mathbf{E}(X_t\ \boldsymbol{\varepsilon_t}\ `) = \mathbf{E}(\boldsymbol{\varepsilon_t}\ X_t`) = 0$, this implies error term has no correlation to the causal variable.
3. $\mathbf{E}(\zeta_t\ \boldsymbol{\varepsilon_t}`) = \mathbf{E}(\boldsymbol{\varepsilon_t}\ \zeta_t`) = 0$, no correlation between error term across the equation.
4. $\mathbf{E}(\eta_t\ \boldsymbol{\varepsilon_t}`) = \mathbf{E}(\boldsymbol{\varepsilon_t}\ \eta_t`) = 0$, the error of measurement model has no correlation to the latent/unobservable variable.

Now we apply the expectation in order to derive the both variance and covariance between the observable variable, it follows the following steps:

$\mathbf{Var\ (y_t) = E\ (y_t\ y_t`) = E}[(\lambda\eta_t + \boldsymbol{\varepsilon_t})(\lambda\eta_t + \boldsymbol{\varepsilon_t})`]$

$= \mathbf{E}\ (\boldsymbol{\lambda}\eta_t\eta_t`\boldsymbol{\lambda}` + \boldsymbol{\lambda}\eta_t\boldsymbol{\varepsilon_t}` + \boldsymbol{\varepsilon_t}\boldsymbol{\lambda}`\eta_t` + \boldsymbol{\varepsilon_t}\boldsymbol{\varepsilon_t}`)$

$= \boldsymbol{\lambda}\mathbf{E}\ (\eta_t\eta_t`)\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon})$

$= \boldsymbol{\lambda}\mathbf{E}[(\boldsymbol{\gamma'X_t} + \zeta_t)(\boldsymbol{\gamma'X_t} + \zeta_t)`]\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon}$

$= \boldsymbol{\lambda}\mathbf{E}(\boldsymbol{\gamma'X_tX_t`\gamma} + \boldsymbol{\gamma'X_t}\zeta_t` + \zeta_t\boldsymbol{X_t`\gamma} + \zeta_t\zeta_t`)\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon}$

$= \boldsymbol{\lambda}(\boldsymbol{\gamma'\Phi\gamma} + \boldsymbol{\psi})\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon}$ ... ... ... . $(\boldsymbol{a})$

$\mathbf{Cov\ (X_t\ y_t) = E\ (X_t\ y_t`) = E}\ [\boldsymbol{X_t}(\lambda\eta_t + \boldsymbol{\varepsilon_t})`]$

$= \boldsymbol{E}\ (\boldsymbol{X_t}\eta_t`\boldsymbol{\lambda}` + \boldsymbol{X_t\varepsilon_t}` + \boldsymbol{\varepsilon_t}\eta_t`\boldsymbol{\lambda}` + \boldsymbol{\varepsilon_t\varepsilon_t}`)$

$= \mathbf{E}\ (\boldsymbol{X_t}\eta_t`)\boldsymbol{\lambda}`$

$= \mathbf{E}[(\boldsymbol{X_t})(\boldsymbol{\gamma'X_t} + \zeta_t)`]\boldsymbol{\lambda}`$

$= \mathbf{E}(\boldsymbol{X_tX_t`\gamma} + \boldsymbol{X_t}\zeta_t`)\boldsymbol{\lambda}`$

$= \boldsymbol{\Phi\gamma\lambda}`$ ... ... $(\boldsymbol{b})$

$\mathbf{Cov\ (y_t\ X_t) = E\ (y_t\ X_t`)} = (\boldsymbol{\Phi\gamma\lambda}`)`$

$= \boldsymbol{\lambda\gamma`\Phi}$ ... ... . $(\boldsymbol{c})$

$\mathbf{Var\ (X_t) = E\ (X_t\ X_t`) = \Phi}$ ... . . $(\boldsymbol{d})$

**Appendix (B): Covariance Matrix of the Long run Part**

Long run equations with I(0) and I(1) causes for EMIMIC model are $\boldsymbol{y_t} = \boldsymbol{\lambda}\eta_t + \boldsymbol{\varepsilon_t}$ and $\eta_t = \boldsymbol{\gamma'X_t} + \boldsymbol{\tau`V_t} + \zeta_t$ . Thus the structure of covariance matrix is given as

$$\sum=\begin{bmatrix} \boldsymbol{Var(y_t)} & & \\ \boldsymbol{Cov(X_t, y_t)} & \boldsymbol{Var(y_t)} & \\ \boldsymbol{Cov(V_t, y_t)} & \boldsymbol{Cov(V_t, X_t)} & \boldsymbol{Var(V_t)} \end{bmatrix}$$

Before applying the expectation operation remember the assumption of MIMIC model as discuss in appendix (A). so the sub-matrices are :

Now we apply the expectation in order to derive the both variance and covariance between the observable variable, it follows the following steps:

**Var ($y_t$) = E ($y_t$ $y_t$`)** = $\mathbf{E}[(\lambda\eta_t + \boldsymbol{\varepsilon_t})(\lambda\eta_t + \boldsymbol{\varepsilon_t})`]$

$= \mathbf{E}\ (\boldsymbol{\lambda}\eta_t\eta_t`\boldsymbol{\lambda}` + \boldsymbol{\lambda}\eta_t\boldsymbol{\varepsilon_t}` + \boldsymbol{\varepsilon_t}\boldsymbol{\lambda}`\eta_t` + \boldsymbol{\varepsilon_t}\boldsymbol{\varepsilon_t}`)$

$= \boldsymbol{\lambda}\mathbf{E}\ (\eta_t\eta_t`)\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon})$

$= \boldsymbol{\lambda}\mathbf{E}[(\boldsymbol{\gamma'X_t} + \boldsymbol{\tau}`\boldsymbol{V_t} + \zeta_t)(\boldsymbol{\gamma'X_t} + \boldsymbol{\tau}`\boldsymbol{V_t} + \zeta_t)`]\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon}$

$= \boldsymbol{\lambda}\mathbf{E}(\boldsymbol{\gamma'X_tX_t}`\boldsymbol{\gamma} + \boldsymbol{\gamma'X_tV_t}`\boldsymbol{\tau} + \boldsymbol{\tau}`\boldsymbol{V_tX_t}`\boldsymbol{\gamma} + \boldsymbol{\tau}`\boldsymbol{V_tV_t}`\boldsymbol{\tau})\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon}$

$= \boldsymbol{\lambda}(\boldsymbol{\gamma'\Phi_1\gamma} + \mathbf{2}\boldsymbol{\gamma}`\boldsymbol{N\tau} + \boldsymbol{\tau}`\boldsymbol{\Phi_2\tau})\boldsymbol{\lambda}` + \boldsymbol{\Theta_\varepsilon} \ldots\ldots\ldots (\boldsymbol{a}`)$

**Cov ($X_t$ $y_t$)** = $\mathbf{E}\ (\mathbf{X_t}\ \mathbf{y_t}`) = \mathbf{E}\ [\boldsymbol{X_t}(\lambda\eta_t + \boldsymbol{\varepsilon_t})`]$

$= \boldsymbol{E}\ (\boldsymbol{X_t}\eta_t`\boldsymbol{\lambda}` + \boldsymbol{X_t}\boldsymbol{\varepsilon_t}` + \boldsymbol{\varepsilon_t}\eta_t`\boldsymbol{\lambda}` + \boldsymbol{\varepsilon_t}\boldsymbol{\varepsilon_t}`)$

$= \mathbf{E}\ (\boldsymbol{X_t}\eta_t`)\boldsymbol{\lambda}`$

$= \mathbf{E}[(\boldsymbol{X_t})(\boldsymbol{\gamma'X_t} + \boldsymbol{\tau}`\boldsymbol{V_t} + \zeta_t)`]\boldsymbol{\lambda}`$

$= \mathbf{E}(\boldsymbol{X_tX_t}`\boldsymbol{\gamma} + \boldsymbol{X_tV_t}`\boldsymbol{\tau})\boldsymbol{\lambda}`$

$= (\boldsymbol{\Phi_1\gamma} + \boldsymbol{N\tau})\boldsymbol{\lambda}` \ldots\ldots (\boldsymbol{b'})$

**Cov ($V_t$ y`$_t$)** = $\mathbf{E}[V_t(\lambda\eta_t + \varepsilon_t)`]$

$= \boldsymbol{E}(V_t\eta_t`)\boldsymbol{\lambda}` = \mathbf{E}[(\boldsymbol{V_t})(\boldsymbol{\gamma'X_t} + \boldsymbol{\tau}`\boldsymbol{V_t} + \zeta_t)`]\boldsymbol{\lambda}`$

$= \boldsymbol{E}(\boldsymbol{V_tX_t}`\boldsymbol{\gamma} + \boldsymbol{V_tV_t}`\boldsymbol{\tau})\boldsymbol{\lambda}` = (\boldsymbol{N}`\boldsymbol{\gamma N} + \boldsymbol{\Phi_2\tau})\boldsymbol{\lambda}` \ldots\ldots (\boldsymbol{c'})$

**Var ($X_t$)** = $\mathbf{E}\ (\mathbf{X_t}\ \mathbf{X_t}`) = \boldsymbol{\Phi_1} \ldots\ldots (\boldsymbol{d'})$

**Var ($V_t$)** = $\mathbf{E}\ (\mathbf{V_t}\ \mathbf{V_t}`) = \boldsymbol{\Phi_2} \ldots\ldots (\boldsymbol{e'})$

**cov ($V_t$ , $X_t$)** = $\mathbf{E}\ (\mathbf{V_t}\ \mathbf{X_t}`) = \boldsymbol{N'} \ldots\ldots (\boldsymbol{f'})$

Finally we obtain the following covariance matrix

$$\sum=\begin{bmatrix} \boldsymbol{\lambda\Phi_2(\gamma'\Phi\gamma + 2\gamma`N\tau + \tau`\tau)\lambda`} + \boldsymbol{\Theta_\varepsilon} & & \\ \boldsymbol{(\Phi_1\gamma + N\tau)\lambda`} & \boldsymbol{\Phi_1} & \\ \boldsymbol{(N`\gamma + \Phi_2\tau)\lambda`} & \boldsymbol{N`} & \boldsymbol{\Phi_2} \end{bmatrix}$$

**Appendix (C): Covariance Matrix of the Short Run Part**

Before applying the expectations on matrix remember the above stated assumption of the MIMIC model. Short run part of the EMIMC model's equations are $\Delta \boldsymbol{y_t} = \boldsymbol{\lambda}\Delta\eta_{\boldsymbol{t}} + \boldsymbol{\varepsilon_t}$ and $\Delta\eta_{\boldsymbol{t}} = \boldsymbol{\alpha}\Delta \boldsymbol{X_t} + \boldsymbol{\beta V_t} + \boldsymbol{kZ_{t-1}} + \zeta_{\boldsymbol{t}}$. So model's general covariance structure is given as:

$$\sum = \begin{bmatrix} \mathbf{Var}(\Delta \mathbf{y}_t) & & & \\ \mathbf{Cov}(\Delta \mathbf{X}_t, \Delta \mathbf{y}_t) & \mathbf{Var}(\Delta \mathbf{X}_t) & & \\ \mathbf{Cov}(\mathbf{V}_t, \Delta \mathbf{y}_t) & \mathbf{Cov}(\mathbf{V}_t, \Delta \mathbf{X}_t) & \mathbf{Var}(\mathbf{V}_t) & \\ \mathbf{Cov}(\mathbf{Z}_{t-1}, \Delta \mathbf{y}_t) & \mathbf{Cov}(\mathbf{Z}_{t-1}, \Delta \mathbf{X}_t) & \mathbf{Cov}(\mathbf{Z}_{t-1}, \mathbf{V}_t) & \mathbf{Var}(\mathbf{Z}_{t-1}) \end{bmatrix}$$

Now we derive the model parameter by applying expectations and we get the following matrix that represents the model parameters.

$$\sum = \begin{bmatrix} \boldsymbol{\lambda}(\boldsymbol{\alpha}'\boldsymbol{\Phi_3}\boldsymbol{\alpha} + 2\boldsymbol{\alpha}\grave{}\boldsymbol{M}\boldsymbol{\beta} + \boldsymbol{\beta}\grave{}\boldsymbol{\Phi_2}\boldsymbol{\beta} + \mathbf{k}\grave{}\Omega\mathbf{k}) & & & \\ (\boldsymbol{\Phi_3}\boldsymbol{\alpha} + \boldsymbol{M}\grave{}\boldsymbol{\beta})\boldsymbol{\lambda}\grave{} & \boldsymbol{\Phi_3} & & \\ (\boldsymbol{M}\boldsymbol{\alpha} + \boldsymbol{\Phi_2}\boldsymbol{\beta})\boldsymbol{\lambda}\grave{} & \mathbf{M} & \boldsymbol{\Phi_2} & \\ (\boldsymbol{\psi}\boldsymbol{\lambda}\grave{}\boldsymbol{k} + \boldsymbol{\psi})\boldsymbol{\lambda}\grave{} + \boldsymbol{\Theta_\varepsilon} & \mathbf{0} & \mathbf{0} & \Omega \end{bmatrix}$$

**Magnitude of Shadow Economy as Percentage of Official Economy**

| Years | MIMIC Results | ARDL Results | Years | MIMIC Results | ARDL Results |
|---|---|---|---|---|---|
| 1974 | 51.5% | 52.1% | 1995 | 38.4% | 59.0% |
| 1975 | 53.8% | 46.6% | 1996 | 37.8% | 56.9% |
| 1976 | 49.7% | 47.6% | 1997 | 35.4% | 67.6% |
| 1977 | 47.6% | 51.3% | 1998 | 34.5% | 68.4% |
| 1978 | 49.1% | 53.5% | 1999 | 35.6% | 59.8% |
| 1979 | 46.5% | 50.0% | 2000 | 35.9% | 51.2% |
| 1980 | 50.7% | 52.9% | 2001 | 35.5% | 54.4% |
| 1981 | 48.5% | 49.9% | 2002 | 36.8% | 47.6% |
| 1982 | 47.5% | 48.9% | 2003 | 35.9% | 48.3% |
| 1983 | 46.2% | 54.8% | 2004 | 34.4% | 44.5% |
| 1984 | 45.6% | 55.8% | 2005 | 32.9% | 48.1% |
| 1985 | 42.6% | 48.9% | 2006 | 31.7% | 40.8% |
| 1986 | 41.8% | 51.4% | 2007 | 33.9% | 40.0% |
| 1987 | 42.6% | 48.4% | 2008 | 32.9% | 52.2% |
| 1988 | 45.0% | 45.7% | 2009 | 30.8% | 54.7% |

| | | | | | |
|---|---|---|---|---|---|
| 1989 | 46.1% | 46.2% | 2010 | 33.1% | 59.0% |
| 1990 | 45.1% | 48.6% | 2011 | 31.2% | 50.8% |
| 1991 | 42.9% | 50.0% | 2012 | 33.0% | 51.5% |
| 1992 | 41.2% | 53.0% | 2013 | 29.4% | 45.5% |
| 1993 | 39.0% | 59.0% | 2014 | 29.4% | 43.8% |
| 1994 | 35.7% | 61.6% | 2015 | 28.7% | 47.2% |