The size of informal economy in Pakistan

Arby, Muhammad Farooq; Malik, Muhammad Jahanzeb and Hanif, Muhammad Nadim
State Bank of Pakistan

07. May 2010

SBP Working Paper Series

No. 33 May, 2010

The Size of Informal Economy in Pakistan

Muhammad Farooq Arby
Muhammad Jahanzeb Malik
Muhammad Nadim Hanif

STATE BANK OF PAKISTAN

SBP Working Paper Series

Editor: Riaz Riazuddin

The objective of the SBP Working Paper Series is to stimulate and generate discussions, on different aspects of macroeconomic issues, among the staff members of the State Bank of Pakistan. Papers published in this series are subject to intense internal review process. The views expressed in the paper are those of the author(s) and do not necessarily reflect those of the State Bank of Pakistan.

Price per Working Paper

Pakistan:	Rs 50 (inclusive of postage)
Foreign:	US$ 20 (inclusive of postage)

Purchase orders, accompanied with cheques/drafts drawn in favor of State Bank of Pakistan, should be sent to:

Chief Spokesperson
External Relations Department,
State Bank of Pakistan,
I.I. Chundrigar Road, P.O. Box No. 4456,
Karachi 74000. Pakistan

For all other correspondence:

Editor,
SBP Working Paper Series
Research Department,
State Bank of Pakistan,
I.I. Chundrigar Road, P.O. Box No. 4456,
Karachi 74000. Pakistan

Published by: Editor, SBP Working Paper Series, State Bank of Pakistan, I.I. Chundrigar Road, Karachi, Pakistan.

ISSN 1997-3802 (Print)
ISSN 1997-3810 (Online)

http://www.sbp.org.pk

Printed at the SBPBSC (Bank) – Printing Press, Karachi, Pakistan

The Size of Informal Economy in Pakistan

Muhammad Farooq Arby
Senior Economist, State Bank of Pakistan, farooq.arby@sbp.org.pk

Muhammad Jahanzeb Malik
Joint Director, State Bank of Pakistan, jahanzeb.malik@sbp.org.pk

Muhammad Nadim Hanif
Senior Economist, State Bank of Pakistan, nadim.hanif@sbp.org.pk

Abstract

This paper estimates the size of informal economy in Pakistan by using monetary approach with some modifications, electricity consumption approach and MIMIC model. Under monetary approach, we take care of the issue of the stationarity of variables and use autoregressive distributed lag (ARDL) model instead of simple OLS and add education as an additional factor affecting the size of informal economy along with some other technical improvements in the standard monetary models. The electricity consumption approach and MIMIC models are used for the first time in case of Pakistan. The results show that the informal economy in Pakistan has been about 30 percent of the total economy which declined considerably in 2000s. Currently, about 20 percent of the economic transactions are taking place in the informal sector.

JEL Codes**:** C10, E26

Key Words**:** Informal Economy, ARDL, MIMIC

Acknowledgment

We appreciate the comments and feedback given by Ali Choudhary, Moinuddin, Amin Lodhi and participants of a seminar in which this paper was presented.

1. Introduction

The term informal economy along with its other synonyms like black, shadow, underground economy etc. is yet to be defined in a precise mode. A common approach to define the informal economy is 'unmeasured and untaxed economic activity' taking place in a country. Wiegand (1992) gives the definition of the informal economy in historical perspectives and other concepts associated with this term. Irrespective of the way it is defined, it is obvious that the informal economy is poorly reflected in official measures of aggregate income and product of a country.

The national account compilers usually rely on secondary data sources of economic activities and information about documented transactions. Their inability to entrap concealed transactions and lack of a precise definition makes the estimation of informal economy difficult. Yet it is important to have some idea of the size of the informal economy because its existence renders the true state of the economy scantily reflected by the official values of economic indicators of growth, consumption, investment, unemployment etc. Given that these statistics are employed to formulate economic policies and to launch social welfare programs and development projects, inaccurate figures may lead to inappropriate policy responses and inefficient use of resources.

This is why the economists have developed a number of direct and indirect methods for estimating the size of the informal economy (see for example Schneider and Enste (2000) for detailed description of different approaches for its measurement). The direct methods are microeconomic in nature and use either voluntary survey data or the results from tax audits to get estimates of unmeasured economic activity. Voluntary surveys typically ask respondents to reveal their incomes, expenditure and labor status. This method has been criticized for its sensitivity to how the questions are posed, and its confidence in the respondents' willingness to truthfully reveal their income. Tax audit-based measures define the magnitude of the informal economy as the difference between the income declared in tax returns and the income actually found after an audit. A potential problem in extrapolating to the national economy is that audits are usually nonrandom and, hence, may not be representative (Perry et al (2007)).

Indirect methods are macroeconomic in nature and include (a) monetary approach that exploits the facts that monetary aggregates are usually measured with fair accuracy and almost all economic transactions are undertaken by using some form of money, (b) physical input approach by which electricity consumption is regarded as the single best physical indicator of overall economic activity

with a unit elasticity between the two and (c) Multiple Indicator–Multiple Cause (MIMIC) model that postulates that magnitude of the unofficial economy can be modeled as a latent or index variable.

The indirect methods also have their weaknesses. The MIMIC model has been criticized on the ground that its results are sensitive to transformations of the data, to the units of measurement, and to the sample used. Another criticism is that no theory is used in order to determine which variables to include as indicators or as causes. Moreover, the estimate of the informal economy through MIMIC model relies on physical input or monetary methods for the initial levels, which make it vulnerable to the criticisms of these two methods. The physical input method has been criticized on grounds that it does not consider technological progress over time which may change the dynamics of electricity consumption per unit of output. On the other hand, the monetary approach has also been criticized particularly for its assumption of a common velocity of money in the official and unofficial economies. Ahumada et al (2007 and 2008) criticize the monetary approach on two grounds: (a) the assumption of equality of income velocity of currency demand in formal and informal economy poses un-necessary restrictions on the model parameter of scale variable; (b) short-run model of monetary approach requires initial condition and thus suffers from arbitrariness. However, both these observations can be addressed if a long-run model is estimated and instead of using level of currency, the currency to M2 ratio is used in the model.

A common criticism to all these methods is lack of theoretical foundations from which such models can be derived. However, despite the criticism these methods are still being used to estimate the size of the informal economy in both developed and developing economies because no alternative is yet available. This paper employs all the three commonly used methods in case of Pakistan; earlier studies on Pakistan use only monetary approach[1].

The rest of the paper has been organized as follows: the next section gives descriptions of the three methods and discusses the data issues. Section 3 presents the results and the last section concludes the paper.

2. Methodology

2.1 Monetary Approach

The monetary approach of estimating informal economy was developed on the Cagan's (1958) proposition that higher tax rates induce people to use currency for transactions to avoid tax reporting.

[1] We have also examined household income and expenditure surveys (HIES) in Pakistan for a direct estimate of the informal economy; however, it is found that per capita income as per HIES is lower than recorded per capita GDP.

The pioneering work in this area include Gutmann (1977), Feige (1979), and Tanzi (1983). 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 that demand. The key assumptions of this approach are (a) informal economic activities are the direct consequence of high taxes; (b) such transactions are mainly carried out by currency (and thus the overall currency in circulation in the economy has two components: currency used for informal economic transactions and formal transactions; and (c) the transaction velocity of money in both the informal and formal economies is the same. Tanzi estimates the following model:

$$In\ (CM) = a_o + a_1\ In\ (1+TW) + a_2\ In\ (WS/NI) + a_3\ In\ (R) + a_4\ In\ (PY) + \varepsilon \qquad (1)$$

CM is the ratio of currency in circulation to *M2*, *TW* is weighted average tax rate[2], *WS/NI* is share of wages and salaries in national income, *R* is the rate of interest paid on time deposits, and *PY* is real per capita income.

However, there are two serious problems with this model: (a) the model is estimated using the ordinary least square (OLS) without considering time series properties of the included variables. If the variables are non-stationary[3] then the OLS estimates of the parameters are spurious, (b) the ratio CM has been used in log form that poses a disaggregation problem when the results are used to bifurcate legal and illegal money (in Tanzi's terminology). This point is explained as follows.

The method usually proceeds with first getting an overall predicted value of Ln(C/M2) on the basis of the above equation and then estimating a predicted level of currency holding ($\hat{C}$). Next, the equation is solved in the same way, assuming that the tax variable is zero while the coefficients of the other variables remain unchanged. This gives an estimate of currency holding in a tax free environment (C_f). Then the currency holding for informal transactions (C_i), induced by taxes is calculated as a difference between $\hat{C}$ and C_f. In principle, one may also proceed through a slightly different route by directly estimating C_i with setting all the coefficients equal to zero except a_1, i.e. directly extracting tax induced currency ratio from the equation (1). Although technically this route is also correct, the results obtained from the two routes are quite different. Taking the work of Tanzi (1983) as an example, we have estimated the informal economy for US by using the alternate route of calculation from the same set of data as used by Tanzi and found that the informal economy was 1.6 to 3.4 percent of the recorded GNP during 1930 to 1980 as compared with Tanzi's results of 0.6 to 8

[2] He also uses income tax rate in an alternate specification.
[3] These variables are found in fact non-stationary in case of Pakistan (results of stationarity test are available on request.)

percent. There is no theoretical reason of which route should be preferred. However, this problem can be avoided if currency to M2 ratio is used in the model without taking its log.

In case of Pakistan, a number of attempts have been made to estimate the informal economy, all through monetary approach. All the papers presented long-run estimates of black economy and used currency ratio (with some variations by different studies). Thus these studies are not subject to Ahumada et al (2007 & 2008) critiques. However, they all overlooked the problem of stationarity in the variables of their models. Thus the results obtained by these studies by using simple OLS may not necessarily be legitimate. Some studies like Ahmed and Qazi (1995), Aslam (1998), Ahmed and Haider (2008) used currency ratio in log form which adds another problem to their estimates as discussed above.

Some studies add foreign currency account (Aslam (1998), Kemal (2007) and bearer bonds (Ahmed and Haider 2008) to currency in circulation with the argument that these instruments are also used as medium of exchange in the informal economy. It is to be noted that both these instruments are primarily used as savings instruments or store of value by both formal and informal income earners; this may be a reason that inclusion of these variables brings insignificant changes in the results of these studies.

In this paper, we use a modified version of the monetary approach through autoregressive distributed lag (ARDL) model suggested by Pesaran and Shin (1999) and Pesaran et al (2001). No such attempt has so for been made for Pakistan. The ARDL model helps us to address some of the problems associated with Tanzi approach. It allows the use of both stationary and nonstationary variables in one model and can also produce long run relationship. Pesaran and Shin (1999) showed that ARDL-based estimators are super-consistent, and valid inferences on the long-run parameters can be drawn using the standard normal asymptotic theory. Similarly, they also find that appropriate modification of the orders of the ARDL model is sufficient to simultaneously correct residual serial correlation and problem of endogenous regressors.

We establish a long run cointegrating relationship between the currency ratio and other related variables and then used it to deduce the size of the informal sector in Pakistan's economy. We have made some modifications to the standard monetary model of estimating informal economy, and have instituted a relationship between the currency to M2 ratio as a dependent variable and tax to GDP ratio, a proxy for financial sector development, market interest rate, and a proxy for higher education as key determinants, as elucidated below:

a) The tax to GDP ratio is expected to affect currency ratio positively as argued by Tanzi (1980);
b) We expect that with the financial sector development, people economize the use of currency (a zero return monetary unit) and switch to other financial instruments for payments;
c) The expected effect of market interest rate on the currency ratio is the same, i.e. negative as suggested by Tanzi (1980);
d) We also include higher education in this model because in a developing country like Pakistan, all the informal activities are not necessarily driven only by taxes. Instead, lack of higher education also is a hurdle in the way of the formal economy. Thus we expect that level of higher education has a negative impact on the informal economy.

An ARDL representation of this relationship is formulated as follows:[4]

$$\Delta CM_t = \lambda_0 + \lambda_1 CM_{t-1} + \lambda_2 T_{t-1} + \lambda_3 F_{t-1} + \lambda_4 R_{t-1} + \lambda_5 E_{t-1} + \sum_{1}^{k} \alpha_{1i} \Delta CM_{t-i} + \sum_{0}^{k} \alpha_{2i} \Delta T_{t-i} + \sum_{0}^{k} \alpha_{3i} \Delta F_{t-i} + \sum_{0}^{k} \alpha_{4i} R_{t-i} + \sum_{0}^{k} \alpha_{4i} E_{t-i} + \varepsilon_t \quad (2)$$

CM_t = currency in circulation to M2 ratio

T_t = ratio of total taxes to nominal GDP

F_t = financial development indicator represented by the ratio of monetary liabilities of the banking system (excluding currency in circulation) to nominal GDP

R_t = market interest rate

E_t = level of higher education proxied by enrolment in professional colleges and universities[5]

The data for all the variables except total taxes and enrolment have been obtained from International Financial Statistics (online). The series of taxes and the enrolment have been obtained from Pakistan Economic Survey.[6]

The presence of a valid long-run relationship is tested by Wald coefficient restriction test with null hypothesis of $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 0$. The calculated F-statistic in this test is compared with the band of critical values tabulated by Pesaran et al (2001); they tabulated two sets of appropriate critical values for different number of regressors. One set assumes that all the variables are integrated of order zero, i.e. I(0) and another set assumes that they all are I(1). A computed F-statistic higher than the

[4] In some alternative specifications, we also included real per capita income as regressors in the model, however, no significant relationship was found between the currency demand and this variable.
[5] We has also used enrolment in some other levels of educational institutions like primary, secondary, college and all institutions, however, none was found significant.
[6] The data from IFS was retrieved in March 2009. The data on monetary variables are 12-month averages for every financial year (July-June). IFS does not report series for taxes prior to 1990.

upper level of the band implies the presence of a cointegrating relationship among the variables. With the approval of the bound test the following long-run model is deduced from (2):

$$\hat{C}M_t = \beta_0 + \beta_2 T_t + \beta_3 F_t + \beta_4 R_t + \beta_5 E_t \qquad (3)$$

$\beta_i = -\hat{\lambda}_i / \hat{\lambda}_1$ for i = 0, 2, 3, 4, 5.

Given the (3) and assumptions of the model as discussed earlier, the informal economy as a ratio to the total size of the economy (formal *plus* informal) is worked out as follows.

$$\mu_t = \frac{Y_i}{Y} = \frac{\beta_2 T + \beta_5 E_t}{m_t} \qquad (4)$$

Y_i is GDP in informal economy, Y is the total size of the economy and m_t is the ratio of *M1* definition of monetary aggregates to *M2*.

2.2 Electricity Consumption Approach

The electricity consumption approach has been presented by Kauffman and Kaliberda (1996) while drawing from the results of a study by Dobozi and Pohl (1995). Kauffman and Kaliberda argue that electric-power consumption is regarded as the single best physical indicator of economic activities in a country. Overall economic activity and electricity consumption have been empirically observed throughout the world to move in lockstep with electricity / GDP elasticity usually close to one. By exploiting this relationship, one can have a proxy measurement for the overall economy; an estimate of the unmeasured GDP can then be obtained by subtracting official GDP from the estimated overall GDP[7]. The following steps are involved in estimating the informal GDP for Pakistan through this approach.

1) Electricity consumption has been obtained from Pakistan Economic Survey FY07 (Supplement) and FY08 for years 1971-72 onward. Assuming the unitary elasticity, the growth in electricity consumption is equal to the growth in the total real GDP.[8]

[7] The official GDP in Pakistan also records economic activities taking place in agriculture and small and household manufacturing sectors which are though informal in the sense that they are poorly documented and their contribution to national exchequer is minimal; yet the Federal Bureau of Statistics is able to estimate their value addition through production approach. Therefore, in electricity approach when we deduct officially measured GDP from overall GDP, we get unmeasured GDP that may not necessarily match our estimates of informal economy through other approaches.

[8] The growth series is smoothed by using Hodrick-Prescott (1997) filter.

2) Having growth rates with our hand, we need an initial value of total real gross domestic product to derive a time series of absolute numbers. For this purpose, we assume that the official GDP of the year 1974-75 reflected all the economic activities of the year. This assumption does not seem implausible when we consider the facts that (a) more than one third of the GDP during mid 1970s consisted of agriculture which was though informal, fairly correctly estimated and (b) most of other businesses in goods and services were nationalized by the government. The pattern of the time series of the overall GDP does not change with taking different initial values.

3) From the overall GDP, we deduct the recorded GDP to obtain an estimate of the unmeasured GDP. For this exercise we have used GDP data estimated by Arby (2008) which is the only source of consistent time series of real GDP since 1970s at a single base; the official estimates are available only for 1999-00 onward.

2.3 MIMIC Model

The Multiple Indicators and Multiple Causes (MIMIC) model is a particular form of Linear Independent Structural Relationship models (LISREL) whereby the informal economy is taken as a latent variable which on the one hand caused by a set of variables and affects other variables on the other. The MIMIC model has been estimated by structural equation models (SEM) that give us an advantage of using all the information contained in the covariance matrix as compared to conventional regression analysis where only the information contained in the variance of variables is exploited. The main focus of SEM is to estimate a covariance matrix that fits the sample covariance matrix. The closeness of two matrices is considered as an indication of good approximation of relationships between the variables that define it. In other words in this methodology we minimize the distance between the observed and model employed covariance matrix (Alanon and Gomez-Antonio, 2005 and Buehn and Schneider, 2008).

There are two components of this model: a structural model which specifies causal relationship(s) between unobservable variable(s) and their causes; and a measurement model that relates unobservable variable(s) to the indicators. Several combinations of causes and indicators were tested to arrive upon a good model in terms of diagnostics. The model finally selected consists of three causes of the informal economy including tax/GDP ratio (T), M2/GDP ratio (F) and the regime durability (P); and two indicators including currency in circulation (as ratio to M2) and growth in electricity consumption. All the variables are in the form of deviations from mean.

The structural equation of the MIMIC model is the following.

$$L = \gamma' X + v \quad (5)$$

Where L is the latent variable, X is the matrix of cause variables including T, F and P in deviation from mean, γ is a vector of parameters, and v is a stochastic error term with zero mean and constant variance.

The set of measurement equations in which latent variable linearly determines observable endogenous variables (the indicators) is the following:

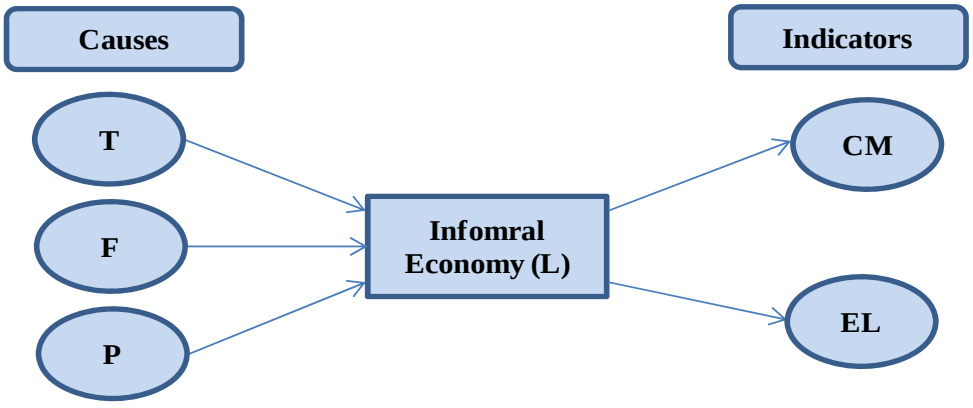
$$Y = \lambda L + e \quad (6)$$

Where Y is the vector of indicators including currency in circulation to M2 ratio and growth in electricity consumption as defined in the preceding section, λ is vector of scalars, and e is the vector of two independent stochastic error terms corresponding to respective equation of each indicator. The vector e is also independent of v of the structural equation and variance of e is a 2 × 2 diagonal matrix. By incorporating (5) into (6) we obtain:

$$Y = \lambda [\gamma' X + v] + e \quad (7)$$

The model is depicted by a diagram linking the latent variable to causes and indicators (Figure 1).

Figure 1: Relationship of Latent Variable with its Causes and Indicators



EL represents growth in electricity consumption and *CM* currency in circulation to M2 ratio; other variables have already been defined.

In order to estimate the model, we need to normalize one of two λ's by setting it equal to 1 such that the effect of other indicator is quantified. We estimate the parameters of model using Maximum Likelihood method. By expressing the model (7) in terms of covariance following Buehn and Schneider (2008)], we proceed as follows.

$$\Sigma = \begin{bmatrix} \text{var}(y_t) & \text{cov}(y_t, x_t) \\ \text{cov}(x_t, y_t) & \text{var}(x_t) \end{bmatrix} \tag{8}$$

$$= E\left[\begin{bmatrix} y_t \\ x_t \end{bmatrix}\begin{bmatrix} y_t \\ x_t \end{bmatrix}'\right]$$

$$E(y_t y_t') = \lambda(\gamma'\phi\gamma + \psi)\lambda' + \Theta_\varepsilon$$

$$E(x_t y_t') = \phi\gamma\lambda'$$

$$E(y_t x_t) = \lambda\gamma'\phi$$

$$E(x_t x_t') = \phi$$

Thus we have,

$$\Sigma = \begin{bmatrix} \lambda(\gamma'\phi\gamma + \psi)\lambda' + \Theta_\varepsilon & \lambda\gamma'\phi \\ \phi\gamma\lambda' & \phi \end{bmatrix} \tag{9}$$

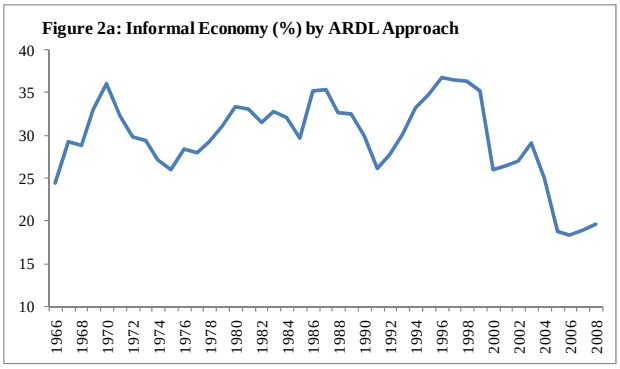
We need to estimate γ, λ, and covariances that produce that estimate of Σ that is as close as possible to sample cov (Y, X). In (9), Θ_ε is the covariance matrix of the error terms in the measurement model, ψ is the variance of the error term in the structural equation and ϕ is the covariance matrix of the causes.

3. Results

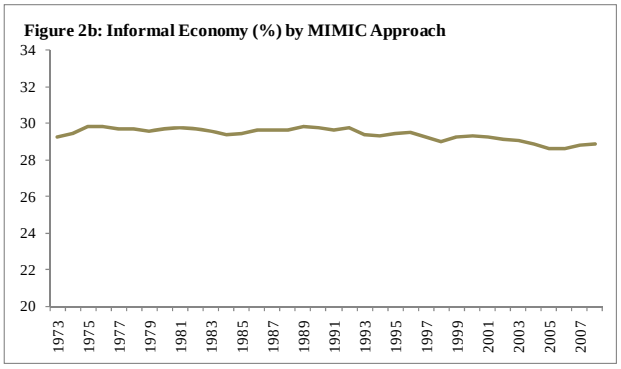
The result of ARDL model is reported in Annexure 1a. We have estimated the model for k = 1 to 4 lags, and selected the model with k = 1 on the basis of minimum Akaike information criterion. The long-run relationship of currency ratio with other variables is deduced as follows:

$$\hat{C}M_t = 37.01 + 1.682\ T_t - 0.675\,F_t - 1.062\,R_t - 0.117\,E_t \tag{10}$$

The computed F-statistic is 8.13 which is above the upper level of the bound test (the bound is 4.93 – 5.73 for lag 1 at 5 percent critical level).

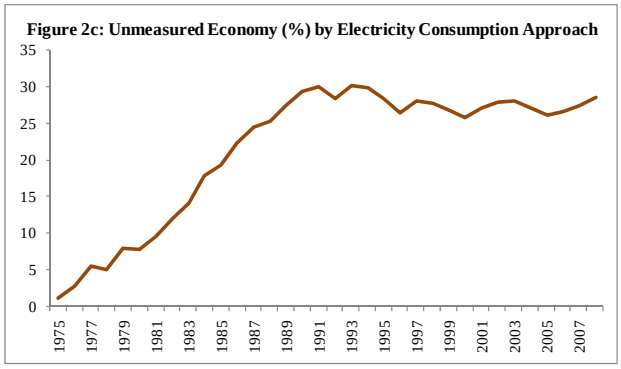


The estimates of informal economy on the basis of the ARDL model and other methods are given in Table 1. The ARDL model shows that the informal economy has increased its share in the Pakistan's economy until end of 1990s and has a declining trend since then. It was below 30 percent during 1960s and 1970s, increased to 33 percent in 1990s and then declined to 23 percent in current decade of 2000s (Figure 2a). Not surprisingly, the informal economy was highest in the decade of 1990s – a period of the most fragile and weak political regimes. It is only 2000s, when although the economy grew with a reasonable average growth rate of 6 percent, the informality declined attributable to activism by the tax authority for documentation, political stability and improved governance.



The results of MIMIC model, on the other hand show that the informal economy has been around 30% of the total economy in Pakistan (Figure 2b). Although the extent of informal economy as deduced through the MIMIC model may hinge upon initial values, the growth path remains steady irrespective of the initial values. It is evident that ratio of informal economy to the recorded economy has been fairly stable in Pakistan which reflects the informal economy has grown with almost the same rate as the recorded economy.[9]

The electricity approach, on the other hand, shows that the extent of the unmeasured economy was less 5 percent during 1970s which then increased sharply until 1990s and remained stagnant after that (Figure 2c). However, this approach may not reflect the actual performance of the economy as



[9] As a corollary to this result, it can be argued that despite the official numbers of absolute GDP may be underestimated; the growth rates of official GDP may be close enough to the overall GDP.

the official numbers of electricity consumption do not incorporate self generation of electricity by economic agents which boomed in mid 1990s onward due to crisis in official sector of power generation and distribution in Pakistan.

The results of this study are generally close to those obtained by other studies on Pakistan (see Annexure 2 for a comparison). However, there are some differences in growth paths as derived by different studies. Both the Kemal (2007) and this study show a rising trend up to the end of 1990s while Ahmed and Haider (2008) shows a declining trend. However, we expect that this study captures the phenomenon more precisely as it addresses most of the weaknesses of earlier studies as discussed above.

Table 1: Estimates of Informal Economy (% of Total Economy)

	ARDL model	MIMIC	Electricity Consumption Approach*		ARDL model	MIMIC	Electricity Consumption Approach*
1966	24.4			1991	26.1	29.7	30.1
1967	29.2			1992	27.7	29.8	28.5
1968	28.8			1993	30.1	29.4	30.2
1969	33.1			1994	33.3	29.4	30.0
1970	36.0			1995	34.8	29.5	28.4
1971	32.3			1996	36.8	29.5	26.5
1972	29.8			1997	36.4	29.3	28.1
1973	29.3	29.3		1998	36.4	29.1	27.7
1974	27.1	29.5		1999	35.2	29.3	26.8
1975	25.9	29.8	1.2	2000	26.0	29.3	25.9
1976	28.4	29.8	2.8	2001	26.3	29.3	27.1
1977	27.9	29.7	5.5	2002	27.0	29.2	28.0
1978	29.2	29.7	5.1	2003	29.0	29.1	28.2
1979	31.1	29.6	7.9	2004	24.9	28.9	27.1
1980	33.3	29.7	7.9	2005	18.7	28.6	26.2
1981	33.1	29.8	9.6	2006	18.3	28.7	26.7
1982	31.6	29.7	12.0	2007	18.9	28.8	27.5
1983	32.8	29.6	14.2	2008	19.6	28.9	28.6
1984	32.1	29.4	17.9				
1985	29.6	29.4	19.4	***Averages***			
1986	35.2	29.7	22.5	1960s	28.9		
1987	35.4	29.6	24.5	1970s	29.7	29.6	4.5
1988	32.7	29.6	25.3	1980s	32.8	29.6	18.1
1989	32.5	29.8	27.5	1990s	32.7	29.5	28.6
1990	30.0	29.8	29.5	2000s	23.2	29.0	27.3

* Unmeasured GDP as % of total GDP.

4. Conclusion

Although there is neither a precise definition of the informal economy nor a method exists that can measure it precisely, it is extremely important to have some idea of its size because its existence renders the official values of macroeconomic indicators poor reflection of the of true state of the economy. Therefore, the economists have developed a number of direct and indirect methods to

estimate the size the informal economy. The most commonly used method is monetary approach which assumes that the informal transactions are carried out by using currency, the transaction velocity of money is equal in recorded and informal economies and that main motive of being informal is tax avoidance. In a popular variant of this approach, a currency demand model is estimated with tax variable as one of the regressors. The level of currency used in informal economy is extracted as a partial impact of tax variable. However, the way in which the earlier studies construct and estimate the model has several shortcomings.

The present study addresses these shortcomings and uses autoregressive distributed lags (ARDL) technique to estimate currency demand equation and also includes a proxy for education as a negative factor for informal economy. We hypothesize that in a country like Pakistan, tax is not the only factor that induces the economic agents to be in informal but the lack of education also keeps them away from a formal and recordable way of economic activity. A negative estimated coefficient of the variable for education supports our hypothesis. We have used the estimated ARDL model to deduce a long run relationship between currency demand and other variables including tax to GDP ratio, proxy for financial development, proxy for education and interest rate. This relationship is then used to estimate the informal economy in Pakistan.

We have also used two other methods to estimate the size of the informal economy including electricity consumption approach and MIMIC model which have not been used earlier in case of Pakistan. The ARDL approach shows that the underground economy increased from less than 30 percent in 1960s to 33 percent in 1990s and then declined to 23 percent on average in 2000s. The electricity consumption approach shows that unmeasured economy increased from about 5 percent of total size of the economy in 1970s to 29 percent in 1990s and then declined to 27 percent in 2000s. The MIMIC model shows that the informal economy has been around 30 percent.

References

Ahmed, Mehnaz and Qazi Masood Ahmed (1995). "Estimation of the Black Economy of Pakistan through the Monetary Approach." *Pakistan Development Review*, 34:4, 791–807.

Ahmed, Qazi Masood and M. Haider Hussain (2008). "Estimation of the Black Economy of Pakistan through the Monetary Approach: A Case Study of Pakistan." *Economic Issues*, 13:1.

Ahumada, Hildegart, Facundo Alvaredo and Alfredo Canavese (2007). "The Monetary Method and the Size of the Shadow Economy: A Critical Assessment." *Review of Income and Wealth,* 53:2.

Ahumada, Hildegart, Facundo Alvaredo and Alfredo Canavese (2008). "The Monetary Method to Measure the Shadow Economy: The Forgotten Problem of the Initial Conditions." *Economics Letters*, 101, 97–99.

Alañón, A., and M. Gómez-Antonio (2005). "Estimating the Size of the Shadow Economy in Spain: A Structural Model with Latent Variables." *Applied Economics*, 37:9, 011-1025.

Arby, M Farooq (2008). "Some Issues in the National Income Accounts of Pakistan (Rebasing, Quarterly and Provincial Accounts and Growth Accounting)." *PhD Dissertation*, Pakistan Institute of Development Economics, Islamabad.

Aslam, Salman (1998). "The Underground Economy and Tax Evasion in Pakistan: Annual Estimates (1960-1998) and the Impact of Dollarisation of the Economy." *The Pakistan Development Review* 37:4, Part II (Winter) 621–631.

Buehn. A, and Schneider. F (2008). "MIMIC Models, Cointegration and Error Correction: An Application to the French Shadow Economy." *IZA Discussion paper* No. 3306

Cagan, Philip (1958). "The Demand for Currency Relative to Total Money Supply." *Journal of Political Economy,* 66, 303-328.

Dickey, D.A. and W.A. Fuller (1979). "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association*, 74, 427–431.

Dobozi, Istvan, and Pohl Gerhard (1995). “Real Output Decline in Transition Economies - Forget GDP, Try Power Consumption Data.” *Transition Newsletter*, 6.

Feige, E. L. (1979). “How Big is the Irregular Economy?” *Challenge,* 22, 5-13.

Guttman, P. M. (1977). “The Subterranean Economy.” *Financial Analyst Journal,* Nov.-Dec, 26-34.

Hodrick, R. J. and E. C. Prescott (1997). “Postwar U.S. Business Cycles: An Empirical Investigation.” *Journal of Money, Credit, and Banking*, 29, 1–16.

Kaufmann, Daniel and Aleksander Kaliberda (1996). “Integrating the Unofficial Economy into the Dynamics of Post-Socialist Economies: A Framework of Analysis and Evidence.” in B. Kaminski (ed.) *Economic Transition in Russia and the New States of Eurasia*. Armont, NY: M. E. Sharpe, Inc. (also available as *World Bank Policy Research Working Paper* No 1691, December 1996).

Kemal, M. Ali (2007). “Fresh Assessment of the Underground Economy and Tax Evasion in Pakistan: Causes, Consequences, and Linkages with the Formal Economy.” *PIDE working paper* No. 2007:13.

Perry, Guillermo E., William F. Maloney, Omar S. Arias, Pablo Fajnzylber, Andrew D. Mason, Jaime Saavedra-Chanduvi (2007). ”Informality: Exit and Exclusion.” The World Bank, Washington DC.

Pesaran, M. Hashem and Yongcheol Shin (1999). “An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis.” Chapter 11 in Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, Strom, S. (ed); Cambridge University Press, Cambridge.

Pesaran, M. Hashem, Yongcheol Shin, and Richard J. Smith (2001). “Bounds Testing Approaches to the Analysis of Level Relationships.” *Journal of Applied Econometrics*, 16, 289-326.

Schneider, Friedrich and Dominik H. Enste (2000). “Shadow Economies: Size, Causes, and Consequences.” *Journal of Economic Literature*, 38:1, 77-114.

Shabsigh, Ghiath (1995). “The Underground Economy: Estimation, and Economic and Policy Implications, The Case of Pakistan.” *IMF Working Papers*, WP/95/101.

Tanzi, Vito (1980). “The Underground Economy in the United States: Estimates and Implications.” *Quarterly Review*, 35, 427-53.

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

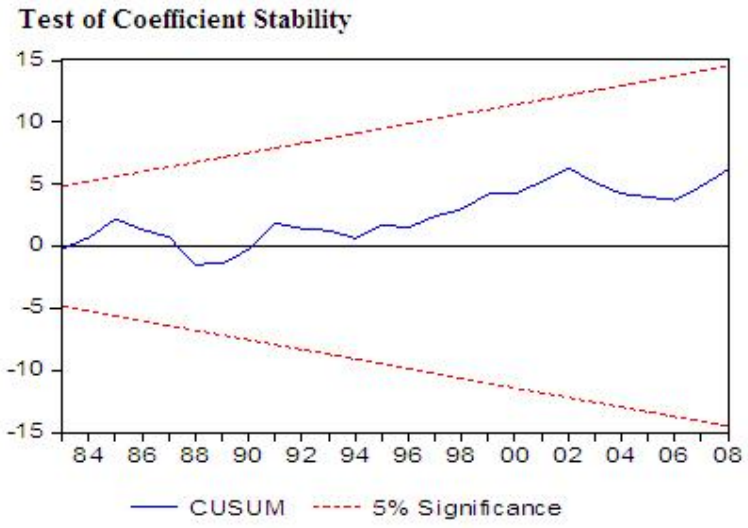
Wiegand, Bruce (1992). “Off the Books: A Theory and Critique of the Underground Economy.” Published by Rowman & Littlefield.

Annexure 1a: Results of ARDL Model

The model was estimated in EViews (v5) using ordinary least square. The estimated values of the parameters are the following.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Intercept	19.0802	4.542361	4.200504	0.0003
CM(-1)	-0.5155	0.091985	-5.60415	0
T(-1)	0.867144	0.198992	4.357685	0.0002
F(-1)	-0.34803	0.101478	-3.42959	0.002
R(-1)	-0.54746	0.11965	-4.5755	0.0001
E(-1)	-0.06049	0.045063	-1.34222	0.1911
ΔT	0.562054	0.195751	2.871264	0.008
ΔF	-0.61698	0.090162	-6.843	0
ΔR	-0.24625	0.111496	-2.20859	0.0362
ΔE	-0.59858	0.212965	-2.8107	0.0093
ΔCM(-1)	0.120674	0.127147	0.949092	0.3513
ΔT(-1)	0.060014	0.184081	0.326018	0.747
ΔF(-1)	0.20049	0.120632	1.661997	0.1085
ΔR(-1)	0.331716	0.13897	2.386954	0.0246
ΔE(-1)	0.054326	0.221229	0.245564	0.8079
R-squared	0.83561			
Breusch-Godfrey Serial Correlation LM Test:		F-statistic		0.26 (0.77)

The stability of the model was tested by the CUSUM test (Brown, Durbin, and Evans, 1975) which is based on the cumulative sum of the recursive residuals. The following chart depicts the cumulative sum together with the 5% critical Lines. The test clearly indicates towards the coefficient stability.



Annexure 1b: Results of MIMIC Model (LISERAL output)

	Estimated Coefficients	t-statistics
Cause Variables		
Tax/GDP (%)	0.16	4.26
M2/GDP (%)	-0.06	-3.55
Durable (regime durability)	-0.019	-1.53
Indicator variables		
Currency to M2 ratio (C/M2) (%)	11.13	6.68
Growth in electricity consumption (%)	1.00	
Diagnostic Tests		
Chi-square (p-value) = 1.40 (0.50) RMSEA= 0 90% confidence interval for RMSEA = (0:0.29) p-value for test of close fit (RMSEA<0.05) = 0.57 N=36 Goodness of Fit Index (GFI) = 0.98 Adjusted Goodness of Fit Index (AGFI) = 0.88		

Annexure 2: Informal Economy (%) Estimated by Different Studies						
	Shabsigh (1995)	Ahmad and Qazi (1995) 1	Aslam (1998)	Ahmed and Haider (2008) [1]	Kemal (2007) [2]	Arby (2009) [3]
1960		52.06	29	60.2		
1961		55.21	29.3	51.3		
1962		54.68	31	45.1		
1963		47.57	29.4	40.4		
1964		46.34	30.5	37.2		
1965		50.97	33	35.2		
1966		42.06	31	33.8		24.4
1967		47.81	37	32.9		29.2
1968		42.09	35	31.7		28.8
1969		48.56	41	30.2		33.1
1970		48.31	40.6	27.9		36.0
1971		39.31	32.4	25.1		32.3
1972		38.48	44.4	22.5		29.8
1973		37.76	42	20.4		29.3
1974		38.59	34.7	19.3	38	27.1
1975	20.74	34.5	30.6	18.8	33.1	25.9
1976	22.92	35.44	27.1	19.3	31.6	28.4
1977	22.06	34.53	27.5	21.1	30.9	27.9
1978	22.01	38.26	46.3	24.6	34.9	29.2
1979	21.98	41.08	46.7	28	39.2	31.1
1980	22.53	49.46	52.6	31	45.6	33.3
1981	24.19	51.46	45.3	32.9	43	33.1
1982	21.91	47.51	43.1	33.1	47.8	31.6
1983	25.64	48.8	46.8	34.2	42	32.8
1984	23.13	49.92	42.5	33.5	49.3	32.1
1985	21.63	45.49	40.2	33.1	39.3	29.6
1986	21.55	41.08	43	33.2	44.7	35.2
1987	21.39	43.39	38.8	32.9	50.5	35.4
1988	24.73	42.89	45	32.3	45.5	32.7
1989	23.31	43.29	46	32	42.7	32.5
1990	23.56	39.27	43.9	32	39.2	30.0
1991	20.46		53	31.3	36.1	26.1
1992			45.3	31.3	44.4	27.7
1993			44.5	31.8	45.5	30.1
1994			42.7	32.3	56.6	33.3
1995			45.7	33.1	60.6	34.8

1996	43.8	32.8	68.7	36.8
1997	38	23.8	74.9	36.4
1998	35.5	23.2	69	36.4
1999		22.5	46.1	35.2
2000		21.9	56.5	26.0
2001		22.2	65.7	26.3
2002		24	64.3	27.0
2003		27.1	68.2	29.0
2004			66.6	24.9
2005			64.8	18.7
2006				18.3
2007				18.9
2008				19.6

1 Based on equation which uses currency and bearer bonds

2 Based on equation 2 of the paper which produces highest estimates of underground economy

3 Results from ARDL model