INFORMAL ECONOMY RESEARCH CENTRE

THE HIDDEN ECONOMY ESTIMATES AND THEIR IMPLICATIONS FOR GOVERNMENT EXPENDITURE

Dilip K Bhattacharyya

Working Paper No: 1

SPONSORED BY: DELOITTE & TOUCHE

THE HIDDEN ECONOMY ESTIMATES AND THEIR IMPLICATIONS FOR GOVERNMENT EXPENDITURE: UK (1960-1990)

D.K.Bhattacharyya, University of Leicester, and IERC England

Abstract

In this paper we examine the problem of updating the 'hidden economy' estimates for the U.K. The general methodology followed in this paper is the same as Bhattacharyya (1990). However, while updating the series we encounter a number of problems. As the generating process is determined by the sample in hand the estimates obtained for the whole period (1960-1990) differ from the earlier estimates. Thus, we can obtain a number of competing estimates for the 'hidden economy' series for the same time period. To discriminate between the two competing series we used a short-run government expenditure function (GEF) as a guideline. GEF has been estimated first by using the published official data. The GEF is then re-estimated incorporating **two** different versions of the estimated 'hidden economy' series with similar behavioural assumptions. This experiment allowed us to examine the validity of **Wagner's law** and at the same time helped us to conduct discriminatory tests to choose among alternative competing series of the 'hidden economy'. As a by-product of GEF estimation we also found an implicit tax rate which can be used to estimate the tax losses to the Government.

JEL Classification: C10, E62

KEY WORDS: Hidden Economy, Wagner's Law, Government Expenditures, DA Criterion, Data Consistency.

Acknowledgement

An earlier version of this paper was presented at the European Economic Association's Congress, the Inland Revenue, UK, CERGE/EI, Prague and to the Money Macro and Finance Conference, London. The author is grateful to the participants of the seminars for their comments. However, the author is responsible for all errors that may remain.

Revised November 1998

Contact address

The Informal Economy Research Centre
Deloitte & Touche
Hill House
1 Little New Street
London EC4A 3TR

Tel: (+44-171) 303 4343
Email: ierc.london.uk@deloitte.co.uk

THE HIDDEN ECONOMY ESTIMATES AND THEIR IMPLICATIONS FOR GOVERNMENT EXPENDITURE: UK (1960-1990)

D.K.Bhattacharyya, University of Leicester

1. Introduction:

All recent available evidences suggest that the unrecorded economy is growing for most countries. This observation raises many questions both in terms of economic theory and economic policy. However, to answer questions of empirical importance it is necessary to have some idea of the size of the 'hidden economy'. Although it is possible to obtain estimates of the 'hidden economy' following the procedure suggested in Bhattacharyya (1990) updating the series creates a number of statistical problems. One of the problems we face in updating the hidden economy series is the changing nature of the data generating process (DGP). The changes in the DGP can be attributed to two components: (1) intrinsic changes in the DGP;[1] and (2) changes in the estimated DGP due to sampling fluctuation.

In this paper we are examining the problem of updating the 'hidden economy' series using the U.K. data and suggest empirical procedures to overcome these difficulties.[2] As a part of the updating exercise we also study the implication of having the 'hidden economy' on the government expenditure function. We obtain dual benefit from this exercise. First, this exercise allows us to discriminate between two competing series of the 'hidden economy' estimates. Secondly, we obtain useful results from the short-run government expenditure function to comment on the existence of 'Wagner's Law' in the presence of the 'hidden economy'. As an extension of this exercise we have also estimated the average tax losses due to 'hidden economy'.

---

1 This may be described as the structural changes in the DGP.

2 The problem we face here has similarities with the issues dealt with in Berndt and Savin(1977).

The paper is planned in the following way. In section 2 we present the model and the general methodology of the paper. The estimated currency demand equations and the details of the 'hidden economy' estimates are presented in section 3. A particular version of 'Government Expenditure Function' was estimated and presented in section 4 where we also studied the interrelationship between the estimated 'hidden economy' and 'Wagner's Law' derived from the published data. The 5th section is devoted to conclusive comments. The two competing series of the estimated 'hidden economy' are presented at the Appendix.

2 Model and Methods

The assumptions used to specify the model and to derive the method of estimating the 'hidden economy' are taken as maintained hypothesis in this study. These assumptions are:

i. the size of the economy measured in terms of national income or GDP has a one to one correspondence with the notes and coin in circulation, ie we assume that the ‘Cambridge Equation’ is valid;

ii. the transactions in the underground economy are primarily conducted in currency;

iii. at any given point of time the total currency in circulation is the sum of (a) demand for currency by the recorded economy and (b) demand for currency by the unrecorded economy.

iv. there is no misspecification in the currency demand equation for the recorded economy.

From the above assumptions we have,

$M_t = M_{Rt} + M_{URt}$ (A)

where M_{Rt} and M_{URt} are the demand for currency by the recorded sector and the unrecorded sector respectively, and M_t is the total demand for currency by the whole economy.

The currency demand equation is specified in a flexible form following Baumol and Tobin (1989). Thus, for the recorded economy the currency demand equation is,

$$M_{Rt} = A\, Y_{Rt}^{\beta_1} \Pi_t^{\beta_2} P_t^{\beta_3} \; e^{F(L)u_t} \qquad (1)$$

where M_{Rt} is as defined before, Y_{Rt} is a recorded income variable, Π_t is a short-term interest rate, P_t is the general price level. A, β_1, β_2 and β_3 are the parameters of the model and F(L) is a polynomial in lag operator L. We also assume that u_t is 'white noise'.

The demand for currency by the unrecorded economy is,

$$M_{URt} = Y_{ht}^{\beta_4} + w_t \qquad (2)$$

where Y_{ht} is a measure of the unrecorded economy and w_t is a 'white noise'. We assume u_t and w_t are independent of each other.

For most economy M_t is observable but not M_{Rt} and M_{URt}. Hence, we write the approximate relation

$$\ln M_t \approx \ln M_{Rt} + (M_{URt}/M_{Rt}) \qquad (3)$$

Substituting for M_{Rt} and M_{URt} from equation (1) and (2) and after some algebraic manipulation (3) can be written as,[3]

$$m_t = \alpha + \beta_1 y_{Rt} + \beta_2 (\pi)_t + \beta_3 p_t + (Y_{ht}\beta_h/H(.)) + \varepsilon_t + \nu_t \qquad (4)$$

where J_{Rt}, π_t and p_t are logarithm of Y_{rt}, Π_τ and P_τ respectively,

$$H(.) = A\, Y_{Rt}^{\beta_1} \Pi_t^{\beta_2} P_t^{\beta_3} \qquad (5)$$

and ε_t is 'white noise' and ν_t is a serially correlated disturbance term.

In this model Y_{ht} is not observable and that we proxy by

$$\sum_{i=2}^{4} \alpha_i Y^i_{Rt} \qquad (6)$$

This proxy will produce asymptotically desirable estimates when the conditional expectation of Y_h given Y_R is a non-linear function of Y_R and which can be approximated by the quartic function of Y_R.[4] Substituting Y_{ht} in (4) by (6) we have

[3] The derivation of the specification in the relation (4) is available in Bhattacharyya (1990 and 1989).

[4] The decision to use a quartic function of Y_R is justified on two grounds. First, this fits the specification of RESET proxy suggested by Thursby and Schmidt(1977). Secondly, in our search process quartic function produced the best results in terms of diagnostic tests. It is also possible to argue that $E(Y_{ht}|Y_{Rt})$ can be represented a polynomial in Y_{Rt}. With this interpretation it is possible to conjecture, following Bhattacharya and Bhattacharyya (1993), that the estimates obtained with the proxy have desirable large sample properties.

a estimable equation where all variables are observable. However, it should be noted the estimating equation presented in (4) is a hybrid model after the substitution of Y_{ht} by (6) and the elasticities are not constant. By estimating this derived model in observed variables we obtain estimates of α_2, α_3 and α_4. The estimates for the 'hidden economy' is then calculated as,

$$Y_{ht} = \sum_{i=2}^{4} d_i \, Y^{i}_{Rt} \qquad (7)$$

and the conditional standard error of Y_{ht} is given by

$$\text{S.E. of } (Y_{ht} \mid Y_{Rt}) = \left[\sum_{i=2}^{4} \sum_{j=2}^{4} \text{cov}(_{1i,1j}) \; Y^{i}_{Rt} \; Y^{i}_{Rt} \right]^{1/2} \qquad (8)$$

The model is estimated by following a two-stage non-linear estimation procedure similar to that suggested by Durbin (1970). However, the error structure of the model allowed us to develop another diagnostic criterion that we found very useful in selecting the final estimates. According to our specification of the model the dynamic structure is related to the currency demand from the recorded sector whereas the demand for currency by the unrecorded sector is static. In our two-stage estimation procedure the dynamic effect is estimated in the second stage by including the lag residuals as an explanatory variable. The estimates for α's should therefore remain the same in two stages of the estimation. We, therefore, choose the estimated series for which

$$DA = (1/n) \; (\Sigma |Z_{1t} - Z_{2t}| / |Z_{1t}|) \times 100.0$$

is closest to zero where Z_{1t} is the estimate of the 'hidden economy' without the correction for dynamic structure and Z_{2t} is the estimates of the 'hidden economy' after the dynamic effect is included in the specification.

As we expect β_4 to be within the range 0 to 1 we obtained estimate for this parameter by grid search.[5] The final estimates of the parameters were chosen where DA has the lowest value and the estimated model passes all other standard diagnostic tests.

3. Empirical Results

In Bhattacharyya(1990) the 'hidden economy' estimates are reported for the period 1960:1 to 1984:4. The choice of the period was dictated by the data availability. It is well known that the official statistics are revised after its first publication and the revisions continue for many years. This prompted us to restrict our estimation to 1984:4.[6] For the same reason we are restricting our updating to 1990:4 in this paper. In fact, we observed that the 1996 Annual Supplement of the Economic Trends produced revised data even for dates prior to 1990. Therefore, we restricted our study for the period 1960 to 1990. The basic data were collected from the Economic Trends Annual Supplement and Financial Statistics.[7]

However, updating of an estimated series of the 'hidden economy' following the procedure suggested here has a special problem. Like all estimates the estimated series of the 'hidden economy' will also change from sample to sample. Thus, the addition of more samples in the updating procedure will produce a different series of the 'hidden economy' estimates for the best fitted model. However, most

---

5 If the hidden economy is created totally by cash transaction then the β_4 should be equal to 1. Hence, the search has been made within the grid 0 and 1. With changing financial markets the parameter β_4 may be changing and the structure of that changes will be of great interest to the policy makers. However, in this paper we are still treating β_4 as a fixed parameter of the model.

6 The revision of published data is much more problematic than what we believed to be the case in 1990. Comparing recent annual supplements of *Economic Trends* we found that the data on GNP current market prices reported in 1996/97 supplement are noticeably different from the figures presented in 1995 supplement from 1984 onwards. This is also true for other series published in the *Economic Trends*. Thus, in our current estimation we have 28 data points that have changed in the most recent publications of *Economic Trends(1996/97 Annual Supplement)*.

7 A full list of the sources of data is presented at the end of the paper.

empirical researchers would prefer to have an updated series that will be consistent with earlier series. Hence, in this paper we compare the best estimated series with a 'near consistent' series of the 'hidden economy' where the estimates for 1960 to 1984 are approximately the same as Bhattacharyya(1990).[8] We believe that this comparison will provide us information that will generally be useful to all empirical researchers. In judging the superiority of a model on the basis of a large number of diagnostic tests may create problems unless all tests for model 1 are numerically lower (higher) than the corresponding test statistics for model 2. In this paper, we decided to rank the two models on the basis of the mean values of DA criterion.[9] While updating the series we observe that there is structural shift in the DGP of the 'hidden economy' after 1987 crash. This, we think, is an important finding, because this may suggest that a great proportion of adjustments due to the 'shock' of 1987 crash are absorbed in the unrecorded economy. This implies that the recorded economy will move relatively smoothly after the 'shock' and may not exhibit any substantial changes in the movement. The estimated parameters of the 'best' model in terms of DA criterion and the 'series consistent' estimates are presented in tables 1 and 2 respectively. To incorporate the structural changes in 1987 we introduce a dummy variable D_t that takes the values 1 for the years 1988 to 1990, but zero otherwise. The parameter α_2 is replaced by $\alpha_2 + \delta_2 D_t$, α_3 by $\alpha_3 + \delta_3 D_t$ and α_4 by $\alpha_4 + \delta_4 D_t$ at the time of estimating the model. Similarly, in the calculation of the 'hidden economy' estimates and their standard errors, of the estimates 2_is are replaced by $(L_i + \delta_i D_t)$'s and cov $(LI, \delta i D_t)$ by $(L_i + \delta_i D_t)$

Table 1

Estimated Parameters of the Model for the Period 1960-1990

(Mean DA minimum)

---

8 The 'near consistent' series presented here is slightly different from Bhattacharyya(1990) as some corrections had to be done due to certain mistakes in the data. The new estimates are available in Bhattacharyya(1996) or by request from the author.

9 The initial ranking of the models can be achieved by choosing any diagnostic tests. We have chosen DA criterion as we use that for final selection of the estimates.

Parameters	Estimates	Standard Errors
lnA	0.710585	0.204719
β_1	0.622130	0.066077
β_2	0.052482	0.015344
β_3	-0.444357	0.108164
ρ	0.701508	0.067222
α_2	0.835934-02	0.206719-03
α_3	-0.140243-03	0.651042-05
α_4	0.705253-06	0.454546-07
δ_2	-0.574915-02	0.104107-02
δ_3	0.125832-03	0.195461-04
δ_4	-0.693263-06	0.949071-07
β_4	0.634	

log likelihood = 350.715; R^2 = 0.9995; = 0.01465;
DW = 2.2061;
Ljung-Box(χ^2(10))=12.9; LM(χ^2(10))=15.00;
ARCH-F(5,112)=0.790;
Innovation Error Test-F(15,102)= 1.523; Mean(DA)=0.1300;
Var(DA)=0.02903;

Table 2

Estimated Parameters of the Model for the Period 1960-1990
(Estimates approximately consistent with earlier series)

Parameters	Estimates	Standard Errors
lnA	0.639027	0.206021
β_1	0.621736	0.062156
β_2	0.042349	0.013757
β_3	-0.395874	0.106624
ρ	0.705071	0.067130
α_2	0.696060-02	0.232298-03
α_3	-0.116611-03	0.631447-05
α_4	0.582846-06	0.419720-07
δ_2	-0.455790-02	0.725658-03
δ_3	0.100473-03	0.137121-04
δ_4	-0.556406-06	0.681135-07
β_4	0.696	

log likelihood = 350.009; R^2=0.9995; =0.01473; Dw = 2.2019;
Ljung-Box(χ^2)=14.3; LM(χ^2)=16.03; ARCH-F(5,112)=0.8265;
Innovation Error Test-F(15,102)= 1.378; Mean(DA)=0.5125;
Var(DA)=0.07416;

It is clear from the diagnostic tests and from the estimates of the parameters and their standard errors that both estimates fit the data well. In fact, purely on the basis of the statistical tests presented in the tables the estimates cannot be differentiated.[10] However, in numerical terms the estimated parameters are different and will

[10] To determine whether the estimates presented in Tables 1 and 2 are the same in statistical sense or not is a very complex issue and therefore we avoided that line of analysis in this study.

produce different 'hidden economy' estimates although both estimates suggest a significant behavioural shifts on the demand for currency by the unrecorded sector after the 1987 crash. We consider this finding as an interesting by-product of the study as this may imply that following the crash most of the adjustments in the economy are done in the unrecorded economy, thus showing near stable structures of the recorded economy. Given that we cannot effectively discriminate the two competing results presented in tables 1 and 2, it is necessary to examine them in different ways. This examination of the estimated models are conducted in three different ways namely:

i. by comparing the forecasting ability of the two fitted equations;

ii. by considering the plausibility of the two series of the 'hidden economy' estimates;

iii. by assessing the relationships between the government expenditure function and the 'hidden economy' series.

A comparison of the forecasting ability of the two competing estimated equations is presented at the Table 3 below. The forecasts are obtained for the period 1991:1 to 1994:2, and we used the dynamic one period ahead forecasting procedure. The results suggest that the ability to forecast is not very different for the two fitted equations. The maximum difference between a forecast and the observed value of the log(M_0) is only 3.59% of the observed value and the minimum difference is only 0.07% of the observed value using the estimates of Table 1. The corresponding differences using the results of Table 2 are 3.60% and 0.04% respectively. Thus in terms forecasting ability the two fitted equations cannot be distinguished.

Table 3

Statistics Related to the Actual and Forecast Values of LnM_0

	LM_0	LM_0F1	LM_0F2	DIF1	DIF2
Mean	2.81166	2.79960	2.79959	0.01206	0.01207
Variance	0.00297	0.00136	0.00136	0.00289	0.00284

LM_0 = Logarithm of currency in circulation;

LM_0F1 = Forecasts of LM_0 using estimates of Table 1;

LM_0F2 = Forecasts of LM_0 using estimates of Table 2;

DIF1 = LM_0 - LM_0F1; and DIF2 = LM_0 - LM_0F2.

We also have conducted the forecasting exercise for the first eight quarters separately and observed no qualitative differences in the forecasts when compared with the forecasts for 14 quarters presented in Table 3.

Failing to discriminate between the two fitted equations on the basis of diagnostic tests and forecasting abilities we examined the estimated 'hidden economy' series in terms of actual magnitudes and the movements over time using the results of Table 1 and Table 2. The estimated 'hidden economy' series are presented in the Appendix Tables A and B. The estimates presented in Table A are based on the estimated parameters in Table 1. Here we observed that the mean and variance of DA is minimum. In absolute terms the sizes of the 'hidden economy' varied from 0.21726 billion pounds sterling in 1960:2 to 15.58833 billion pounds in 1990:4; which as a percentage of GNP is 3.36% for 1960:2 whereas for 1990:4 it is 11.16%. The estimated 'hidden economy' series is increasing, almost, monotonically during this period. In 1978:4 the 'hidden economy' became 13.01% of GNP and attained the maximum value of 13.57% of GNP in 1980:3. The size of the 'hidden economy' fluctuated but stayed above 13.0% of the GNP until 1982:3. Since then the 'hidden economy' is continuously decreasing as a percentage of GNP until 1989, but some increases are observed during 1990. The estimated 'hidden economy' presented in the Appendix Table B are consistent with the starting estimates obtained for the period 1960:2 to 1984:4 in Bhattacharyya(1996)[11]. Thus the estimates from Appendix Table B can be taken as the updated series of the 'hidden economy' published earlier. In fact, according to these estimates the absolute value of the 'hidden economy' in 1960:2 was 0.18093 billion pounds and risen to 8.6918 in 1990:4. This series also attained the maximum (10.42%) as a percentage of GNP in 1979:4, and the figures were above 10% for the period 1975:3 to 1980:3. Thus the general pattern of movement for the two series are broadly similar, although the actual levels and the timings of peaks and troughs are different. We have no obvious reasons to believe that the actual sizes of the 'hidden economy' are closer to the estimates presented in Appendix Table A. It is possible to argue that the estimates consistent with the earlier estimates are more likely to be useful updated estimates of the 'hidden economy'. Hence, the estimates presented in Appendix Table B may be considered a more useful than the estimates in Appendix Table A as that implies no revision is required for the data produced in Bhattacharyya (1996).

4. Government Expenditure Function

In the previous section we failed to decide which of the two possible series of the estimated 'hidden economy' will be more acceptable on statistical ground. In this section, we compared the two series by judging the plausibility of the effects they may have on the short run government expenditure function (GEF). The literature on the study of GEF is fairly large and exists in many different names. However, the existing literature has one common feature, i.e. most of the studies estimated long run relation between the national income and the government expenditure (see, Chrystal and Alt (1979), Ram(1986,1987)). However, Gemmell (1990) extended the earlier approach by including the price effect in the GEF (this is often known as a study of Wagner's Law). We are following Gemmell's (1990) basic specification but we assume that three groups of people are directly or indirectly involved in the

[11] These estimates are slightly different from Bhattacharyya (1990). For further details see footnote 9.

government expenditure:

(a) elected government;

(b) civil servants; and

(c) the consuming public.

The elected government has a planned expenditure to satisfy the consuming public. We assume that the desired real government expenditure (G^*) is determined by the real per capita GDP, population, prices of government output and prices of non-government output. This specification can be written as:

$$G^* = A\, y^{\beta_1} \Pi^{\beta_2} \; P_g^{\beta_3} \; P_c^{\beta_4} e^{u} \qquad (9)$$

where G^* is the desired real government expenditures

y is the real per capita GDP,

Π is the population

P_g is the prices of the government output,

P_c is the prices of non-government output,

u is the disturbance term with mean 0 and variance σ^2.

However, the government executes their planned expenditure through civil servants. We assume that civil servants have the primary objective to increase the real government expenditures. However, unbounded increases are not possible either due to general economic condition of the country or due to political commitment of the elected government. When the increase in the real government expenditures is not possible then the civil servants will adopt an adjustment procedure for the government expenditure plan. Thus, the desired real government expenditure G^*_t is adjusted to a time path through the intervention of the civil servants as well as due to short run fluctuation of the economy. We assume that the adjustment to the desired level is achieved through the relation $g_{t^*} = D(L)g_t$, where D(L) is a polynomial in lag operator L and g_t is the logarithm of G_t.

By taking logarithm the equation (9) can be written as:

$$g_{t*} = \ln A + \beta_1 \ln y_t + \beta_2 \ln \Pi_t + \beta_3 \ln P_{tg} + \beta_4 \ln P_{tc} + u_t \quad (10)$$

Substituting for g_{t*} the equation (10) can be written as:

$$g_t = \ln A+\beta_1 \ln y_t+\beta_2 \ln \Pi_t+\beta_3 \ln P_{tg}+\beta_4 \ln P_{tc}+(1-D(L))g_t+u_t \quad (11)$$

The interpretation of the parameters of this model plays an important role in the testing of the Wagner's Law. It is believed that if the prices are affecting the government expenditure decisions then $\beta_3<0$ and $\beta_4>0$. If the Wagner's Law is valid then $\beta_1>1$ and the parameter β_2 signals whether the government outputs are 'pure' public goods or not. If $\beta_2= 0$, then the government expenditure is on pure public goods; if $\beta_2= 1$, then the government spending is on pure private goods; and if $0<\beta_2<1$ then the government spending is on a mixture of public and private goods. In our empirical work we use the flexible version of equation (11), where y_t is taken as the real GDP and the per capita effect is included in Π. We assume that (1-D(L)) can be approximated by finite lag structure and with normalisation (1-D(L))g_t can be written as:

$$\sum_{i=1}^{k} \alpha_i L^i g_{t-i}$$

where k is any arbitrary finite number.

In choosing the lag structure we followed 'general' to 'specific' approach in the results presented below.

The estimated relation with the published data is:

$$\begin{aligned} g_t = &-3.5647 + 0.1249\ln y_t + 0.3466\ln\Pi_t - 0.1229\ln P_{gt} \\ &(12.565) \quad (0.093) \qquad (1.294) \qquad (0.255) \\ &+ 0.1192\ln P_{ct} + 0.1651 g_{t-1} + 0.6805 g_{t-4} \qquad (12) \\ &\quad (0.292) \qquad (0.079) \qquad (0.077) \end{aligned}$$

Adj.R^2 = 0.865; log-Likelihood = 192.5; = 0.0475; DW-h =1.7178;

Ljung-Box-Q(4) = 8.5518; ARCH test = 0.4371; ADF(4) = -9.558;

Breusch-Pagan het. test = 3.876.

It is easy to see from (12) that the equation fitted well in terms of overall diagnostic tests. The signs of the estimated parameters correspond with prior expectations, although three of the parameters are not significantly different from zero. The estimate of β_1 is much lower than that was reported in Gemmell(1990) for the developed countries. However, the Wagner's Law is a long run relation, hence we calculated the long run solution for $_1$ from our estimated equation which we found is equal to 0.8089 with asymptotic standard error 0.9414. Thus we do not find any support for Wagner's Law from the published statistics.

A natural question arises from these findings. Will the government expenditure function (GEF) be any different when the estimates of the 'hidden economy' included in the study? The issue is complex. The expenditure plan of the government will depend on the type of knowledge they have about the size and working of the 'hidden economy'. There are two obvious ways the 'hidden economy' can affect the government's expenditure decision. It is possible that the government has a very clear idea about the size of the 'hidden economy'. The government may think that when the income is generated within the economy the indirect tax revenue will increase although the income is hidden from the national statistics. In this situation we consider that the 'hidden economy' is fully anticipated by the government and they include this in measuring the real GDP while deciding their expenditure plan.

There is an alternative way one could treat the fully anticipated 'hidden economy' in the GEF. It is possible that the observed government expenditure is based on the recorded real GDP and the 'hidden economy' not only changes the size of the real GDP this also changes the real government expenditure by the amount of real tax revenue that can be collected from the 'hidden economy'. Assuming a balanced budget situation this interpretation allows us to calculate the average tax losses that is due to the presence of the 'hidden economy'.

To examine the effect of the 'hidden economy' we start with the specification of the fitted GEF presented in (12). We then include the 'hidden economy' in that specification in the ways suggested above.

5a. 'GEF' With The 'Hidden Economy': Part I

We first re-estimate the GEF (as in (12)) with the assumption that the government fully anticipate the size of the 'hidden economy' and includes that in the calculation of the real GDP but we assume that the published government expenditure figures incorporate the expenditure that can be attributed to the 'hidden economy'. This experiment is conducted separately with the two plausible 'hidden economy' series presented at the Appendix. The estimated equations based on data for the period (1962 - 1990) are:

$$g_t = \underset{(4.414)}{-1.0415} + \underset{(0.078)}{0.1730} \ln y_{h1t} + \underset{(0.439)}{0.0701} \ln \Pi_t - \underset{(0.180)}{0.1208} \ln P_{gt}$$

$$+ \underset{(0.199)}{0.1064} \ln P_{ct} + \underset{(0.067)}{0.1683} g_{t-1} + \underset{(0.068)}{0.6807} g_{t-4} \qquad (13)$$

Adj.R^2 = 0.867; log-Likelihood = 193.19; = 0.0472; DW-h = 1.374;

Ljung-Box-Q(4) = 8.5276; ARCH test = 0.3876; ADF(4) = -9.651;

Breusch-Pagan het. Test = 4.423.

$$g_t = \underset{(4.377)}{-1.1408} + \underset{(0.080)}{0.1966} \ln y_{h2t} + \underset{(0.434)}{0.0589} \ln \Pi_t - \underset{(0.180)}{0.1364} \ln P_{gt}$$

$$+ \underset{(0.198)}{0.1156} \ln P_{ct} + \underset{(0.067)}{0.1675} g_{t-1} + \underset{(0.067)}{0.6799} g_{t-4} \qquad (14)$$

Adj.R^2 = 0.868; log-Likelihood =193.75; = 0.0470; DW-h =1 .295;

Ljung-Box-Q(4) = 8.3692; ARCH test = 0.3736; ADF(4) = -9.697;

Breusch-Pagan het. test = 4.576.

The estimated equation (13) is obtained by adding the real 'hidden economy' estimates derived from Appendix Table B to the real GDP where we assume that the government has perfect knowledge of the size of the 'hidden economy'. It is clear that from the results presented here that the equation (13) fits the data very well and the diagnostic statistics are slightly improved when compared with the equation (12). The major difference between the equation (13) and (12) are the point estimates associated with the variable income and population. In addition from equation (13) we find that the long run β_1 is 1.1457 with asymptotic standard error 1.3288. This is substantially higher than 0.8089 derived from equation (12). Thus the results in (13) supports Wagner's Law whereas there was no such support from the results obtained in equation (12). These findings have serious implications for many macroeconomic studies as the estimated macro relations will be considerably different when the estimates for the 'hidden economy' are explicitly included in the specification of the model.

Similarly, we estimated the GEF by adding the 'hidden economy' estimates from Table A to the GDP and the estimated equation is presented in (14). The diagnostic statistics are improved in (14) compared to both (12) and (13) although the improvement from (13) is very small. The estimated equation (14) is different from equation (13). The differences are even larger when compared with estimates presented in (12). The implied long run estimate for β_1 is 1.288 with the asymptotic standard error 1.3916. This long run estimate is slightly higher than 1.1457 obtained from the estimates in (13). Thus, once again the inclusion of the 'hidden economy' estimates provided support for the existence of Wagner's Law. There is no clear indication in the literature about the empirical validity of Wagner's Law. However, if we believe that the 'Wagner's Law' is valid for the U.K. then the 'hidden economy' estimates are very useful in predicting the government expenditure. It is interesting to note that the long run estimate of β_1 as obtained from (14) is numerically close to the value presented in Gemmell (1990) for the developed countries. Hence, the weight of the evidence suggests that the 'hidden economy' estimates obtained in Table A are more reliable. This implies that re-estimation of the data generating process using the complete sample is better than updating the previously obtained estimates.

5b. 'GEF' With The 'Hidden Economy'

Part II

Our assumption of perfect knowledge of the 'hidden economy' may not be very realistic, although the results in (13) and (14) are plausible. Hence, we have conducted the second experiment described earlier. In this experiment we acknowledge that if the 'hidden economy' was not hidden then a percentage of the 'hidden economy' would have been added to the government expenditure and while the total 'hidden economy' would have been a part of the GDP. Therefore, we conducted this experiment with the two estimated 'hidden economy' series in the following way:

i. We assumed that a mean tax rate can be defined within the period that can be used to calculate the tax revenue from the 'hidden economy'. We denote that by λ.

ii. Then we assume balanced budget condition and added λY_{ht} to the government expenditure data and Y_{ht} to the GDP figures and estimated relations similar to (13) and (14).

iii. We searched for λ within the range 0 to .5 and accepted the results that appeared most plausible. Thus, in this exercise we indirectly find the amount of tax losses due to the presence of the 'hidden economy'.

While searching over a range of λ we observed that the main diagnostic statistics are not very helpful in choosing the best fitted value of λ. Therefore, we had to look for an alternative criterion for choosing the best λ. In equation (13) and (14) we observed that the long run solution of β_1 are plausible and generally matches the findings of other studies. Therefore, we decided to choose the values of λ which produced the fitted equations with long run solutions of β_1 close to that we obtained from fitted equations (13) and (14).[12]

The estimated relations are,

$\lambda = 0.28$.

$$g_{1t} = -\underset{(4.281)}{0.3312} + \underset{(0.077)}{0.1729}\ln y_{h1t} - \underset{(0.419)}{0.0172}\ln\Pi_t - \underset{(0.176)}{0.1168}\ln P_{gt}$$

$$+ \underset{(0.194)}{0.1017}\ln P_{ct} + \underset{(0.065)}{0.1779}g_{t-1} + \underset{(0.066)}{0.6941}g_{t-4} \qquad (15)$$

Adj.R^2 = 0.888; log-Likelihood = 197.41; = 0.0455; DW-h = 1.836;

Ljung-Box-Q(4) = 7.937; ARCH test = 0.4969; ADF(4) = -9.755;

Breusch-Pagan het. test = 2.84.

---

12 The search for λ is made with a step length of 0.02. Hence, our preferred value of λ can be 0.01 different from the best fitted value. In reality λ will not be fixed over the whole period, hence the loss precision due to step length may not very significant.

$\lambda = 0.38.$

$$g_{2t} = \underset{(4.153)}{0.8607} + \underset{(0.101)}{0.2004} \ln y_{h2t} - \underset{(0.403)}{0.1587} \ln \Pi_t - \underset{(0.172)}{0.1131} \ln P_{gt}$$

$$+ \underset{(0.188)}{0.0933} \ln P_{ct} + \underset{(0.064)}{0.1847} g_{t-1} + \underset{(0.064)}{0.6913} g_{t-4} \qquad (16)$$

$R^2 = 0.912$; log-Likelihood = 201.09; = 0.0441; DW-h = 1.853;

Ljung-Box-Q(4) = 7.9704; ARCH test = 0.3225; ADF(4) = -9.856;

Breusch-Pagan het. test = 2.833.

Where

g_{1t} is the log of published government expenditure plus λ times the 'hidden economy' estimates in Appendix Table B.

y_{h1t} is the log of real GDP plus the 'hidden economy' estimates in Appendix Table B.

g_{2t} is the log of published government expenditure plus λ times the 'hidden economy' estimates in Appendix Table A.

y_{h2t} is the log of real GDP plus the 'hidden economy' estimates in Appendix Table A.

Once again we find that the two estimated equations (15) and (16) fit well to the data and the long run solutions from the equations support Wagner's Law. The long run solutions for β_1 are 1.3508 and 1.6154 derived from the estimated equations (15) and (16) respectively. However, the main difference between the estimated equation (15) and (16) are the implied tax rates (the estimates of λ) which can be used to calculate the tax losses due to 'hidden economy'. The implied tax rates figures are interesting on two counts: (1) the implied tax rates 28% and 38% are much higher than those predicted in the literature before and definitely above the average direct tax rate calculated from the published data; (2) the lower rate is related to the lower estimates of the 'hidden economy' derived to make it consistent with the estimates published in Bhattacharyya(1990,1996). In all the tests conducted so far the new estimates obtained for (1960 - 1990) and presented in Table B have stronger supports. The 38% tax loss rate implies the total tax losses to the government would be even higher than £30 billions. The 38% figure is much higher than the marginal tax rate faced by most individuals and substantially higher than the standard VAT rate. However, as discussed earlier that there could be unrecorded economy which would be 100% taxable. Hence, the figure 38% is quite plausible and this may suggest that the amount of evasion through VAT system is very high. This interpretation also fits with many anecdotal evidence suggested in the media regarding VAT evasion.

6. Conclusion

The main aim of this paper was to examine the problems of updating the 'hidden economy' estimates. However, in the process of this investigation we obtained a few other results that would interest many empirical researchers.

(a) We concluded that the total revision of the 'hidden economy' estimates are better than updating the series while keeping the earlier estimates fixed. This finding provides a support for the complete revision of the published data even after few years. In fact, we found that if the earlier series was extended only for two or three years that was not much different from the new estimates for the whole period. We conjecture that the revised data published in the 1996/97 *Economic Trends*: Annual Supplement, and presented here as Appendix C, can also be justified by reasons similar to ours.

(b) We observed a distinct shift in the parameters of the data generating process for the 'hidden economy' after the stock market crash of 1987. This may suggest that the shock of the crash had been mainly absorbed in the unrecorded economy, therefore to study the effect of the crash on the economy will require a model, which incorporate the behaviour of the unrecorded economy.

(c) The estimated short-run government expenditure functions were improved when the 'hidden economy' estimates are included in the specification. This suggests that the government made their expenditure plan by implicit consideration of the size of the 'hidden economy'. The inclusion of the 'hidden economy' estimates also brought support to justify the presence of 'Wagner's Law' in the UK economy. This finding suggests that similar experiments should be undertaken for other industrialised countries.

(d) We suggested a new method of calculating the tax losses due to the 'hidden economy'. Traditionally, using the average tax rate implied by the published data does this calculation. However, as the relevant average tax rate for the 'hidden economy' could be anywhere between 0% to 100%, the procedure suggested in this paper is likely to be more accurate.

Data Sources

1) Currency in circulation (M) - *Financial Statistics*, Bank of England (issue department);

2) Personal income in current prices (Y_R) - *Economic Trends*, 1992 Annual Supplement;

3) Retail Price index (P) - *Economic Trends*, 1992 Annual Supplement;

4) London clearing banks' base rate (?) - *Financial Statistics*;

5) Gross national product in current prices (GNP) - *Economic Trends*, 1996 Annual Supplement;

6) Government expenditure (G) - *Economic Trends*, 1995 Annual Supplement;

7) Population - *Annual Abstract of Statistics*, several issues;

8) Price indices for government output and other output - *Annual Abstract of Statistics*, several issues.

References

Berndt, E., and Savin, E, (1977). 'Conflict among criteria for testing hypothesis in the multivariate linear regression model,' *Econometrica*, vol.45, pp. 1263-1277.

Bhattacharya, Rabi and Bhattacharyya, Dilip (1993). 'Proxy and instrumental variable methods in a regression model with one of the regressors missing,' *Journal of*

Multivariate Analysis, vol. 47, pp. 123-138.

Bhattacharyya, D.K.(1989). 'An econometric method of estimating the "hidden economy", United Kingdom (1960-1984): Estimates and Tests,' Wayne State University Working Paper, No. 89-11.

Bhattacharyya, D.K.(1990). 'An econometric method of estimating the "hidden economy", United Kingdom (1960-1984): Estimates and Tests,' *The Economic Journal*, vol.100, pp. 703-717.

Bhattacharyya, D.K.(1996). 'The hidden economy of the UK: revised and updated estimates and tests', paper presented at the Royal Economic Society's annual conference at Swansea, Wales.

Chrystal, K.A. and Alt, J.E.(1979). 'Endogenous government behaviour: Wagner's Law or Gotterdammerung,' S.T.Cook and P.M.Jackson (eds.), *Current issues in fiscal policy*, Martin Robertson, Oxford.

Durbin, J. (1970). 'Testing for serial correlation in least squares regression when some of the regressors are lagged dependent variables,' *Econometrica*, vol. 39, pp. 410-421.

Gemmell, N. (1990). 'Wagner's law, relative prices and the size of the public sector,' *The Manchester School*, vol.58, pp. 361-377.

Ram, R. (1986). 'Comparing evidence on Wagner's hypothesis from conventional and "real" data,' *Economics Letters*, vol.20, pp. 259-262.

Ram, R. (1987). 'Wagner's hypothesis in time-series and cross-section perspectives: evidence from "real" data for 115 countries,' *Review of Economics and Statistics*, vol.69, pp.194-204.

INFORMAL ECONOMY RESEARCH CENTRE

CORRUPTION IN INDIA AND THE HIDDEN ECONOMY

Dilip K Bhattacharyya

Working Paper No: 2

SPONSORED BY: DELOITTE & TOUCHE

CORRUPTION IN INDIA AND THE HIDDEN ECONOMY

By Dilip K. Bhattacharyya, University of Leicester and IERC,and Susmita Ghose, CASAD, Pune

Abstract

In this paper we argue that firms due to regulatory conditions primarily generate corruption, in terms of bribes and kickbacks. To recover the cost of bribery, firms hide their production output that then remains unrecorded in the official statistics. Hence, by estimating the unrecorded income of the industrial sector it is possible to examine the growth of corruption. A method for estimating sectoral unrecorded income is suggested in this paper, and from the estimated unrecorded income of the industrial sector we demonstrate that the large increases in corruption signalled by recent reported cases are justifiable. Our empirical results also suggest the disaggregated 'hidden economy' estimates are more informative than aggregated estimates.

Acknowledgement

This paper is based on a larger study supported by the Overseas Development Administration of the British Government under the research contract R5905. The main data collection and analysis was done in Pune, India with the support from CASAD, Pune. We are grateful to the ODA and CASAD for their support. We are grateful to Sonali Bhattacharyya for editing the preliminary draft. An earlier version of the paper was presented in the ESRC/Development Economics Study Group Conference at Reading. However, the authors are responsible for any errors or omissions.

Revised, March 1998

(Published in Economic and Political Weekly of India)

CORRUPTION IN INDIA AND THE HIDDEN ECONOMY

By Dilip K Bhattacharyya and Susmita Ghose

Introduction

In recent years, much corruption involving large sums of money has been uncovered in India. To the people in India, corruption is not a new phenomenon, but the size and scale of corruption that has been uncovered now is beyond the imagination of many people. One school of thought believes that the scale of corruption has always been very high, but has remained undetected. However, the sum of money reported by the CBI as 'corrupt money' cannot be justified for the size of the Indian economy that was recorded twenty years ago. Hence, it is reasonable to assume that the large levels of corruption are only recent phenomena. The CBI reports suggest that the current revelation of 'corrupt money', to a large extent, is associated with foreign currency . This is not surprising as the regulations (official and unofficial) to control the economic activities in India are extensive. For example, to invest in India, a foreign company needs clearance from the state government, central government, and the Reserve Bank of India. This procedure of clearance for investment is meaningful given that once a firm enters India, in general, they do not abide by the laws and regulations using bribes and kickbacks. This is possible partly because the laws are enforced by low paid officials who can be bought at a very low foreign price by a rich foreign company. In this context, the unofficial regulations imposed by civil servants are also important. For example, a clerk can hold up a clearance simply by pretending that the file is lost.[1] The foreign price for the levels of bribery that is necessary to move the file generally small and, therefore, attractive to the firms.

The corruption involving foreign firms is sometimes even termed as 'welfare enhancing' activities. Suppose the return from investment in India is 20%, whereas in the Europe or in the United States the returns are only 10%. In such cases, the firm can share one half of

[1] It is well known in India that after retirement one may have to wait many years to receive the pension. If the retired person is not ready to pay kickbacks to the dealing clerk the pension file will not move to the next level. This type of activities is identified as unofficial regulations.

the extra return with Indian officials and politicians and still make a substantial extra return on the investment. Thus, an investment of $100 million in India will create an additional return of $5 million to the firm. If India's regulations allow a firm to repatriate the full return on investment, the firm will repatriate $20 million. Through bribes and kickbacks India's earnings gone up by $5 million, hence it is 'welfare enhancing' for India. However, in theory such circumstances could lead India to a gain of $5 million earnings by taxing the profit repatriated from India at a rate of 25%. In practice, this may lead to tax evasion and thus reducing the total revenue to the country. Suppose, for simplicity we assume that the tax evasion is stoppable with zero cost. In that situation, it is possible to argue that the extra tax revenue will improve the welfare of the society if the money is spent on non-excludable public goods. It is always difficult to identify completely non-excludable public goods. A practical approximation could be spending on defence. However, in a poor country like India the immediate tangible consumption and growth of the economy may be more desirable than the increase in defence expenditures. Therefore, on pure economic grounds, it is extremely difficult to evaluate which method of collecting $5 million would be more beneficial to the country. For our present discussion we assume that the firms bribe the officials to obtain the licence to operate in the country and therefore, look for an opportunity to recover that amount through under reporting of output.

Suppose some tax concessions or subsidies are available to the foreign firms during the initial stages of their operations in India, when the production level is lower than a certain pre-assigned level. This will encourage the foreign firms to under-report their activities. The consequence of this act is that the firm will produce a 'hidden economy' or will have under-recorded production.

Similarly, Indian firms enjoy subsidies (e.g. export subsidies) and tax relief based on the export as a percentage of the total production of the firm. Thus, hiding a part of the production allows the firm to benefit from the export subsidies. However, to an Indian firm, a more pressing desire to under-report the productions will be custom and excise tax evasion as well as evasion of profit tax. The evasion of taxes is only possible with collusive agreement between the firms and the people representing the tax authorities. The clear outcome of this collusion would be the bribing of people who enforce the laws and regulations. The Indian firms are also bound by the licensing laws of the country. Thus, to obtain licenses, firms offer kickbacks or bribes to the employees of the licensing authorities. To compensate for these losses, firms will not declare total productions, and thus will produce a hidden economy.

The discussion in this section establishes relationships between corruption and the hidden economy that exist mainly through laws and regulations. However, it is possible to argue that a profit maximising firm will include bribes and kickbacks in the cost structure and will produce output following standard marginal condition. In this situation there is no reason for hiding the total production. We argue below that for certain market conditions hiding a part of the production is compatible with profit maximising behaviour of the firm.

Firm's Profit Maximising Behaviour and the 'Hidden Economy'

In the previous section we argued that the regulatory and legal conditions would encourage firms to operate simultaneously in the formal and informal markets. We have argued that a profit maximising firm can maximise profit if the costs of bribes and kickbacks are included in the cost. This implies that given the constraints the firm can maximise profit by deciding the level of output. However, this does not prevent firms to find ways to reduce the number of constraints, which will increase, profit and market share. In the present context the firm can reduce the number of constraints by operating in the informal sector as we assumed that bribes and kickbacks are directly related to the recorded output. We substantiate these assertions by using standard textbook analysis.

First, consider that a firm is operating in a perfectly competitive industry and is in a long run equilibrium after incurring costs for bribes and kickbacks. If the firm can avoid bribes and kickbacks the average cost curve will move downward and as a consequence the firm will produce more output. The increased output is likely to be produced at a lower unit cost. Hence, in this situation the firm can release part of the output to the recorded market and consequently paid the bribes and kickbacks. However, the firm will sell the additional output through informal markets. It is possible to show that by selling simultaneously in 'formal' and 'informal' markets the firm will have increased profit even when the output price in the informal market is lower than the recorded market. It is possible to show that this argument will be also valid when the firm is facing a downward sloping demand curve. More interestingly it can be shown that when the output market is in a depressed condition, for example the demand curve is below the average cost curve, the firm will operate in the informal market as price discrimination will be easier in the informal market. Thus a profit maximising domestic firm have reasons for under-reporting the output due to bribes and kickbacks.

The primary market for the industrial products is very limited in India. Effectively, most industrial products face more than one market. The richer people of the society mainly subscribe the standard recorded markets, whereas the demand from the poorer people is far below the demand schedule of the richer people. In this situation, the poorer people can be served through the informal market where bribing of officials is not necessary. Thus, profit maximising behaviour is quite consistent with the firms' participation in the 'hidden economy'.

India pursued a slow liberalisation process from 1984 when Rajiv Gandhi was the Prime Minister. The liberalisation was only partial and slow, when older laws and regulations were still operative. To take initial advantage of this liberalisation process, both foreign and Indian firms became active participants in corrupt activities by offering bribes and kickbacks. The recent revelations of a large number of corruption cases involving even senior ministers of the government justify this conjecture. This suggests that both foreign firms and Indian firms have an incentive to use informal markets to sell at least some of their products. Hence, by estimating the 'hidden' production of the industrial sector, one could get an idea of the size of corruption, or more precisely the rate of change of corruption.

The estimation of the unreported income in India goes back to 1956 when Lord Kaldor, as a part of the Wanchoo Committee, tried to estimate it from the national income estimates, by subtracting the income assessed for income tax. However, major interest in this area of research came only during the period 1981-83 when a number of people published their findings in the *Economic and Political Weekly*. The majority of authors followed the approaches suggested by Gutmann (1977) and Feige (1979). However, there are exceptions, Chopra (1982) used Kaldor's method with certain extensions and Gupta and Mehta (1981) used a physical input approach. Finally, Acharya (1985) modified Tanzi's method to estimate the unrecorded economy in India. The methods used in these articles are not suitable to estimate the sectoral estimates of the unrecorded economy. Hence, the rest of the paper is planned in the following manner.

In the next section we give a brief description of the modified Tanzi's method of estimation used by Acharya (1985). Then, we examine the estimates of the 'black economy' and the growth of corruption implied by the Acharya's (1985) estimates, which are available only up to 1981. As we conjectured that a major increase in corruption occurred after 1984, we updated the estimates of the 'black economy' following the method used by Acharya (1985) and studied its movement. We then suggest a method of estimating the sectoral 'hidden economy' by extending the method suggested in Bhattacharyya (1990). Our estimates suggest that the 'hidden economy' associated with the industrial sector is a better indicator to measure the rate of change in corruption in India.

The Modified Tanzi Method

Acharya (1985) reported that the direct application of Tanzi's method to estimate the unaccounted income of India is not suitable on the grounds that the assumptions made when studying the US economy are not valid when studying the Indian economy. This argument leads Acharya (1985) to consider the real cash balance as the relevant dependent variable, instead of the cash balance as a proportion of M_2 used in Tanzi's study of the United States. Acharya also observed that the log-linear equation used by Tanzi does not produce plausible estimates of the unrecorded income for India and, therefore, he uses a linear model . Our independent experiments with the updated data support Acharya's findings. The sample period in Acharya's study was 1950-51 to 1980-81. The currency demand equation was specified as,

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + U_t \quad (1)$$

Where

$Y_t =$ (C_t/P_t), where C_t is the annual currency holdings with public and P_t is the index number of wholesale prices at time t;

$X_{1t} =$ Real GNP in market prices at time t;

$X_{2t} =$ Number of commercial bank branches at time t;
$X_{3t} =$ Ratio of total tax revenue to GNP market prices at time t;

$X_{4t} =$ Real rate of interest (i.e. Nominal rate - expected inflation) at time t;

$X_{5t} =$ Annual expected rate of inflation at time t;

$U_t =$ The disturbance term at time t, which is assumed to be serially correlated to encompass the dynamic structure of the currency demand equation.

The expected inflations were calculated by using the adaptive expectation rule, and weights were searched within the interval 0 and 1. The disturbance term was assumed to follow an AR (1) process and consistent estimates were obtained by using Cochrane-Orcutt procedure. In Acharya (1985), two rates of nominal interests were tried which produced quite diverse estimates of the unaccounted income. Acharya then followed the method suggested by Tanzi to estimate the unaccounted income. In brief, the procedure is as follows:

a) estimate the model in (1) using the published data and predict Y' from the estimated equation;

b) estimate model (1) assuming the values of X_{3t} as zeroes and obtained the predicted values Y'';

c) calculate $(Y' - Y'')$ which is tax induced demand for currency (TICD);

d) subtracting TICD from M_2 the "legal money" is obtained;

e) using the recorded income and the "legal money" income velocity is calculated;

f) multiplying the calculated income velocity by the TICD, estimates for unaccounted income are obtained. For further details of this procedure see Acharya (1985, chapter 4). Estimated unaccounted income with the updated data is consistently higher than the estimates reported in Acharya(1985) for the overlapping years.

The estimated figures from Acharya are:

Table 1:

Estimates from Acharya (1985)

Year	**Unaccounted Income (in crores)**
1960-61	5472.89
1965-66	11912.52
1970-71	16240.36
1975-76	43482.03
1980-81	48648.29

(Source: Acharya(1985), table 4.5.1 page 71)

The growth of unaccounted income between 1960 to 1965 was 117.66% but between 1965 to 1970 the growth was only 36.33%. If the growth of unaccounted income and the growth of corruption have one to one correspondence then this implies that corruption was increasing at a much faster rate between 1960 to 1965, but the rate of increase slowed down between 1965 to 1970. Similarly, the figures for the unaccounted income suggest that between 1970 to 1975 the corruption went up by 167.74% but increased by only 11.88% during 1975 to 1980. These fluctuations in the growth of corruption implied by the estimates of the unaccounted income can be justified from individual events, but we propose to examine this further using the updated and disaggregated estimates of the unaccounted incomes.

First we obtained estimates of the unaccounted income following Acharya's procedure but using the data for the period 1950-51 to 1992-93. Our experiment differs from Acharya (1985) as we have used a different interest rate series. In our updated study we find the signs of the estimated parameters are the same as those reported in Acharya (1985) except for the intercept term. However, we find the effect of tax rate on the currency demand is much higher than the result reported in Acharya. Consequently the estimates of the unaccounted income are higher in our experiment when compared with the estimates reported in Acharya. The estimates are obtained in two different ways for the updated data. First, we replicated the procedure used by Acharya (1985) described earlier. Then, we consider that the agriculture sector is not fully monetized as the farmers following past tradition keep some of the productions for own consumption and some of the exchanges follow barter exchange. Hence, we adjust the reported national income data by subtracting 60% of agricultural GDP from the total GDP. This adjusted national income data thus incorporate the non-monetization effect of the agricultural sector on currency demand. We have presented below two series of estimated unaccounted income one with the correction for monetization in the agricultural sector and the other without such correction.

Table 2:

Estimates obtained from updated data following Acharya's method (figures are in crores)

Year	Y_{unag}	% increase	Y_{unwag}	% increase
1960	16138.99		12320.18	
1965	30903.62	91.48	23277.95	88.94
1970	35811.17	15.88	26879.44	15.47
1975	98178.55	174.16	3638.24	173.96
1980	125124.44	27.45	94152.75	27.86
1985	71472.41	37.04	132110.14	40.31
1990	201677.97	17.62	159226.75	20.53
1991	210260.20	4.08	166241.89	4.41
1992	249337.52	18.58	195184.67	17.41

Y_{unag} = Unaccounted income calculated without any monetization correction for agricultural sector (figures are in crores).

Y_{unwag} = Unaccounted income calculated with the assumption that 40% of the agricultural sector is non-monetized.

The general movement of these estimates is similar to what have been observed using Acharya's estimates for the period 1960-1980. Our estimates of the unaccounted income also suggest a big increase during the period 1970 to 1975. This can be attributed, at least partly, to the creation of Bangladesh and the increase in corruption leading to the 'Emergency' of 1975. The other significant increase in the unaccounted income was observed between 1991 and 1992. This suggests that there has been some increase in corruption during the year 1991-93 compared to the previous eight years. However, it should be noted that the total unaccounted economy has at least two components: (1) that generated due to corruption in the economy; (2) that generated for other reasons. Thus, the increase in the unaccounted income ($Y_{u,t}$) can be written as equal to:

$$\frac{\Delta Y_{u,t}}{Y_{u,t-1}} = \frac{\Delta Y_{c,t}}{Y_{c,t-1}} [(1 + \frac{\Delta Y_{o,t}}{\Delta Y_{c,t}})/(1 + \frac{Y_{o,t-1}}{Y_{c,t-1}})]$$

where, $Y_{c,t}$ is the unaccounted income generated through corruption and $Y_{o,t}$ is the unaccounted income generated by actions not included in corruption at time t. It is easy to see the that the rate of increase in the unaccounted income will coincide with the rate of increase in the corruption if $Y_{c,t}$ and $Y_{o,t}$ have a constant rate of change over time. However, in practice the changes will not be constant and therefore the rate of increase of corruption measured from the estimated unaccounted income will be, to some extent, an approximate measure. Using the aggregate data on unaccounted income, the estimated rates of increase in corruption were not very high except for the period 1970 to 1975. However, the anecdotal evidences suggest that the increase in the corruption rates have been much higher during the last twenty-five years. Hence, a further examination of corruption through the unaccounted income of the industrial sector will be more informative. This we present in the following sections.

A Method of Estimating the Sectoral Hidden Economy

To obtain estimates of the sectoral hidden economy we extend the procedure suggested in Bhattacharyya(1990). Suppose there exists an economic activity M that appears in k different sectors of the economy and satisfy the relation:

$$M_t = M_{1t} + M_{2t} + ... + M_{kt}. \qquad (1)$$

In practice we observe only M_t, but we know the variables that stochastically determines M_{1t}, M_{2t}, ..., M_{kt}. The recorded sectoral incomes (or the value of the productions) are Y_{1t}, Y_{2t}, ..., Y_{kt} and the unrecorded incomes (or productions) are $Y_{1t}, Y_{2t}, ..., Y_{kt}$ and the unrecorded incomes (or productions) are $Y_{1ht}, Y_{2ht}, ..., Y_{kht}$ The actual outcome of M's are also affected by other vectors of vaiables Z_1, Z_2 ...., Z_k in k sectors respectively. Thus, we can write $M_{jt} = f_j(Y_{jt}, Y_{jht}, Z_{jt}) + U_{jt}$. Therefore, from equation (1) we have:

$$M_t = f_1(Y_{1t}, Y_{1ht}, Z_{1t}) + U_{1t} + f_2(Y_{2t}, Y_{2ht}, Z_{2t}) + U_{2t} + ... + f_k(Y_{kt}, Y_{kht}, Z_{kt}) + U_{kt} \qquad (2)$$

In equation (2) except $Y_{1ht}, Y_{2ht}, ..., Y_{kht}$ and the disturbance terms $U_{1t}, U_{2t}, ..., U_{kt}$ all other variables are observable. We assume that the U's are 'white noise'. Thus, in equation (2) we can write the disturbance term as U_t where $U_t = \sum_i U_{it}$, which is also 'white noise'.

The method suggested in Bhattacharyya (1990) can be interpreted as a method of finding proxies for $Y_{1ht}, Y_{2ht}, ..., Y_{kht}$ which produces best fit for the equation (2) in terms of standard diagnostic tests. In empirical studies, we first set the form of the proxies and then estimate all the parameters of the model. Among the group of best fit proxies, we choose the one which has the best diagnostic statistics. The proxies can take many forms and this could be the main focus of some future empirical research. We describe here a particular approach which approximately coincides with the proxy used in the Bhattacharyya (1990), where a quartic function of Y_{jt} was used as a proxy for Y_{jht}. This proxy was justified as a special form of RESET proxy. However, in Bhattacharya & Bhattacharyya (1993), we have shown that in a linear regression, if one of the regressor is unobservable and if

$E(Y_{jht}|Y_{jt}) = g(Y_{jt})$ (a non-linear function of Y_{jt}) then using g (Y_{jt}) in place of Y_{jht} we obtain estimates for the parameters of the model which have desirable large sample properties. Therefore, it is possible to interpret the quartic function in Y_{jt} as the conditional mean of Y_{jht} given Y_{jt}.[2]

From the best fitted equation we would be able to choose $g(Y_{jt})$ as an estimate for Y_{jht} (unrecorded income or production). The conditional variance of Y_{jht} (estimated unrecorded income of the j^{th} sector) could be written as:

$$Var(Y_{jht}|Y_{jt}) = Var(g(Y_{jt})).$$

To estimate the unrecorded income of different sectors in India we used the primary indicator as notes and coins in circulation which we have denoted as M_t, while the economy is assumed to have three sectors: (a) Agriculture; (b) Industry; and (c) Service. The total amount of notes and coins in circulation (M_t) are observed and by standard accounting procedure we have:

$$M_t = M_{st} + M_{it} + M_{at} + M_{rt} \qquad (3)$$

Here at time t:

M_{st} = currency in circulation in the service sector;
M_{it} = currency in circulation in the industry sector;
M_{at} = currency in circulation in the agriculture sector;
M_{rt} = residual currency in circulation which is assumed very small relative to other sectors.

2 As a generalization of this procedure one can consider that the conditional mean(s) is(are) derived by conditioning on more than one variable.

For empirical implementation we assumed that:

$$\ln M_t \approx \ln M_{st} + \ln M_{it} + \ln M_{at} + V_t; \qquad (4)$$

where, $E(V_t) = \mu$; and $Var(V_t) = \lambda$; (and both are constant).

The demand for currency by each sectors are specified as:

$$M_{st} = A_1(Y_{st} + Y_{sht})^{\beta_1} P^{\beta_2} R^{\beta_3} e^{u_{1t}} \qquad (5)$$

$$M_{it} = A_2(Y_{it} + Y_{iht})^{\beta_4} P^{\beta_5} R^{\beta_6} e^{u_{2t}} \qquad (6)$$

$$M_{at} = A_3 Y_{at}^{\beta_7} P^{\beta_8} R^{\beta_9} e^{u_{3t}} \qquad (7)$$

The symbols used in equation (5), (6) and (7) are:

Y_{st} - the recorded income of the service sector at time t;

Y_{sht} - the unrecorded income of the service sector at t;

Y_{it} - the recorded income of the industrial sector at t;

Y_{iht} - the unrecorded income of the industrial sector at t;

Y_{at} - the recorded income of the agriculture sector at t;

P_t - the consumer price index at time t;

R_t - short run interest rate at time t;

u's are the disturbance terms with mean zeroes and constant (but different) variances;

A's and β 's are the unknown parameters of the model.

Substituting (5), (6) and (7) in (4) we have:

$$\ln M_t = (A_1+A_2+A_3) + \beta_1 \ln Y_{st} + \beta_4 \ln Y_{it} + \beta_7 Y_{at} + (\beta_2+ \beta_5+ \beta_8)\ln P_t$$
$$+ (\beta_3+ \beta_6+ \beta_9)\ln R_t + \beta_1 \ln(1+Y_{sht}/Y_{st}) + \beta_4 \ln(1+Y_{iht}/Y_{it}) + (u_{1t}+u_{2t}+u_{3t}+V_t) \qquad (8)$$

Denoting $v= V - \mu$, and $A = A_1+A_2+A_3+ \mu$; we replace the disturbance terms in (8) as $(u_1+u_2+u_3+v)$ and the intercept term by A. We also used the linear approximations,

$$\ln(1+Y_{sht}/Y_{st}) \approx Y_{sht}/Y_{st}$$

and $\ln(1+Y_{iht}/Y_{it}) \approx Y_{iht}/Y_{it}$

Following the general procedure described earlier we used a polynomial in Y_{st} as proxy for Y_{sht} and a polynomial in Y_{it} and P_t as the proxy for Y_{iht}. The equation was then estimated by ordinary least squares, and from the best fitted equation we used the estimated proxies as the estimates for the 'hidden economy' of the service and industry sector respectively. The search for proxies is a difficult proposition in this case, as we have no prior information about the functional form of the conditional means of Y_{sht} and Y_{iht}. Therefore, we had to search among a large number of alternatives, and chose two sets of estimates of the unrecorded economy which satisfy the following criteria:

1. The estimates of the sectoral unrecorded economy are derived from the fitted equation that satisfies standard diagnostic tests.

2. The qualitative predictions from the fitted equation are plausible.

3. The size and the trend of unrecorded estimates match the anecdotal evidence published in the news papers.

4. The estimates of unrecorded incomes, when used in a standard macro models (such as consumption function and investment function), provide plausible results.

We obtained estimates and their standard errors for the unrecorded incomes generated in the service and the industrial sectors. However, only the estimates of the unrecorded incomes are presented below which we need for further analysis and discussion.

Table 3:

A set of estimates following Bhattacharyya's method

Year	Y_{sht}	% Increase	Y_{iht}	% Increase
1960	46.8		1.00	
1965	119.2	154.7	4.84	380.0
1970	298.4	150.3	21.82	354.2
1975	1086.9	264.2	159.70	632.6
1980	3047.6	180.4	1055.85	561.2
1985	10725.1	251.9	9435.47	793.6
1988	19736.7	84.0	33980.05	260.1
1989	23272.0	17.9	54464.94	60.9
1990	26195.6	12.6	83426.98	53.2
1991	27074.1	3.4	118779.87	42.4
1992	23276.9	-14.0	169208.66	42.5

The year column in the tables refer to the agriculture year, thus the year 1960 is actually representing July, 1960 to June 1961. Hence, our analysis is based on information for the period July, 1950 to June, 1993. However, to keep the results comparable with Acharya (1985) we have presented estimates only for the period 1960 to 1992. We have observed in Table 2 that the estimates obtained by following a modification of Tanzi's method as suggested by Acharya are much higher than those reported by Acharya (1985). The estimates presented in Table 3 suggest that the total unrecorded incomes, i.e. the sum of unrecorded incomes for the service sector and the industry sector, are lower than those observed in Table 2. Thus, it is clear from the estimates in Table 3 that the estimated size of the 'hidden economy' in India has been overestimated in Acharya's study. However, we notice that the rate of increase in the unrecorded income in both sectors are higher than those derived in Table 2 in terms of the total unrecorded income. The percentage increase in the unrecorded income for both sectors in the 60s and 70s may be misleading because the actual size of the estimates were small. It is clear from the industry sector's estimates that the rate of increase of corruption has fluctuated substantially during this period. However, some of the large increases can be explained by particular events in those periods. For example, the large increase in the unrecorded income from 1970 to 1975 is associated with the Bangladesh freedom war. The continued increase of the unrecorded incomes of the industry sector for the period 1970 to 1985 also seems reasonable when we take into account the amnesty declaration of 'black money' by Rajiv Gandhi in 1984. Examining the yearly estimates we find that the increase in 1984 to 1985 was only about 25% which suggests that due to amnesty the rate of increase had decreased. The average yearly increase of the unrecorded income of the industry sector was about 85% during the period 1985 to 1988. After 1988 the rate of increases started slowing down then went up slightly during the years 1991 to 1992. Again, these movements can be explained through liberalisation policies taken up by the Government of India during 1990-91. Our estimates suggest that the initial effect of liberalisation was the reduction of the rate of increase of unrecorded income in the industry sector, but by 1992-93 the initial effect of liberalisation has started to wear off. If the CBI cases of corruption are studied according to year, our findings will receive further documentary support.

It may be important to note that when the unrecorded income due to corruption is increasing faster than the unrecorded income due to other reasons. The rate of increase measured from the unrecorded income will be much hire than the true rate of increase in corrupt income. Our estimates in Table 3, therefore, support the view that corruption was increasing at a very high rate in India during the period 1960-1992.

Conclusion

As far as we know, this is the first attempt to estimate the sectoral unrecorded income of India. We have suggested an estimation procedure and used this on Indian data. The current detection of corruption in India supports our estimation of the rate of increase in corruption from the estimated unrecorded income of the industrial sector.[3] Although we have mentioned two sets of plausible estimates for the unrecorded income of the industrial sector, only one series of estimates is presented here. Our decision is influenced by the fact that the alternative estimates produce a similar time series movement.

We believe that we have successfully demonstrated the link between corruption and the 'hidden economy' of the industrial sector. However, it is natural to ask how the government can use this information to reduce corruption. It is well known that to stop corrupt acts it is necessary to identify the acts and actions have to be taken through the countries judicial system. Following our approach it is possible to identify the times when corruption is increasing. Hence, our method of identifying corruption will allow the government to identify the timings of the corruption and to take necessary actions to prevent further increase of corruption.

The research in this area is in its early stages, and many more studies are needed both on the methodological issues, and on linking the estimates of the unrecorded income to other economic activities.

[3] We have obtained estimates for the sectoral 'hidden economy' where we followed the exact procedure suggested in Bhattacharyya(1990). These estimates are lower in magnitudes but the rate of increases is very similar to the one reported in this paper. We have chosen these estimates as they satisfy some other auxiliary tests.

Reference

Acharya, S. (1985). *Aspects of Black Economy in India*, National Institute of Public Finance and Policy, New Delhi.

Bardhan, Pranab, (1997). 'Corruption and Development: A Review of Issues', *Journal of Economic Literature*, vol.35, pp.1320-1346.

Bhattacharya, Rabi and Bhattacharyya, Dilip, (1993). 'Proxy and Instrumental Variable Methods in a Regression Model with One of the Regressors Missing', *Journal of Multivariate Analysis*, vol.47, no.1, pp. 123-138.

Bhattacharyya, D.K. (1990). 'An Econometric Method of Estimating the Hidden Economy, U.K. (1960-1984) : Estimates and Tests', *The Economic Journal*, vol.100, pp. 703-717.

-----------------. (1995). *Estimation of the 'Hidden Economy' of India at Macro Level*, report submitted to the ODA, 48 pages.

-----------------. (1996). 'Issues in Measuring the Hidden Economy', paper presented at the ESRC/Informal Economic Activity Workshop, CEPR, London.

Chopra, O.P. (1982). 'Unaccounted Income: Some Estimates', *The Economic and Political Weekly*, April 24.

Feige, E.L. (1979). 'How Big is the Irregular Economy', *Challange*, Nov-Dec.

Gupta, P and Mehta, R. (1981). 'An Estimate of Underreported National Income', *Journal of Income and Wealth*, vol.5, no.2.

Gutmann, P.M. (1977). 'The Subterranean Economy', *Financial Analysts Journal*, November/December.

Tanzi, V. (1983). 'The Underground Economy in the US: Annual Estimates, 1930-1980, *IMF Staff Papers*, vol.30, pp. 283-305.

INFORMAL ECONOMY RESEARCH CENTRE

ON THE USE OF

THE HIDDEN ECONOMY ESTIMATES

Dilip K Bhattacharyya

Working Paper No: 3

SPONSORED BY: DELOITTE & TOUCHE

ON THE USE OF THE HIDDEN ECONOMY ESTIMATES

D.K.Bhattacharyya,
University of Leicester and I.E.R.C., England

ABSTRACT

In this paper we examine the importance of the 'hidden economy' estimates in explaining the consumers' demand for durable goods using UK data over the period 1960(2) - 1984(4). The empirical results establish a significant relation between the current spending on durable goods and one period lag 'hidden economy' estimates. It is observed that the estimated demand function for durable purchases is less complex and improved when the 'hidden economy' estimates are included as an explanatory variable.

Key words: Hidden Economy, Durable Expenditure, Tax Evaded Income, Separate Propensity to Consume.

JEL Classification: C52 and E21.

Acknowledgement:

A part of the results presented here were presented at the Royal Economic Society's annual conference at Swansea. I have benefited from the comments I received from the participants of the conference. I am very grateful to Kevin Lee who has made elaborate comments on an earlier draft which helped me to remove some of the mistakes. I am responsible for any remaining errors.

November, 1998

ON THE USE OF THE HIDDEN ECONOMY ESTIMATES

D.K.Bhattacharyya,
University of Leicester and I.E.R.C., England

1. Introduction

In recent years, there has been world wide concern on the accuracy of the measurement of the Gross Domestic Product (GDP). The reasons for concern vary considerably from country to country and depend on the socio-economic condition of the country. At the moment, a number of pilot studies are being conducted under the general umbrella of an 'exhaustiveness study of the GDP/GNP measurement' initiated by the European Union and by the United Nation. The fruits of these investigations will not be available in the near future. At the same time, different authors continue to produce estimates for unrecorded GDP/GNP under the names 'underground economy', 'hidden economy', 'unrecorded economy' etc, for different countries and for different periods. A non-exhaustive list of such estimates includes Feige(1979), Tanzi(1983), Aigner et. al (1988) and Bhattacharyya(1990). However, these researchers have made very limited use of these estimates in terms of econometric or policy modelling. Potentially, the uses of the 'hidden economy' estimates can be classified into three categories: (1) they provide an indication to the government on the size of the missing tax revenues; (2) they provide useful information to policy makers in evaluating the effectiveness of their policies and regulations; (3) the information contained in these estimates can be profitably used by the empirical researchers and theorists in their modelling activities. All three possible uses of the 'hidden economy' estimates are not developed in this paper. Instead we examined its uses in empirical modelling of consumer expenditure on durable goods.

The 'hidden economy' estimates, and the associate standard errors, published in Bhattacharyya(1990)[1] are used in this paper to estimate the model under study. At the conceptual level, the 'hidden economy' estimates are an aggregate of several sub-activities which are not included in the recorded GDP. However, this type of aggregation problems also exist when we use published GDP figures in empirical modelling. Hence, it is justifiable to use the 'hidden economy' estimates in our empirical modelling as a distinct component of GDP. We assume that the published GDP corresponds to 'income with certainty' for a representative individual while the 'hidden economy' estimates represent the 'uncertain income' of the same individual.[2]

2. A Model of the Demand for Consumer Durables

Modelling the demand for consumer durable has proven to be a difficult task. This is apparent from the works of Stone and Rowe(1957), Ball and Drake(1963), Cuthbertson(1980), Muellbauer(1981) and others.[3] However, it is clear from all past

---

[1] The results presented in this paper are based on slightly revised estimates of the 'hidden economy' of the U.K. for the period 1960-1984. These estimates are available from the author on request. Estimated equations are available from the author based on the data published in Bhattacharyya(1990) which will demonstrate that the results presented in this paper are not significantly distorted due to revision of the 'hidden economy' estimates.

[2] The 'hidden economy' estimates can be used in a wide range of econometric studies. A number of such studies are to be published in the **Economic Journal-Controversies Sections in 1999.**

[3] Examples are given only from the studies relating to the U.K. A large number of other studies for the US and elsewhere have tackled the problem of specifying the model in a very similar fashion.

studies that income is an important determinant for consumer demand for durable goods. The 'hidden economy' estimates, defined as estimates for the unrecorded GDP, are assumed as a component of income and are likely to have a direct effect on consumer demand. In the specification of the model, we use the 'hidden economy' as a separate source of income for the following reasons:

1. the 'hidden economy' income can be considered as tax evaded income as supported by the estimated 'tax evasion' function in Bhattacharyya(1990);

2. given that this is tax evaded income, the way it is spent differs from the way recorded income is spent; in particular, the time pattern of receiving and spending the income will differ from that on recorded income;

3. the time series structure of the 'hidden economy'(income) estimates differs significantly from that of recorded income;

4. as the standard errors of the 'hidden economy' estimates are known, the uncertainty of income is open to explicit modelling in the empirical analysis.

It is already established in the literature that different sources of income have different propensities to consume (see, Holbrook and Stafford(1971) and Pollak, Shelly and Wales(1996)). Following this trend, we specified the model where the recorded income and the 'hidden income' are treated as two separate sources of income. The desired level of expenditure on consumer durable are determined by a number of variables which have been suggested in Muellbauer(1981) and Bhattacharyya(1978). The specification of the model can also be justified by intuitive economic reasoning. Thus, the desired level of expenditure on consumer durable is written as

$$Z^*_t = f(\Pi_t, H_{zt}, Y_{r,t}, Y_{h,t}, A_t, R_t); \qquad (1)$$

where,

Z = the level of durable goods that the individual desires to purchase at time t;[4]

P_t = the price index of consumer non-durable goods at t;

P_{zt} = the price index of consumer durable goods at t;

$\Pi_t = P_t/P_{t+1}$;

$H_{zt} = P_t/P_{zt}$;

$Y_{r,t}$ = the real recorded income at t;

$Y_{h,t}$ = the real value of the risk discounted hidden income at t;

A_t = the real value of liquid assets at t;[5]

R_t = the short term rate of interest at t.

This specification encompasses a variety of alternative models (e.g. 'the relative price model', 'bank credit and interest rate models' and 'liquid assets models') which have been investigated as separate models in Cuthbertson(1980). Our model includes intertemporal price substitution, durable and non-durable price substitution, different sources of income effects, and bank credit effects, as well as the effect of

[4] In defining durable goods we exclude all form of transports (car, planes etc.). In the empirical analysis the real expenditure on durable goods is taken as Z.

[5] In the formal derivation of the model from an intertemporal objective function the total asset holdings is the determinant variable rather than the liquid asset holdings. The implicit assumption behind our specification is that the non-liquid assets are not used to decide the level of expenditure on durable goods included in this model.

liquid asset holdings (although, in our study, we do not use independently measured 'liquid assets' as an explanatory variable). All past savings from the recorded and the unrecorded income constitute the total asset holdings of the individual. We assume that individuals hold their recent savings as 'liquid' assets before transferring into 'non-liquid' assets. Hence, the desired expenditure on durable goods can be written as,

$$Z^*_t = f(\Pi_t, H_{zt}, Y_{r,t}, Y_{h,t}, Y_{r,t-1}, \ldots, Y_{h,t-1}, \ldots, R_t); \qquad (2)$$

The desired level of expenditure on durable goods is not directly observable. However, we assume that Z can be approximated by $D(L)Z_t$, where $D(L)$ is a polynomial in the lag operator L and Z_t is the actual expenditure on durable goods at constant prices.

3. The Empirical Verification of the Model

The past literature (see Cuthbertson(1980)) and the economic theory suggests that income is the major determinant of the expenditure on durable goods. Therefore, it is necessary to establish the role of the 'hidden economy' estimates in defining the income variable. Would it be proper to add the real 'hidden economy' to the recorded real income? The answer depends on the time-structures of the real 'hidden economy' estimates and the real recorded income. Hence, using an autoregressive distributed lag (ADL) approach, and emphasizing 'parsimony' we observe the following time-structures for the two income series. We used the U.K. data for the period 1960-2nd. quarter to 1984-4th. quarter.[6]

$$Y_{r,t} = 0.7916Y_{r,t-1} + 0.5970Y_{r,t-4} - 0.3818Y_{r,t-5}; \quad (3)$$
$$\text{(S.E)} \quad (0.079) \qquad (0.089) \qquad (0.097)$$

Adj. $R^2 = 0.98$; LM-$\chi^2_{(8)} = 12.48$; Ljung-Box-$Q_{(8)} = 8.73$;

ADF(4 lags) = -5.59; AIC = 16.80; Schwarz Bayes IC = 14.04.

$$Y_{h,t} = 1.1934Y_{h,t-1} - 0.4586Y_{h,t-3} + 0.4205Y_{h,t-4}$$
$$\text{(S.E)} \quad (0.074) \qquad (0.140) \qquad (0.122)$$

$$- 0.1554Y_{h,t-8} \qquad (4)$$
$$(0.045)$$

Adj. $R^2 = 0.996$; LM-$\chi^2(8) = 12.58$; Ljung-Box-Q(8) = 12.62;

ADF(4 lags) = -10.66; AIC = -1.12; Schwarz Bayes IC = -3.85.

The fitted equations (3) and (4) confirm that the time-structures of the recorded real income series and the unrecorded real income series (real 'hidden economy' estimates) are distinctly different. This gives us the empirical support to treat the two income series as two separate sources of income.

The model stated in equation (2) is estimated in a log-linear form where all the variables are taken as their first differences. The real 'hidden economy' estimates are obtained by deflating the estimates by the consumer price index. The unrecorded income measured through the 'hidden economy' estimates are subject to chance

6 The period of analysis is restricted to 1984 as the published time series data for the 'hidden economy' estimates are available only for the period 1960 to 1984. As our primary objective is to establish the importance of the 'hidden economy' estimates in econometric modelling, the period of analysis is not our primary criterion.

variation. Hence, the uncertain estimates of the 'hidden economy' are replaced by the risk discounted hidden incomes. The risk discounting have been achieved in two different ways. In the first specification we assumed that only the 'hidden economy' estimates are available and the estimates correspond to the individual's expected income from the unrecorded sources. Individuals discount for risk by multiplying the expected income by a constant α where $0<\alpha<1$. This discounting factor may vary from individuals to individuals but is assumed fixed over time. In the second approach to measuring the risk discounted hidden income, we utilise, the standard errors of the 'hidden economy' estimates. In this case, following the 'certainty equivalence' principle, we assume that the risk discounted hidden income can be written asY_{ht} - b.Var(Y_{ht}), where b is a constant parameter.

The results presented here are obtained using a search procedure similar to the 'general' to 'specific' approach. However, the general approach adopted here satisfy the approach partially as only the recorded income, the unrecorded income(hidden economy) and the expenditure on durable goods are lagged by several periods as the 'general' specification. The empirical results suggest that one period lag with differenced variables were sufficient to produce plausible and statistically acceptable empirical results. We observed that the 'stationarity' for all variables are achieved by first differencing. The estimated expenditure function for durable goods obtained by including the first definition of the risk-discounted 'hidden economy' is presented below,

$$\Delta_1 z_t = -0.0072 + 1.0337\Delta_1 y_{r,t} - 0.1057\Delta_1 y_{r,t-1} - 0.2139\Delta_1 y_{h,t}$$
(S.E) (.008) (.282) (.312) (.264)

$$+ 0.6567\Delta_1 y_{h,t-1} - 0.5325\Delta_1 \pi_t + 1.3009\Delta_1 h_{z,t} - 0.1020\Delta_1 r_t$$
(.238) (.550) (.646) (.043)

$$- 0.1354\Delta_1 z_{t-4} + 0.1257\Delta_1 SPD_t + 0.0685\Delta_1 THD_t \qquad (5)$$
(.073) (.015) (.060)

Adj.R^2 = 0.539; = 0.056; Durbin-h = -0.825; LM$(\chi^2)_{(8)}$ =10.07;

Ljung-Box-$Q_{(8)}$ = 8.77; ADF(4 lags) = -10.373; ARCH = 0.330;

Chow Test (break at 1975:4) = 1.431; Jarque-Bera test = 2.10;

Log likelihood = 144.255.

(All lower case letters are logarithm of the capital letters defined earlier.)

The diagnostic statistics suggest that the model fits the data well. The results presented in (5) show that the changes in the current recorded income have a strong positive effect on durable expenditure while the change in the lagged one period recorded income has a negative but statistically insignificant effect. Similarly, the change in the current 'hidden economy' has a negative effect but it is not statistically significant, whereas the one period lag change has significant positive effect on the durable expenditures. If we consider the 'hidden economy' estimates to be a measure of tax-evaded income, then the spending behaviour associated with the 'hidden economy' is consistent with the risk-averting behaviour expected from a tax evader. The significant negative effect of interest rate supports the credit constraint model of the durable purchases. The sign of the estimated coefficient of π_t suggests that if consumers expect future prices of the non-durable goods to increase[7] then the

[7] In our empirical analysis we assumed perfect knowledge of the next period's price level. We tried some random corrections for the next period price level but the results obtained are not

purchases for the durable goods will increase in the current period (although the estimated parameter is not significantly different from zero). In contrast, the price ratio between non-durable goods and the durable goods ($h_{z,t}$) has a significant effect on the durable purchases and the sign of the coefficient satisfies the standard prediction of the economic theory. As the durable purchases are not everyday purchases, the past expenditures on the durable goods have effect on the current purchases. We tried up to eight period lag effect and found only the four period lag purchases have a significant effect at 10% level of significance. Thus, we find that all the economic variables affect the durable purchases in the ways that have been expected in the economic theory. In addition, the dummy variable SPDUM is included in the model to capture the announcements and increases in the VAT during the period of study. The estimated coefficient of SPDUM is significantly different from zero and, therefore, an integral part of the model.[8]

All of the available evidence suggest that during 'Thatcher' period[9] income inequality had increased. This suggest that a part of the economy became relatively richer with higher income and was in a position to increase durable purchases. However, income inequality changes are slow in macro data and not regularly calculated for each quarter. Therefore, to capture this effect we introduce a dummy variable THADUM which took the value 1 during Thatcher years and 0 elsewhere.

---

significantly different from the results presented here.

[8] The dummy variable was formulated in the same way as in the Davidson et.al(1978).

[9] This is the period 1979 to 1990 when Margaret Thatcher was Prime Minister of the U.K.

This dummy variable has a positive effect on the durable purchases although the point estimates is not statistically significant.

The model of durable purchases was re-estimated after dropping all non-significant variables to obtain a more 'specific' model. However, to keep the theoretical foundation valid we included π_t in this reduced model.

$$\Delta_1 z_t = -0.0098 + 0.8901\Delta_1 y_{r,t}$$
(S.E) (.008) (.208)

$$+ 0.5920\Delta_1 y_{h,t-1} - 0.3884\Delta_1 \pi_t + 1.5323\Delta_1 h_{z,t} - 0.0964\Delta_1 r_t$$
(.198) (.462) (.611) (.042)

$$- 0.1371\Delta_1 z_{t-4} + 0.1235\Delta_1 SPD_t \qquad (6)$$
(.072) (.015)

Adj.R^2 = 0.543; = 0.055; Durbin-h = -1.138; LM$(\chi^2)_{(8)}$ =12.50;

Ljung-Box-$Q_{(8)}$ = 9.24; ADF(4 lags) = -10.73; ARCH = 0.258;

Chow Test (break at 1975:4) = 2.192; Jarque-Bera test = 1.503;

Log likelihood = 143.056.

The estimated equation presented in (6) still fits the data well and demonstrates the strong effect of the 'hidden economy' on the durable purchases. By dropping some regressors only the Chow test and the propensity to consume from the recorded income have been affected. The estimated coefficient of $\Delta_1 y_{r,t}$ in equation (6) seems to be more plausible than greater than 1 observed in equation (5).

The model was re-estimated using the alternative definition of the risk-discounted hidden income mentioned earlier. The other specifications of the model remained the same as (5). However, due to additive nature of the risk discounting the variance of the hidden economy estimates entered in our specification as an additional explanatory variable. In this approach the parameter 'b' inY_{ht} - b.Var(Y_{ht}) is not identifiable. As our specification of the model is in logarithmic variables the risk-discounted hidden income is approximated as

y_{ht} - b(Var(Y_{ht})/Y_{ht}) in the empirical analysis.

The estimated model parallel to (5) is:

$$\Delta_1 z_t = \underset{(.009)}{-0.0067} + \underset{(.289)}{1.0349}\Delta_1 y_{r,t} - \underset{(.319)}{0.0946}\Delta_1 y_{r,t-1} - \underset{(.277)}{0.2312}\Delta_1 y_{h,t}$$
(S.E)

$$+ \underset{(.257)}{0.6701}\Delta_1 y_{h,t-1} - \underset{(.562)}{0.5264}\Delta_1 \pi_t + \underset{(.678)}{1.2972}\Delta_1 h_{z,t} - \underset{(.044)}{0.1063}\Delta_1 r_t$$

$$- \underset{(.073)}{0.1369}\Delta_1 z_{t-4} + \underset{(137.7)}{75.6192}\mu_t - \underset{(133.3)}{92.6554}\mu_{t-1}$$

$$+ \underset{(.015)}{0.1258}\Delta_1 SPD_t + \underset{(.061)}{0.0755}\Delta_1 THD_t \qquad (7)$$

where, μ_t = (Var(Y_{ht})/Y_{ht})

Adj.R^2 = 0.530;   = 0.056; Durbin-h = -0.666; LM$(\chi^2)_{(8)}$ =10.38;

Ljung-Box-$Q_{(8)}$ = 8.51; ADF(4 lags) = -10.211; ARCH = 0.471;

Chow Test (break at 1975:4) = 1.338; Jarque-Bera test = 2.25;

Log likelihood = 144.536.

The diagnostic statistics for the fitted equation (7) are very similar to what we have observed in equation (5). In addition, we observe that the estimated parameters associated with μ's are not significantly different from zero. Hence, it can be concluded that the variance of the estimated hidden economy provides no additional information for this model.[10] The implied estimates of 'b' calculated from the coefficients of μ_t and μ_{t-1} are 327.11 and 138.28. The large difference between the two estimates brings doubt about their robustness. However, it is comforting to note that for the both estimates of 'b' in the risk-discounted hidden income (y_{ht} - b.Var(Y_{ht})) are positive for all t and thus satisfy the theoretical restriction on the parameter 'b'.

Finally, we re-estimated the version of the model in equation (6) with the risk discounted hidden income asy_{ht} - b(Var(Y_{ht})/Y_{ht}). The estimated equation that is parallel to (6) is:

$$\Delta_1 z_t = -0.0095 + 0.8875\Delta_1 y_{r,t} + 0.5866\Delta_1 y_{h,t-1}$$
(S.E) (.008) (.210) (0.204)

$$- 0.3884\Delta_1\pi_t + 1.5323\Delta_1\pi_{z,t} - 0.0964\Delta_1 r_t$$
(.462) (.611) (.042)

$$-0.1371\Delta_1 z_{t-4} \quad -9.7930\mu_{t-1} \quad + \quad 0.1235\Delta_1 SPD_t \qquad (8)$$
(.072) (83.66) (.015)

Adj.R^2 = 0.538; = 0.056; Durbin-h = -1.120; LM$(\chi^2)_{(8)}$ =12.33;

Ljung-Box-$Q_{(8)}$ = 9.23; ADF(4 lags) = -10.72; ARCH = 0.245;

Chow Test (break at 1975:4) = 1.98; Jarque-Bera test = 1.518;

Log likelihood = 143.064.

---

[10] The variance of the hidden economy appeared in our model in a particular way. However, in a different modeling context the variance may appear in a different manner in the estimated model. Hence, the inference drawn in this case should be read with caution.

The results presented in equation (8) are very similar to the results presented in (6). Once again we observe that the variance of the hidden economy has no significant effect on the durable purchases. However, the implied estimate of 'b' from equation (8) is 16.69. This is much lower than what we estimated from equation (7) and appears to be more plausible.

4. Conclusion

The empirical results presented in this paper demonstrate the importance of the 'hidden economy' in modelling durable purchases. This is one of the many possible uses of the 'hidden economy' estimates in econometric modelling. However, the empirical results allow us to assert that:

1. the 'hidden economy' as an explanatory variable has a significant effect on the durable purchases by the consumers;

2. the 'hidden economy' effect takes at least a quarter to influence consumer's decision;

3. there is a clearly established link between the 'hidden economy' and the recorded economy;

4. the empirical results replicate earlier findings that there are different propensity to consume out of income arising from different sources;

5. with the use of the 'hidden economy' estimates we obtain a simple model for durable purchases which fits the conventional theory.

On the basis of the empirical results presented here, we believe that the 'hidden economy' estimates will be useful information for model builders in many situations. For example, in Bhattacharyya(1996b) it was observed that the use of the 'hidden economy' estimates substantially improved the estimated government expenditure function in terms of diagnostic tests and offer clearer explanation on the existence of 'Wagners Law'. The present study does not find any significant effect on the durable purchases by the variances of the 'hidden economy' estimates, although the uncertainty of the 'hidden income' is adequately captured by the constant factor discounting used in the first specification. It is possible that with different models the variance of the 'hidden economy' estimates may exhibit significant effect on the outcome.

DATA SOURCES

■ The 'hidden economy' estimates are close to that was reported in Bhattacharyya(1990).

■ The price indices are collected from the Department of Employment publication, *Retail Prices - Indices 1914-1986*, HMSO, London.

■ Short term interest rates are collected from *Financial Statistics* as the London clearing banks' base rate.

■ Data for expenditures on durable goods and recorded income are taken from the *Economic Trends,* 1992 Annual Supplement.

References

Aigner,D.J., Schneider, F and Ghosh, D. (1988). 'Me and my shadow: estimating the size of the hidden economy from time series data', in W.A.Barnett et al. (eds.), *Dynamic Econometric Modelling: Proceedings of the Third International Symposium in Economic Theory and Econometrics*, Cambridge University Press, pp.297-334.

Ball,R.J., and Pamela S. Drake. (1963). 'The impact of credit control on consumer durable spending in the United Kingdom, 1957-1961', *Review of Economic Studies*, vol.30, pp.181-94.

Bhattacharyya,D.K. (1978). *Demand for financial assets: an econometric study of the U.K. personal sector*, Saxon House, England.

Bhattacharyya,D.K. (1990). 'An econometric method of estimating the "hidden economy", United Kingdom(1960-1984): estimates and tests', *Economic Journal*, vol.100, pp.703-717.

Bhattacharyya,D.K. (1996a). 'The hidden economy of the U.K.: revised and updated estimates and tests', paper presented at the *Royal Economic Society's* annual conference, Swansea, U.K.

Bhattacharyya,D.K. (1996b). 'The hidden economy estimates and their implications to government expenditure: U.K. (1960-1990)', paper presented at the *European Economic Association's Congress*, Istanbul, revised version out as *I.E.R.C* working paper no.1.

Bhattacharyya,D.K. (1999). ' On the Economic Rationale of Estimating the Hidden Economy', *Economic Journal*, vol.109, pp.F348-F359.

Cuthbertson, Keith. (1980). 'The determination of expenditure on consumer durables', *National Institute Economic Review*, pp.62-72.

Davidson,J.E.H., Hendry,D.F., Srba,F. and Yeo,S. (1978). 'Econometric modelling of the aggregate time series relationship between consumers' expenditure and income in the UK', *Economic Journal,* vol.88, pp.661-692.

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

Holbrook, Robert and Stafford, Frank. (1971). 'The propensity to consume separate types of income: a generalized permanent income hypothesis', *Econometrica*, vol.38, pp.1-21.

Muellbauer, J.N.J. (1981). 'Testing neoclassical models of the demand for consumer durable', in Deaton, A. (ed.), *Essays in the theory and measurement of consumer* *behaviour*, Cambridge University Press, pp.213-236.

Pollak, Robert, Lundberg, Shelly, and Wales, Terry. (1996). 'Do husbands and wives pool their resources? Evidence from the U.K. child benefit', paper presented at the annual meeting of the *American Economic Association*, San Francisco.

Stone, Richard, and Rowe,D.A. (1957). 'The market demand for durable goods', *Econometrica*, vol.23, pp.423-443.

Tanzi, V. (1983). 'The underground economy in the United States, 1930-80', *I.M.F. Staff Papers*, vol.30, pp.283-305.