# CHAPTER 7

# UNDERGROUND ECONOMY - VARIABLE RELATIONSHIPS

## Introduction

This chapter discusses on the regression models of UE estimates on CPI, GDP per capita, federal expenditure, crime rate, unemployment rate, corruption index, tax revenue, electricity consumption, and cash in circulation and M1 aggregates (as discussed in paragraphs 3.9 of Chapter 3). The structural characteristics (symmetrical distribution, linearity relationship, stationary and variable association) of UE estimates in three level time series (low, moderate and upper bound size) were first examined and corrected for appropriate regression models. UE models of best fit imply appropriate size and recommended for policy measures.

## 7.1 Structural characteristics of data

All time series data showed significant evidence of unit roots in level form. But their first differencing (taking the difference between succession values of a time series) showed no significant evidence of unit roots. In other words, all variables selected at level, have unit root and are non-stationary, but are stationary when they are first differenced at 5% level of significance. Table 7.4 in the Appendix of Chapter 7 summarises the results of unit root tests.

Line graph in Figure 6.3 of Chapter six shows that UE growth over the period of 1980-2009 fluctuates distinguishably high in bad time. However its volatility is based on variance test is modest. The variance of UE is relatively lower (0.4169 – 0.4630) than other macro variables; GDP (0.5082), M1 aggregates (0.6525), electricity (0.7306), GDP per capita (0.5921), total crime (0.4212), direct tax revenue (0.8988) and indirect tax revenue (0.5426). However, it is larger than the variance of perception bribery index (0.0315), CPI (0.2315) and employment (0.2683). All time series data plotted in log form approximate linear relationship but exhibit an upward trend in level form.

Figures 7.1 and 7.2 in the Appendix of Chapter 7 illustrates linear graph in level and log forms respectively. Table 7.1 in the Appendix of Chapter 7 summarises the results of normality tests.

The UE series over the period of 1980-2009 showed no significant evidence of no-co-integration with the hypothesized variables. In other words, there is significant evidence to reject the null hypothesis of no co-integration As there is no evidence of co-integration relationship, the correlation coefficients of the variables in OLS double log models are appropriate. Table 7.5 in the Appendix of Chapter 7 summarises the co-integration tests.

All time series data of UE estimate and other explanatory variable displayed strong time trend and are highly correlated, by the correlation matrix statistics. Table 2 in the Appendix of Chapter 7 summarises the results.

However, close correlations between paired variables (UE and explanatory variables) lack evidence of causal relationship by pair wise - Granger causality

test. Nevertheless, considering limitation of Granger test, it is not conclusive to imply mere causal relationship as the variables in potentially complex economic transactions could interact in a causal manner. Table 7.6 in the Appendix of Chapter 7 summarises the results of Granger causality. Only the upper UE level shows uni-causal relationship with Federal expenditure.

The appropriate lag forms are selected between paired relationships, based on Schwartz-Bayesian Criterion to estimate the number of optimal lags in this model. UE is individually significant to lag of; tax revenue (lag of 1), GDP (lag of 1), Federal expenditure (lag of 1 and 2) and M1 and cash in circulation (lag of 1 – 7), while others have been attributed with no lag at all.

## 7.2 Full double log model

The estimated UE series (low, moderate and upper level or series) were first regressed on two sets of maximum number of explanatory variables. These two full models differ by the components of GDP per capita and CPI as they are highly correlated. Basic equations of full model functions are as follows:

$UE_i = F$ ([pc*, c, M1, fe, el, dtx/indtx, cr) and
$UE_i = F$ ([CPI*, c, M1, fe, el, dtx/indtx, cr)

Where, pc = GDP per capita income; CPI = consumer price index; c = cash in circulation; M1 = money aggregates; fe = federal expenditure; el = electricity consumption; dtx = direct tax revenue; indtx = indirect tax revenue; and cr = pre-post 1998-99 crises (dummy of 0,1). Table 7.7 in the Appendix of Chapter 7

summarises the definitions of these variables. Table 7.8 of Appendix of Chapter 7 summarises the estimated coefficients and statistical. Although the coefficients of federal expenditure, per capita income, CPI and electricity consumption are significant, models indicate some evidence of miss specification. As the full models are not well specified likely due to multi collinear effects UE estimates were then examined in reduced models.

### 7.3 Reduced double log model

Reduced models were constructed in accordance with priori believes such that better estimates and stable ECM models could be obtained based on parsimonious principle (Gujarati 1995). The distant correlated variables were grouped together in a set of explanatory variables and the highly correlated variables were separated in different set of explanatory variables as alternative models. The estimated UE series (low, moderate and upper level) were then regressed on various combinations of distant explanatory variables and evaluated for any significant economic association. Basic equations of reduced model functions are as follows:

$UE_i$ = F ([pc or CPI]* or [c or M1]*, fe, el, *[dtx or indtx], cr)

Note:

- i represents UE either in the lower, moderate or upper series.
- *Highly correlated variables are inserted in different models.

The econometric tests of the reduced models; indicate coefficients are statistically significant at 5% level with highly explained variation; consistently showed expected signs as priori; no evidence of "abnormalities" [serial correlation, heteroskedascity, model specification error]. Variable correlations (coefficient, significant level and variances) are of the same signs but differ in magnitudes across UE sizes and constructed sets of explanatory variables. The variable coefficients of these models are likely to exhibit long run relationship as VECR or auto regression correction (AR(1)) insertion is insignificant. The models with the least variation [AIC and SIC], are assumed to exhibit reasonably well behaved residuals, generate valid and robust models of reliable estimate.

Table 7.11 and Table 7.12 in the Appendix of Chapter 7 summarise alternative "nominal models" and "real models" respectively. Table 7.9 and Table 7.10 in the Appendix of Chapter 7 summarise the statistical tests for the estimated coefficients for nominal models and real models respectively. Table 7.3 denotes significant economic coefficients for easy reference.

Tables 7.1 and 7.2 summarise the equations of six best fit models consisting of a set of common explanatory variables in nominal and real term respectively. Table 7.2 summarises equations model of strong positives association between UE and CPI and economic crises.

Most variable coefficients indicate significant t-values (as in the parenthesis of the respective coefficients) and show the expected signs at 5% significant error level. Approximately 90% of the variances in the selected UE models are statistically explained.

Table 7.1 UE models – valid in both nominal and real terms

| | | |
|---|---|---|
| **GDP model** | | |
| Nominal term | Us UE | = 28.2238 + 1.8154 pc – 1.2282 fe + 0.9104 el<br>( 6.7711) (3.3702) (-4.3766) (2.3238) |
| Real term | Us UE | = 18.4577 + 0.5783 pc – 0.8941 fe + 1.2665 el<br>(7.1543) (1.5357) (-4.0232) (5.8975)<br>With evidence of heteroskedasticity |
| **Indirect Tax Model** | | |
| Nominal term | Ls UE | = 9.3417 + 4.2642 cpi – 0.8089 fe +0.5482 indtx -1<br>(2.7585) ( 5.57574) (-3.8456) (2.3782) |
| Real term | Ls UE | = 5.4216 + 3.1446 cpi – 0.5574 fe + 0.4615 indtx -1<br>(1.7468) (7.3728) (-3.3429) (2.1266) |
| **Cash Model** | | |
| Nominal term | Us UE | = 23.8182 + 1.0182 c – 0.8365 fe + 1.0706 el<br>(5.7650) (2.7667) (-3.1555) (2.6226) |
| Real term | Us UE | = 5.9858 + 0.9676 c – 0.7266 fe + 0.9210 el<br>(1.0810) (2.2078) (-4.1495) (3.0512 |
| **Lag model (Lag-1)** | | |
| Nominal term | Us UE | = 34.2208 + 1.9160 pc – 1.5774 fe + 1.0843 el<br>(8.6891) (3.7644) (-5.9523) (2.9308) |
| Real term | Us UE | = 22.0541 + 0.7401 pc – 1.1579 fe + 1.2881 el<br>(8.2867) (2.1582) (-5.3431) (5.8631) |

Note: Us UE = upper UE series estimates; Ls UE = lower UE series estimates
pc = GDP per capita, fe = federal expenditure, ei = electricity consumption, indtx-1 = indirect tax revenue of lag one year, c = cash in circulation, dtx-1 = direct tax revenue of lag one year, cr = pre and post 1998-99 economic crises dummy.

Table 7.2: UE models – only valid in nominal term

| **CPI Model** | | |
|---|---|---|
| Nominal term | Ls UE | = 20.1298 + 3.3076 cpi – 0.9844 fe + 1.1031 el<br>(5.5961) (3.5931) (-4.3486) (3.3513)<br><br>= 9.3417 + 4.2642 cpi – 0.8089 fe +0.5482 indtx -1<br>(2.7585) ( 5.57574) (-3.8456) (2.3782)<br><br>= 18.2523 + 3.9585 cpi – 1.3826 fe +0.8266 dtx -1<br>(6.8273) (4.9037 ( -4.1006) (2.4515) |
| Real term | Us UE | = 11.6719 + 2.5870 cpi – 0.8984 fe +0.6276 dtx -1<br>(5.9555) (3.8819) ( -3.3050) (1.8305)<br>With evidence of serial correlation and model specification error |
| **Crises Model** | | |
| Nominal term | Us UE | = 25.1938 + 3.6721 cpi – 1.5110 fe + 1.0193 el + 0.3464 cr<br>(6.1916) (4.1970) (-5.1472) (3.2929) (2.2024)<br><br>= 33.6325 + 1.4784 c – 1.2793 fe + 0.7366 el + 0.5352 cr<br>6.9350) (4.1694) ( -4.6798) (1.9809) (3.0121) |
| Real term | Us UE | = 9.4286 + 1.5741 cpi – 0.4305 fe + 0.9665 el + 0.1753 cr<br>(1.8336) (1.6088) (-1.4420) (2.8375) (0.8559) |

Note: Us UE = upper UE series estimates; Ls UE = lower UE series estimates
pc = GDP per capita, fe = federal expenditure, ei = electricity consumption, indtx-1 = indirect tax revenue of lag one year, c = cash in circulation, dtx-1 = direct tax revenue of lag one year, cr = pre and post 1998-99 economic crises dummy.

Table 7.3: Summary of significant economic coefficients

| | pc | fe | el | CPI | Indtx1 | c | cr |
|---|---|---|---|---|---|---|---|
| **UE in upper level series** | | | | | | | |
| GDP model | | | | | | | |
| Nominal term | √ | √ | √ | | | | |
| Real term | **X** | √ | √ | | | | |
| Cash model | | | | | | | |
| Nominal term | | √ | √ | | | √ | |
| Real term | | √ | √ | | | √ | |
| Lag model (UE-1) | | | | | | | |
| Nominal term | √ | √ | √ | | | | |
| Real term | √ | √ | √ | | | | |
| CPI model | | | | | | | |
| Nominal term | | √ | √ | √ | | | |
| Real term | | √ | √ | √ | | | |
| Crises model | | | | | | | |
| Nominal term | √ | √ | √ | √ | | | √ |
| Real term | **X** | **X** | √ | **X** | | | **X** |
| **UE in lower level series** | | | | | | | |
| Tax model | | | | | | | |
| Nominal term | | √ | | √ | √ | | |
| Real term | | √ | | √ | √ | | |

Note: Us UE = upper UE series estimates; Ls UE = lower UE series estimates
pc = GDP per capita, fe = federal expenditure, ei = electricity consumption, indtx-1 = indirect tax revenue of lag one year, c = cash in circulation, dtx-1 = direct tax revenue of lag one year, cr = pre and post 1998-99 economic crises dummy.

The magnitudes of the variable coefficients in the double log multivariate models are interpreted as growth elasticity. They vary across size of UE and differ between nominal and real term. Tax models indicate that UE in low level is responsive to tax burden. Crises models indicate that UE in the upper levels is responsive to economic crises. Differences between nominal and real models implied CPI effect. As models of UE in the upper level exhibit smaller variances

(based on AIC and SIC estimates) compared to UE in the lower level, UE in the upper level should be the appropriate estimate.

### 7.4 Model interpretation - nominal compared with real models

In general, UE in nominal terms correlates negatively with federal expenditure at level but positively with GDP per capita, CPI, cash in circulation, M1 aggregates, and electricity consumption. The correlations between UE and GDP per capita, federal expenditure and electricity consumption are also significant in one year delayed effect. UE in low level correlates positively with prior burden of indirect taxes but not economic crises. UE in upper level correlates with economic crises but not with indirect taxes.

UE in both nominal and real terms associate negatively with federal expenditure that are characterized by evidence of robust correlation in both level years and in one year delayed effect. Both nominal and real models indicate that UE is elastic to GDP per capita in one year delayed effect, cash in circulation in level years and electricity consumption at level year and in one year delayed effect. Both models show UE in low series correlate with indirect taxes in one year delayed effect.

As the variable coefficients in the models constructed on UE of upper level are larger and more persistent, with smaller variances compared to models of lower series, it is proposed that the upper estimates represent the actual size of UE that escape taxes.

Overall association showed that the main difference between nominal and real models of the same set of explanatory variables is the magnitude of correlation

coefficients. Real models are characterized by smaller to insignificant coefficients. Any differences between nominal and real models are likely due to CPI incremental as real variables were derived from nominal variables which is partly deflated by CPI.

### 7.4.1 Federal expenditure

UE at all level size estimates correlate negatively and highly with federal expenditure in both nominal and real models in level year and in lag of one year. The consistent and elastic correlation between UE and federal expenditure, before and after controlling for CPI implies a robust negative relationship. This suggests that the government has successfully interacted to suppress the UE. In other words the negative immediate and delayed significant effects indicate that federal expenditure is an effective remedy to suppress the UE. In the nominal models, the coefficient for federal expenditure coefficient is between 0.8089 and 1.6363, implying that for every one percent growth of federal expenditure, UE growth increases to as high as 1.6363%, provided other controlled variables in the models are constant (CPI, GDP per capita, tax, electricity, cash and economic crises, wherever relevant). In the real models, the coefficient for federal expenditure is between 0.5574 and 1.1579, indicating that for every one percent growth of federal expenditure, the UE growth increases to as high as 1.1579%, while other controlled variables in the model are constant.

Correlation between UE and all the explanatory variables show no causality relationship except for federal expenditure. UE at upper level indicate evidence of uni-causality with federal expenditure as follows:

Pairwise Granger causality tests (Sample 1:30), Lags 2 (Obs 28)

Ho: Ln Fe does not granger cause $LnUE_{high}$

[F-statistic; 0.6459. Probability; 0.5335];

(Federal expenditure does not cause UE to increase)

Ho: $LnUE_{high}$ does not granger cause Ln Fe

[F-statistic; 3.9619. Probability; 0.0332]

(UE cause federal expenditure to increase)

The significant uni causal association between UE in the upper level series and federal expenditure suggests that UE in large size (economic constrain) had cause government to intervene or to combat the UE. UE growth reduced with increasing growth of federal expenditure. The opposite UE growth or negative robust correlation suggests that government intervention had suppressed the UE growth and size of UE relative to GDP. This suggests that the rising federal expenditure growth result to long term federal budget deficit is partly due to a flourishing UE.

A negative interpretation of this uni causality relationship is that government could have not incurred this remedial expenditure in an official economy or if UE was at acceptable low level. In other words, the remedial expenditure would have

been “unnecessary expenses”, instead employed on subsidies and the welfare of citizens.

### 7.4.2 GDP per capita

UE in the nominal model correlates positively with GDP per capita, in level year and in lag of one year with elasticity between 1.8154% and 2.1541% provided that other controlled variables in the models are constant. However UE in the real model is insignificant to GDP per-capita at level year but positively significant in lag of one year. This suggests that UE is likely to increase with GDP nominal; due to CPI incremental. CPI which is the increase in price of goods and services that associate positively with GDP nominal is a cost to consumers. This would result to some erosion of cost of living as people would have to pay more for the same product. Past studies showed that UE could either associate positively or negatively with GDP. Priori explanations do not conclusively explain how UE growth is affected by GDP growth.

But naturally, under economic constrain people would seek opportunity of gaining “private benefit”, one way is to opt out of the official economy and enter UE to avoid regulation costs. However, this contention is a delayed effect, as supported by the insignificant correlation between UE and GDP real in level years. This positive correlation is consistent with positive association between UE and income level of individual participants, states wealth and GDP contribution of economic industries or business sectors.

### 7.4.3 Inflation rate

UE in the nominal model correlates positively with CPI in level year and in lag of one year. UE is most positively elastic to CPI compared to other hypothesized variables, with elasticity between 3.3076% growth and 4.2642%. It means that UE growth is between 3.3076 % and 4.2642% for every one percent growth of CPI, with other relevant variables in the model constant. As UE is most elastic to CPI and considering that CPI is the main deflation factor that converts the GDP nominal to GDP real per capita, association differences between UE with GDP nominal and GDP real is mainly due to CPI incremental

This positive elastic correlation conforms to the consensus view that the CPI is an "unlegislated tax" which is an adequate economic constrain that drive people to opt out of the official economy for underground activities. It is believed that CPI effect works in a chain mechanism such that money increases with CPI, resulting to low purchasing power and insufficient goods. Shortage supply of goods would initiate demand to increase, and eventually price of goods and services to increase.

### 7.4.4 Cash in circulation

Priori studies indicate that cash increases with CPI and associate with UE as it is the preferred medium of transactions. As for the "cash model" (Table 7.1) where UE is a function of cash, federal expenditure and electricity consumption, UE correlates positively with Malaysian currency in circulation both in nominal and real terms, with elasticity between 1.0182 and 0.9676 respectively. One possible reason to the significant cash coefficients after controlling for CPI and population

in an environment of various financial instruments (an economy of a cashless society), is a flourishing UE. Cash is often said the preferred medium of exchange in UE as it leaves minimal audit trails.

### 7.4.5 Electricity consumption

The model of the official economy as summarized in Table 7.14 in the Appendix of Chapter 7 showed no significant association with electricity consumption. One possible reason to this insignificant correlation is because country's electricity consumption constitutes only about 18% of the national physical consumption.

However, there is evidence of positive association between UE and the physical consumption (electricity in KWh), in all models constructed both in nominal and real terms. This indicates that consumption of electricity is a significant physical resource for economic production in UE.

Significant electricity consumption in UE ranging between 0.9104 and 1.2881 implies that subsidized supply of electricity had been used in an unrecorded economy. If its consumption is also underground (non compliant with electricity rules -"tabbed" or unpaid bills), it would mean that the country also loses revenue from prices of power supply.

### 7.4.6 Taxes

Over the years, generally direct tax reform has transformed to a lower tax burden but indirect tax reform transform to a higher tax burden. As tax transformation differs between direct and indirect taxes, tax revenue variable is used to proxy tax burden instead of tax structure. This is based on the believe that increase tax

revenue implies increase burden of tax payment, that drive people to escape tax for “private benefit”.

In nominal term, UE model in low level series showed positive correlation with the previous year tax burden with coefficients at 0.8266% and 0.5482% for direct tax and indirect tax respectively. It implies that for every one percent growth of tax paid in the previous year (as inferred by increase of tax revenue), controlled for CPI and federal expenditure. However, in real term, UE in low level is positively significant only to indirect taxes with coefficient of 0.4615%. It implies that controlling for CPI, the role of direct tax burden is no longer significant, but indirect tax burden remains.

The pro-cycle relationship implies that the amount of tax paid in the previous year partly drives people to evade taxes in the following year. The positive previous effect suggests that by nature people would want to compensate for the income lost to tax made in the previous year for “private benefit”. In other words, they would try to adjust future private profits and loss from “detection risks” to compensate for income lost made on tax liability in the previous year. However this inelastic positive correlation is no longer significant when UE is in the upper level, suggesting that other determinant variables are displacing tax burden in bad time.

7.4.7 Economic crises (dummy between pre and post 1997-98 financial crises)

The common consequences of economic crises are negative impacts on economic structures. UE in the upper level series but not in the moderate and lower series is positively significant to economic crises. The positive correlation with economic

crises implies that UE associate with macro structural changes of impact of economic crises. The insignificant association between UE in upper level series with tax burden instead significant association with economic crises; is consistent with significant association of UE in low series with tax burden but insignificant association with economic crises. The economic association shift according to size of UE implies that tax role is of less importance than "other macro structural changes" of economic crises in bad time.

The evidence of displacement association between tax burden and other variables is also consistent with shift of UE composition mix between irregular economy and illegal economy during good and bad time as discussed in Chapter Six. Illegal activities have displaced a portion of legal activities during recession times. Nevertheless, the positive macro structural influence is fairly modest with coefficients between 0.3464 and 0.5352, after controlling for federal expenditure, electricity consumption and cash in circulation. Indeed association is no longer significant after having controlled for CPI, as in real model. This suggests that economic crises influence is likely due to CPI incremental effect.

### 7.4.8 Other variables and overall interactions

The unemployment, bribery and crime rates variables are only significant in the uni-variate regression UE models, and with the insertion of AR factor (autoregressive factor). The significant insertion implies a short run relationship and there is no evidence of significant association between UE and these variables collectively. This implies that they are somewhat unreliable indicators of UE, likely due to crude average estimates of indexes.

Transactions of economic activities in the real world are potentially complex; the variables used can be highly correlated and integrated among themselves and with other numerous stimulating factors that are not considered in this study. Table 7.2 in the Appendix of Chapter 7 summarises the variable closed proximity (correlation matrix) that reflects potential multi collinear problem. There could also be confounding variables such as other government intervention and macro variables that induce or mask UE and the selected explanatory variables. The potentially complex interactions could form part of a vicious circle causing difficulties in assessing and quantifying the origins, an artifact multi collinear. Economist believe that multi collinear effect in regression model could generate spurious variable coefficient estimates. According to Bradsley (1994), good economic model is complex enough to be relevant, yet simple enough to be understood. Richard and Scurya (2001) indicate the difficulties of achieving the right balance between simplicity and complexity within the economic models, as the models generally failed to feed equilibrium. Hence it is justifiable that study showed no evidence of causal association between UE and most of the significant variables (except for federal expenditure).

To wrap up UE models constructed indicate that CPI is the most elastic variable followed by cash in circulation and electricity consumption. The government has intervened onto the issues of UE as indicated by the robust negative association with federal expenditure. The robust federal expenditure elasticity suggests that it is; a leading indicator (provides advance warning for a flourishing underground activities); a coincident indicator (reflecting current economic performance); and

a lagging indicator (signaling previous changes). The uni-causal relationship suggests that federal expenditure can be as index or ratio of inputs to UE models. However the positive correlation between UE and GDP per capita, CPI, Cash, M1 aggregates, electricity consumption and post 1998-99 economic crises, with no evidence of causal relationship suggests their "mere" correlation, ignoring the limitation of paired Granger tests.

As UE in the upper series model established the best fit model compared to UE in the lower and moderate series, UE in the upper series is proposed to be the appropriate size. The average size examined in 5 year period fluctuates between 8.69% (1980-84) and 26.94% (1995-99), which now hovers around 15% relative to GDP.

**Summary**

UE and explanatory variables revealed typical characteristics of time series data, exhibiting trend, non stationary characteristics and autocorrelation among successive observations. UE estimates were evaluated for any economic association in double regression model for appropriate and practical interpretation of coefficients; buoyancy, elasticity and growth.

UE correlate with macro variables as hypothesized, with correct signs and in long run association. In nominal term, UE in the upper series grew most with CPI, followed by GDP per capita, cash in circulation, electricity consumption, pre and post 1997-98 crises and pro-cycle to taxes. The most elastic "in immediate effect" to CPI in nominal term, and GDP per capita loses its positive role in real term

strongly suggest that CPI is a strong determinant of UE. Nevertheless, a "delayed effect" of one year with GDP per capita in real term implies that income level is also a significant determinant of UE, as it is more "profitable" to participate in UE. With regard to tax role, other economic macro changes are stronger determinants of UE in bad time. Government had intervened and successfully suppressed the growth of UE but the rising growth of federal expenditure has resulted to a long run federal deficit in the post 1997-98 crises.

The effects are reasonably explained such that; as GDP per capita increases the "complexity and invisibility of income" increase with UE opportunities; CPI as the "illegal tax" that create economic constrains are incentives of UE; cash in circulation is the preferred medium of exchange in UE transactions as it leaves minor audit trails; electricity consumption is the common physical resource in UE activities; and economic crises that result to economic macro changes are incentives of UE. As economic interaction can be potentially complex and people participate in the UE for many reasons, result indicate measures of output (indicators) rather than measures of input (determinants), except for federal expenditure could possibly prove otherwise.

The correct signs and persistent significant coefficients of variables as priori could reduce any skeptical views on estimation procedure and sample based estimates or unique estimates. Thus size of UE estimates that escape taxes are reasonably appropriate. As UE in the upper level series generate best fit models, it is recommended the appropriate size – UE that escape taxes fluctuate between 8.69% (in good time) and 26.94% in bad time.