The Hidden Economy and Tax-Evasion Prosecutions in New Zealand

David E. A. Giles

Department of Economics, University of Victoria

P.O. Box 3050, MS 8532, Victoria, B.C.

CANADA, V8W 3P5

April, 1996

***Proposed Running Head*:** *The hidden economy and tax prosecutions*

***Abstract*:** We consider non-stationary time-series data for the size of the hidden economy and for tax-related prosecutions in New Zealand. The two series are found to be cointegrated, and there is strong evidence of Granger causality from prosecutions to hidden activity. There is no significant evidence of reverse causality.

***Author Contact*:** Professor David E. A. Giles, Department of Economics, University of Victoria, P.O. Box 3050, MS8532, Victoria, B.C., CANADA, V8W 3P5

FAX: +1-604-721 6214 **Phone:** +1-604-721 8540

e-mail: DGILES@EC1.ECON.UVIC.CA

I. INTRODUCTION

Measuring the size of the hidden economy is an important policy problem which has attracted considerable international attention. Following Feige (1979,1982) and others, several methods have been used to construct such measures for different countries. Frey and Weck-Hanneman (1984) and Aigner *et al*. (1988) view the size of the hidden economy as a "latent variable" and use the MIMIC ("Multiple Indicators, Multiple Causes") model of Zellner (1970), Jöreskog and Goldberger (1975), and others, for their analysis. The MIMIC model belongs to the LISREL ("Linear Interdependent Structural Relationships") class of models (*e.g.*, Jöreskog and Sörbom (1993)).

Giles (1995) uses[1] MIMIC modelling to estimate the size of the hidden economy in New Zealand, and finds it to be 8.8% of measured GDP, on average, over the period 1968 to 1994. The time-series for the hidden economy that he generates has a cyclical pattern which is similar to, but more pronounced than, that of the measured economy. The issue of Granger (1969) causality between measured and hidden output in that country is analyzed by Giles (1996), who finds that while there is clear evidence of causality from the former variable to the latter, there is only mild evidence of causality in the reverse direction.

In this paper we focus on the issue of tax evasion, as one component of hidden economic activity. In particular, we use the time-series for the New Zealand hidden economy constructed by Giles (1995) to test for the presence of a causal relationship between unmeasured economic output, and prosecutions associated with tax evasion in that country. It is interesting to know whether or not the activities of the agency charged with implementing and policing the taxation laws are having an impact on the degree of illicit activity. Equally, it is interesting to know if that agency is aware of, and responding to, the extent of non-compliance.

II. DATA FEATURES

We have annual data for the size of the New Zealand hidden economy (HE), in real 1982/83 $Millions, for the period 1968 to 1994. These data are created from ratio data generated by[2] MIMIC Model 2 of Giles (1995), and measured real GDP. The *relative* size of unrecorded economic activity increased from 6.8% of measured real GDP in 1968 to a peak of 11.3% in 1987. It then fell to 8.7%

of measured real GDP in 1992 before increasing to 11.3% in 1994. Other features of this series are noted by Giles (1996).

We also have annual data for 1967 to 1993 for the total number of *prosecutions* (PROS) relating to offenses against the Inland Revenue Acts. This information, in disaggregated form, is a matter of public record and is published regularly in *The New Zealand Gazette*. At this stage we have (annual) data on the associated *convictions* only for the period 1967 to 1986, so this information has not yet been considered in detail. However, it is worth noting that over that period the simple correlation between the numbers of prosecutions and convictions was 99.97%, and the (geometric) average of the conviction *rate* over this period was 96.3%.

The PROS series exhibits a marked structural break in both its level and trend in the early 1980's. From the disaggregated data it is clear that this is due primarily to changes in the number of prosecutions in the "Failure to Furnish Returns" category of non-compliance. Other noteworthy underlying characteristics of these data are that the number of prosecutions in the "Wilfully False Returns" and "Failure to Deduct PAYE" categories both approximately quadrupled in 1983, and both approximately halved in 1987 (in each relative to the numbers in the previous year). Prosecutions relating the Goods and Services Tax (GST) are recorded[3] only from 1988, with 205 cases in that year. Interestingly, the corresponding numbers of such prosecutions in each successive year up to and including 1993 were 1,036, 526, 259, 256 and 443 respectively.

The simple correlation between HE and PROS over the (common) sample period of 1968 to 1993 is 69.5%. The positive sign for this figure is interesting in itself - at least superficially it suggests that the *raw number* of prosecutions, taken in isolation from other effects, does not act *concurrently* as a deterrent to unrecorded activity. This is a matter that warrants closer examination, perhaps in the context of a model of a structural model of tax evasion[4] which takes account, *inter alia*, of a suitable measure of the "probability of detection".

III. STATIONARITY ISSUES

Giles (1995, 1996) paid particular attention to the non-stationarity of his time-series data, and this feature of the econometric analysis is also emphasised here. Using the SHAZAM (1993) package, we have used the "augmented" Dickey-Fuller (ADF) test (*e.g.*, Dickey and Fuller (1979,1981); Said and Dickey (1984)), to test the data for stationarity. Dods and Giles (1995) show that (for samples as small as ours) the default method of choosing the augmentation level, p, in the SHAZAM package performs well in terms of minimizing the size-distortion associated with a pre-test selection of this level. We have followed the strategy suggested by Dolado *et al.* (1990) (and also used by Giles *et al.* (1992), Giles (1994, 1996), and Mandeno and Giles (1995)) to determine whether or not drift and/or trend terms should be included in the Dickey-Fuller regressions. To test that the series x_t is integrated of order 1 (*i.e.*, is I(1)), against the alternative that x_t is integrated of order zero (*i.e.*, is I(0), or stationary) the level of augmentation, p, is determined as above, in the context of the following ADF regression model:

$$\Delta \mathbf{x_t} = \alpha + \beta \mathbf{t} + \gamma \mathbf{x_{t-1}} + \theta_1 \Delta \mathbf{x_{t-1}} + \ldots\ldots + \theta_p \Delta \mathbf{x_{t-p}} + \epsilon_t \qquad (1)$$

We then test $H_o : \gamma = 0$ *vs.* $H_A : \gamma < 0$ using the Dickey-Fuller "t" test (denoted "t_{dt}" below) and MacKinnon's (1991) critical values. If H_0 is rejected, we conclude that x_t is stationary, otherwise we test $H_0 : \beta = \gamma = 0$, using the "F-test" (denoted "F_{ut}" below) of Dickey and Fuller (1981). Rejection, leads us to conclude that x_t is I(1), otherwise we remove the trend from the ADF regression and test $H_0 : \gamma = 0$ *vs.* $H_A : \gamma < 0$. The ADF "t-statistic" is denoted "t_d". If we cannot reject H_0, we test $H_0 : \alpha = \gamma = 0$ using the "F-test" (denoted "F_{ud}" below) of Dickey and Fuller (1981). Rejection suggests

that x_t is I(1), otherwise we remove the drift, re-estimate, and test $H_0 : \gamma = 0$ *vs.* $H_A : \gamma < 0$. This "t-statistic" is denoted "t" in Table 1. A rejection of H_0 suggests that x_t is I(0), or stationary[5], while failing to reject H_0 suggests that x_t is I(1). Table 1 shows the results of testing the order of integration of the HE and PROS series. To allow for the structural break in the latter series we have also followed Perron's (1989) modification of the ADF test according to his "Case (C)".

IV. GRANGER CAUSALITY AND COINTEGRATION

We have considered two simple empirical issues arising from the above discussion. First, it is interesting to ask, "Is the size of the hidden economy cointegrated with the number of tax-offense prosecutions?". An affirmative reply, implies there is a long-run equilibrium relationship between these two variables which ensures that they will not continue to drift apart indefinitely in the event of an exogenous "shock" to one or both of them. It also implies that there is either uni-directional or bi-directional Granger causality between HE and PROS. So, the second issue that we have addressed is the nature and direction of such causality. (It should be noted that this is a *potential* issue in any case, regardless of whether HE and PROS are cointegrated or not. Cointegration implies (Granger) causality, but not *vice versa*.)

As both HE and PROS are I(1), we have tested for cointegration, and the results appear in Table 2. As well as the ADF test for cointegration we have also used the method of Phillips and Perron (1998) and Phillips and Ouliaris (1990) (denoted PPO). To allow for the structural break in the PROS series we have conducted a modified ADF (MADF) test for cointegration by including, in the "drift and trend" cointegrating regressions[6], the trend and drift shift-dummies that are used in Perron's (1989) modification of the ADF test for a unit root, as in Table 1. The evidence in Table 2 is mixed, depending on the normalization of the cointegrating regressions. We conclude in favour of cointegration, by focussing on the regressions with the largest R^2 values (*e.g.*, Banerjee *et al*. (1986)), and by taking special note of the MADF results. This, in turn, suggests there is Granger causality, one way or the other, between the size of the hidden economy and the number of prosecutions for tax-related offenses.

Tables 3 and 4 show the results relating to the specification of a simple VAR model between HE and PROS. The results of Toda and Yamamoto (1995) justify modelling in the *levels* of the data (despite their non-stationarity), as well as choosing the lag specification on the basis of a sequence of Wald tests. The latter have been performed sequentially for the significance[7] of the maximum lag of each variable in each equation (four restrictions). The HE equation includes a linear trend, and the PROS equation includes the intercept dummy for the structural break. Their inclusion is also based on Wald test for zero restrictions, which Toda and Yamamoto (1995) show to be asymptotically valid. The results in Table 3 are based on joint estimation of the two-equation system, and we conclude that five lags of each variable should appear in each equation[8]. Following Toda and Yamamoto (1995) we then add one[9] extra lag of each variable to each equation, but *do not* include the coefficients of these extra lags in the subsequent causality tests. In this way, the latter Wald test statistics have their usual asymptotic chi-square distribution.

In the joint estimation of the final system, the Breusch-Pagan LM Test and the Likelihood Ratio Test statistics took the values of 0.31 and 0.93 respectively (and are each $\chi^2(1)$). So, we cannot reject the hypothesis of a diagonal error covariance matrix, and OLS estimation was adopted for the rest of the analysis. We see from the Wald test statistics in Table 4 that we strongly *reject* the hypothesis of no causality form PROS to HE, but we *cannot reject* this hypothesis in the reverse direction. Although a very simple VAR model has been estimated, the results of the diagnostic tests in Table 4 are very encouraging. JB is the Jarque-Bera Normality test; LM1-5 is the Lagrange Multiplier test for serial independence against AR(5) or MA(5) alternatives; and F11 and F22 are asymptotic "FRESET" (Fourier-RESET) tests based on 1 and 2 sine and cosine terms.[10]

Focussing on the crucial Wald test for Granger causality, we have conducted a small bootstrap simulation experiment (with 10,000 replications) to see how reliable it is in a sample of this size and with a model of this particular type. The exact finite-sample p-values for the causality tests are also shown in Table 4 and we see that they support our earlier conclusions.

V. CONCLUSIONS

Our finding that the size of the hidden economy and the number of taxation prosecutions are cointegrated implies that there is no long-run tendency for these two series to diverge if one or other of them receives an exogenous shock. This seems very plausible in the New Zealand context. Second, we have found strong evidence that the number of prosecutions for tax-related offenses Granger-causes the size of the hidden economy. (It is important to note that the modelling procedure which generated the latter data *did not* incorporate any prosecution-related data.) On the other hand, there is no evidence of reverse causality from the size of the hidden economy to the number of prosecutions. This suggests that the auditing and other compliance-related activities of the New Zealand Inland Revenue Department are impacting on the underground economy in that country. It also suggests that the government agents are pro-active, rather than reactive, at least in a long-run average sense. These conclusions open up some interesting areas of further investigation, especially in relation to a disaggregation of hidden output into "avoidance" and "evasion" components. This is the subject of research in progress.

ACKNOWLEDGEMENTS

The author is particularly grateful to Patrick Caragata of the New Zealand Inland Revenue Department for initiating and supporting this research. Thanks are also due to Linda DeBenedictis, Johannah Dods, Ed Feige, Judith Giles, Ewen McCann, Knox Lovell, Gerald Scully, John Small and Ken White for their comments and/or data assistance.

REFERENCES

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

Allingham, M. G. and A. Sandmo (1972) Income tax evasion: a theoretical analysis, *Journal of Public Economics*, **1**, 323-338.

Banerjee, A., J. J. Dolado, D. F. Hendry and G. W. Smith (1986) Exploring equilibrium relationships in econometrics through static models: some Monte Carlo evidence, *Oxford Bulletin of Economics and Statistics*, **48**, 253-277.

Crane, S. E. and F. Nourzad (1986) Inflation and tax evasion, *Review of Economics and Statistics*, **68**, 217-223.

DeBenedictis, L. F. and D. E. A. Giles (1996) Testing the specification of regression models using Fourier approximations: the FRESET test, mimeo., Department of Economics, University of Victoria.

Dickey, D. A. and W. A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, **74**, 427-431.

Dickey, D. A. and W. A. Fuller (1981) Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, **49**, 1057-1072.

Dickey, D. A. and S. G. Pantula (1987) Determining the order of differencing in autoregressive processes, *Journal of Business and Economic Statistics*, **15**, 455-461.

Dods, J. L. and D. E. A. Giles (1995) Alternative strategies for 'augmenting' the Dickey-Fuller test: size-robustness in the face of pre-testing, *Journal of Statistical Computation and Simulation*, **53**, 243-258.

Dolado, J. J., T. Jenkinson and S. Sosvilla-Rivero (1990) Cointegration and unit roots, *Journal of Economic Surveys*, **4**, 249-273.

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

Feige, E. L. (1982) A new perspective on macroeconomic phenomena: the theory and measurement of the unobserved economy in the United States: causes, consequences and implications, in *International Burden of Government* (Ed.) M. Walker, Vancouver, pp. 112-136.

Frey, B. S. and H. Weck-Hannemann (1984) The hidden economy as an 'unobserved' variable, *European Economic Review*, **26**, 33-53.

Giles, D. E. A. (1995) Measuring the size of the hidden economy and the tax-gap in New Zealand: an econometric analysis, Working Paper 5, Inland Revenue Department, Wellington.

Giles, D. E. A. (1996) Causality between the measured and underground economies in New Zealand, *Applied Economics Letters*, forthcoming.

Giles, D. E. A., J. A. Giles and E. McCann (1992) Causality, unit roots and export-led growth: the New Zealand experience, *Journal of International Trade and Economic Development*, **1**, 195-218.

Giles, J. A. (1994) Another look at the evidence on foreign aid led economic growth, *Applied Economics Letters*, **1**, 194-199.

Granger, C. W. J. (1969) Investigating causal relations by econometric models and cross spectral methods, *Econometrica*, **37**, 434-438.

Jöreskog, K. and A. S. Goldberger (1975) Estimation of a model with multiple indicators and multiple causes of a single latent variable, *Journal of the American Statistical Association*, **70**, 631-639.

Jöreskog, K. and D. Sörbom (1993) *LISREL 8: Structural Equation Modeling With the SIMPLIS Command Language*, Scientific Software International, Chicago.

1n9 9 1K) cCarMtical values for co-integration tests, in *Long-Run Economic Relationships* (Eds.) R. F. Engle and C. W. J. Granger, Oxford University Press, Oxford, pp. 267-276.

Mandeno, R. J. and D. E. A. Giles (1995) The expectations theory of the term structure : a cointegration/causality analysis of U.S. interest rate data, *Applied Financial Economics*, **5**, 273-283.

Mizon, G. E. (1977) Inferential procedures in non-linear models: an application in a U.K. industrial cross section study of factor substitution and returns to scale, *Econometrica*, **45**, 1221-1242.

Perron, P. (1989) The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, **57**, 1361-1401.

3 .l alnd S .P Ouliaris (1990) Asymptotic properties of residual based te cointegration, *Econometrica*, **58**, 165-193.

Phillips, P. C. B. and P. Perron (1988) Testing for a unit root in time series regression, *Biometrika*, **75**, 335-346.

Said, S. E. and D. A. Dickey (1984) Testing for unit roots in autoregressive-moving average models of unknown order, *Biometrika*, **71**, 599-607.

SHAZAM (1993) *SHAZAM Econometrics Computer Program: User's Reference Manual, Version 7.0*, McGraw-Hill, New York.

Toda, H. Y. and T. Yamamoto (1995) Statistical inference in vector autoregressions with possibly integrated processes, *Journal of Econometrics*, **66**, 225-250.

Zellner, A. (1970) Estimation of regression relationships containing unobservable variables, *International Economic Review* , **11**, 441-454.

FOOTNOTES

1. The work reported by Giles (1995,1996) results from a major on-going research program into various aspects of the New Zealand tax system which is being undertaken by the Policy Advice Division of the New Zealand Inland Revenue Department.

2. While Giles (1995) reports several MIMIC models for the relative size of the hidden economy, the time-series predictions for that variable are quite *insensitive* to the choice of model specification. The quality of the model specifications is also robust to the choice between annual and quarterly data in their estimation.

3. This (consumption) tax was introduced in New Zealand in October 1986.

4. For example, see Allingham and Sandmo (1972) and Crane and Nourzad (1986).

5. We actually follow Dickey and Pantula (1987) and test I(3) against I(2), using the doubly first-differenced data; I(2) against I(1), using the first-differenced data; and then I(1) against I(0). The results of testing I(3) against I(2) are omitted to conserve space.

6. As these dummies are exogenous, the MADF statistic will be *asymptotically* equivalent to the ADF statistic in the "drift and trend" case. The "drift/no trend" cointegrating regression is *not* modified to allow for the break in PROS, as this break affects *both* the level and the trend.

7. Sequential testing involves a sequence of st
independence of the successive tests allows the actual significance level to be controlled at each step (Mizon (1977). This also allows us to compute the "effective" p-values at each step.

8. The choice between five and six lags is moot. Our choice preserves degrees of freedom and the associated diagnostic test results in Table 4 are "clean". Details of the estimated models are available on request.

9. The one extra lag corresponds to our finding that each series is I(1).

10. DeBenedictis and Giles (1996) propose this test and show that it has excellent power in cases where the conventional RESET test has negligible power. The asymptotic Chi Square distribution for the FRESET test follows from the results of Toda and Yamamoto (1995). A shortage of degrees of freedom precludes implementing the test with more terms in our case.

Table 1. *ADF and Perron tests for unit roots* [a]

Var.	Test	T	p	t_{dt}	F_{ut}	t_d	F_{ud}	t	Outcome
					I(2) vs. I(1)				
HE									
	ADF	24	1	-2.94	4.72	-3.13	n.a.	n.a.	Reject I(2)
PROS									
	ADF	24	1	-3.24	n.a.	n.a.	n.a.	n.a.	Reject I(2)
	Perron	24	1	-5.20	n.a.	n.a.	n.a.	n.a.	Reject I(2)
					I(1) vs. I(0)				
HE									
	ADF	26	0	-2.78	3.99	-0.68	1.72	1.52	I(1)
PROS									
	ADF	26	0	-1.79	1.65	-1.61	1.33	-1.11	I(1)
	Perron	25	1	n.a.	n.a.	n.a.	n.a.	-3.76	I(1)

[a] T = sample size. The other notation is defined in the text. For the Perron Test, $\lambda = 0.6$; 5% (10%) asymptotic critical values are -4.24 (-3.95) respectively. See Dickey and Fuller (1981) and MacKinnon (1991) for the other critical values.

Table 2. *ADF, PPO and MADF cointegration tests* [a]

Dependent Variable	Drift, No Trend				Drift &Trend			
	T	p	t	R^2	T	p	t	R^2
				ADF [b]				
HE	26	0	-3.04	0.48	26	0	-3.76	0.85
PROS	26	0	-2.36	0.48	26	0	-2.58	0.49
				PPO [b]				
HE	25	1	-2.38	0.48	25	1	-3.77	0.85
PROS	25	1	-2.45	0.48	25	1	-2.64	0.49
			(-3.58)					(-4.17)
			[-3.21]				[-3.78]	
				MADF [c]				
HE	n.a.	n.a.	n.a.	n.a.	26	0	-3.92	0.88
PROS	n.a.	n.a.	n.a.	n.a.	26	0	-4.94	0.75
							(-3.78)	
							[-3.50]	

a "Drift, No Trend" and "Drift & Trend" refer to the cointegrating regression. In each case the Dickey-Fuller regression for the residuals has no drift and no trend.

[b] *Exact* 5% and 10% critical values (from MacKinnon (1991)) appear in parentheses and brackets, respectively.

[c] *Asymptotic* 5% and 10% critical values appear in parentheses and brackets, respectively.

Table 3**.** *Wald tests for lag specification in the VAR* [a]

	Maximum Number of Lags [b]				
	6	5	4	3	2
Wald (χ^2_4)					
	8.64	10.35	5.55	5.08	3.56
	(0.07)	(0.03)	(0.24)	(0.28)	(0.47)
	[0.07]	[0.09]	[0.31]	[0.51]	[0.74]

[a] Nominal asymptotic p-values appear in parentheses; "effective" asymptotic p-values appear in brackets.

[b] The lags apply to both variables in both equations.

Table 4. *Wald causality tests and VAR diagnostics* [a]

Equation	Wald	JB	LM1-5	F11	F22
	(χ^2_5)	(χ^2_2)	(χ^2_5)	(χ^2_2)	(χ^2_4)

HE	13.56	1.21	5.58	3.44	3.30
	(0.02)	(0.55)	(0.35)	(0.18)	(0.51)
	[0.05]				
PROS	1.25	3.08	5.50	1.39	5.37
	(0.94)	(0.21)	(0.36)	(0.50)	(0.25)
	[0.93]				

[a] Asymptotic p-values appear in parentheses. Exact bootstrapped p-values for the Wald tests appear in brackets.