EXOGENEITY AND MEASURES OF PERSISTENCE

Guglielmo Maria Caporale*
Nikitas Pittis**

RESUMEN

El artículo argumenta que el hecho que la evidencia empírica sobre persistencia es mixto no es muy sorprendente, ya que la teoría económica esta limitada al tener que especificar el modelo estadístico. Este punto se ilustra de dos maneras. Primero, resaltamos el hecho que el concepto de persistencia depende del modelo, i.e. es una función del modelo adoptado a partir de la teoría económica. Segundo, se analiza el tema relacionado con la definición de un conjunto de parámetros de interés. En particular, consideramos un caso bivariado simple. Dada la exogeneidad débil de los regresores, los parámetros del único vector de cointegración pueden ser estimados equivalentemente dentro de un sistema completo de ecuaciones o con un modelo uniecuacional. Por el contrario, si la persistencia está siendo medida, la exogeneidad débil de los regresores no se mantiene, ya que los parámetros de interés no pueden ser escritos solo como una función de aquellos del modelo condicional y el concepto de exogeneidad (ya no débil) del modelo se convierte en mas relevante. Una vez mas, la teoría económica puede ser vista como esencial en el ejercicio de especificación del modelo.

ABSTRACT

This paper argues that the fact that the empirical evidence on persistence is mixed is not very surprising, as economic theory is bound to be drawn upon in order to specify the statistical model. This is illustrated in two ways. Firstly, we highlight the fact that the concept of persistence is model dependent, i.e. it is a function of the maintained model adopted on the basis of economic theory. Secondly, we analyse the related issue of the definition of a set of parameters of interest. In particular, consider a simple bivariate case. Given weak exogeneity of the regressors, the parameters of the unique cointegrating vector can equivalently be estimated within a full system or a single equation framework. On the contrary, if persistence is being measured, weak exogeneity of the regressor does not hold any longer, as the parameters of interest cannot be written as a function of those of the conditional model only, and the concept of model (rather than weak) exogeneity becomes more relevant. Once again, economic theory can be seen to play an essential role in model specification.

Keywords: Persistence, Weak Exogeneity, Autocorrelation Function, Conditioning Information Set, Dynamic Models, Cointegration
JEL classification: C22, C32

---

* Centre for Monetary and Financial Economics South Bank University London
** Department of Banking and Financial Management University of Piraeus
We are grateful to George Alogoskoufis, Stephen Hall, Michael Magdalinos and Aris Spanos for useful comments and suggestions. The usual disclaimer applies.
Corresponding author: Professor Guglielmo Maria Caporale, South Bank University, Business School, 103 Borough Road, London SE1 OAA, UK. Tel +44(0)20 7815 7012. Fax +44 (0)20 7815 8226
E-mail: g.m.caporale@sbu.ac.uk

I. INTRODUCTION

The issue of persistence in macroeconomic time series has been extensively investigated in the last decade because of its implications for macroeconomic theory and policy. For example, in their seminal paper Nelson and Plosser (1982) argued that the presence of unit roots meant that shocks were persistent, and hence that the data were consistent with Real Business Cycle (RBC) models, in which most shocks to GNP are technology shocks. Similarly, Campbell and Mankiw (1987a, 1987b) suggested that an ARMA(2,2) model provided the best description of the data for US real GDP, which is therefore generated by a difference-stationary (DS) (or unit root) process, and also that the long-run response of US GDP to a unit shock, given by the cumulative response function A(1), is greater than 1, which implies that there is no trend-reversion.[1]

Various statistics have been proposed to measure persistence. For instance, Cochrane (1988) argued that, because any time series with unit root can be decomposed into a stationary series and a random walk, and the latter can have arbitrarily small variance, persistence should be defined as the ratio of the variance of the change in the random walk component to the variance of the actual change. Lo (1991) introduced a modified rescaled range statistic, which converges to a well-defined random variable under the null hypothesis of short-term dependence, and can distinguish between short-run and long-run dependence.

Cochrane and Sbordone (1988) provided a measure of persistence for GDP and stock prices which makes use of multivariate information. Cochrane (1991) showed that the persistence of univariate and multivariate prediction error shocks can be very different.[2] Lupi (1993) also suggested that measures of persistence are not invariant to the information set, and that in a general probabilistic framework they are inadequate to capture persistence in terms of non-mixing properties. Evans and Reichlin (1994) demonstrated that the Beveridge-Nelson (BN) decomposition into trend and cycle is non-increasing in the number of conditioning variables, and it is strictly decreasing if additional conditioning variables Granger-cause the variable of interest. Cochrane (1994) illustrated the empirical importance of this insight: the addition of Granger causal variables radically alters measures of transitory components in US GDP and stock prices.

This paper argues that as economic theory is bound to be relied upon in order to specify the statistical model used to measure persistence, the inconclusiveness of the available empirical evidence is hardly striking. This is illustrated in two ways. Firstly, we higlight the fact that a concept such as the degree of persistence of a time series is

---

1 Christiano and Eichenbaum (1989), however, pointed out that this inference was very sensitive to the choice of ARMA specification from a set of models which had equally good fits.

2 Lippi and Reichlin (1992) pointed out that another measure of persistence often used in empirical studies (see, e.g., Clark (1987), or Watson (1986)), which is based on the standard unobserved components models (UCARIMA) developed in Beveridge and Nelson (1981), is necessarily less than one as a mathematical consequence of the structure of these models.

model dependent. For instance, suppose that the *maintained model* is a VAR – persistence can be shown to be a function of the eigenvalues of the system. By contrast, if economic theory suggests that some variables are *model exogenous*, then the maintained model is a linear regression, and a valid measure of persistence is given by the autoregressive coefficient.

Secondly, we show that in the case of persistence the appropriate concept to use is *model* (rather than weak) *exogeneity*, with economic theory again playing an essential role in the specification of the model. This is related to the notion of *parameters of interest*. More specifically, consider a simple bivariate case, and assume that the researcher is first focusing on the coefficients of the unique cointegrating vector between the variables y_t and x_t, ant then on the degree of persistence of y_t. Under the assumption that x_t is weakly exogenous with respect to the cointegrating vector, asymptotically equivalent estimates of the cointegrating parameters can be obtained either from the full system of from the conditional model. However, if persistence is being measured, x_t cannot be weakly exogenous, as the parameters of interest cannot be written as a function of those of the conditional model only. Therefore, the full model has to be employed. At this stage, though, economic theory can be used, not only to select the variables to be included in the VAR, but also to define a *structure*, thereby obtaining a new maintained model, in the context of which the appropriate measure of persistence will be different.

The layout of the paper is the following. Section II considers persistence estimates in a cointegrated system. Section III emphasizes the role economic theory plays in model specification by deciding on the exogeneity status of the variables of interest. It is shown that model exogeneity and the definition of the set of parameters of interest affect crucially the inference on persistence. Some conclusions are offered in section IV.

II. PERSISTENCE ESTIMATES IN A COINTEGRATED SYSTEM

We define persistence as the effect of a 1 percent innovation on the long-run level of a series, say y_t - it can therefore be seen as the memory of y_t, which is the rate at which the autocovariance (or, alternatively, the autocorrelation) function decays to zero (see Priestley (1981)). Intuitively, this is because the memory of a series indicates the rate at which a dynamic system returns to its initial stage, which could be any steady state after being perturbed by a shock.[3] If the memory of the process gradually dies out, then persistence is small and eventually becomes zero, whereas in the case of constant memory, i.e. when the system never returns to its initial state after a shock, persistence is infinite. In other words, the faster the rate at which the autocorrelations

[3] In the literature, persistence is normally defined as the value towards which the impulse response function converges in the case of I(1) variables, and as the area under the impulse response function which would be infinity if the variable was non-stationary) for an I(0) series. In both cases, this corresponds to the sum of the coefficients of the moving average representation of the process. Our definition is basically equivalent.

1. Caporale.p65 3 16/02/03, 03:54 p.m.

function vanishes, the smaller the degree of persistence. Let us now consider the following bivariate VAR(1) representation of the DGP of y_t and x_t, which is assumed to be a Gaussian-Markov process:[4]

$$Z_t = AZ_{t-1} + E_t \qquad E_t \sim NIID(0,\Omega) \tag{1}$$

$$A=[a_{ij}], \qquad \Omega=[\omega_{ij}], \qquad i,j=1,2$$

We can then show that the vector autocovariance matrices takes form:[5]

$$C(1) = C(0)A'^{1} \tag{2}$$

Where

$$C(1) = \begin{bmatrix} c_{11}(1) & c_{12}(1) \\ c_{21}(1) & c_{22}(1) \end{bmatrix} \tag{3}$$

for 1= $\pm 1, \pm 2, ...$

Also, the corresponding vector autocorrelation matrix at lag 1 is denoted by

$$\rho(1) = V^{-1/2}C(L)V^{1/2}\begin{bmatrix} \rho_{11}(1) & \rho_{12}(1) \\ \rho_{21}(1) & \rho_{22}(1) \end{bmatrix} \tag{4}$$

where $V^{-1/2} = Diag\left(c_{11}(0)^{-1/2}, c_{22}(0)^{1/2}\right)$

Therefore, the vector autocorrelation matrix for the BVAR(1) has the form:

$$\rho(1) = V^{-1/2}C(L)V^{1/2} = \rho(0)[V^{-1/2}A'^{1}(L)V^{1/2}] \tag{5}$$

---

4 Note that even if normality in the errors is not explicitly assumed, the linear conditional mean and the homoscedastic conditional variance brings the whole framework very close to normality. This is due to the characterization result according to which if both conditional means $E(y_y \mid x_t)$ and $E(x_y \mid y_t)$ are linear in x_t and y_t respectively, and only one conditional variance $Var(y_y \mid x_t)$ is homoscedastic, then the joint distribution $f(y_t, x_t)$ is normal (see Spanos (1995a)).

5 Note that all the results derived below are also valid in a more general multivariate framework, where the vector Z_t in the stochastic process $Z_t = \left[Z_t, t \in T\right]$ is equal to $Z_t = (y_t, X_t)$, and $X_t = (X_{1t}, X_{2t}, \ldots, X_{kt})$.

1. Caporale.p65 4 16/02/03, 03:54 p.m.

now the VAR(1) coefficient matrix A' can be expressed in the *Jordan canonical form* as $A' = P\Lambda P^{-1}$, where P is a nonsingular matrix and Λ is a special upper triangular matrix that has the eigenvalues λ_1 and λ_2 of A' on its diagonal and (possibly) one in the position just above the diagonal (see Luktepohl (1991)). Therefore we have:

$$A'^{1} = \left(P\Lambda P^{-1}\right)^{1} = P\Lambda^{1}P^{-1}$$

which means that:

$$\rho(1) = \rho(0)[V^{-1/2}PL'P^{-1}(L)V^{1/2}] \tag{7}$$

It can be seen that the behavior of the vector autocorrelation function $\rho(1)$ depends on the absolute value of the two eigenvalues λ_1 and λ_2. If both λ_1 and λ_2 are less than one, then we are dealing with a stable VAR in the levels and the autocorrelation functions for both y_t and x_t will vanish as $1 \to \infty$, since Λ^1 tends to a matrix with zeroes in its diagonal, exhibiting a mixture of decaying exponential and damping sinusoidal behavior depending on the nature (real and / or complex conjugate values) of the eigenvalues of A.

In the case of a cointegrated VAR, however, one eigenvalue is equal to one and the other is less than one. Therefore, Λ^1 does not tend to a matrix with zeroes on its diagonal, which in turn implies that both y_t and x_t are characterized by an infinite degree of persistence.

This can be seen more clearly by examining the nature of the univariate models for y_t and x_t from the VAR(1) process. Reinsel (1991) shows that in the case of a K-1dimensional VAR(1) process, the individual series will follow univariate ARMA(k,k-1) models, where k and k-1 are the maximum orders for the individual ARMA models. In the case of the bivariate VAR(1) considered here, the individual series y_t and x_t will follow, at most, ARMA(2,1) models, with the modulus of one of the two foots of the characteristic polynomial of the autoregressive part being equal to one. This root will dominate the behavior of the autocorrelation function, thus resulting in an infinite degree of persistence.

III. THE IMPLICATIONS OF WEAK EXOGENEITY

In this section we examine the implications of conditioning on a subset of variables, which are treated as weakly exogenous with respect to a particular set of parameters of interest (see Engle et al (1983). We shall see how defining clearly the parameters of interest is of paramount importance for valid inference in this context. Initially, we assume that we are dealing with two *I(1)* series, say y_t and x_t, which exhibit cointegration. The parameters of interest are taken to be the elements of the unique cointegrating vector $b = [b_y, b_x]'$ between y_t and x_t. These parameters can be

1. Caporale.p65 5 16/02/03, 03:54 p.m.

estimated in the context of a system error correction formulation as described in Johansen (1988, 1991). Alternatively, by assuming that x_t is weakly exogenous for **b**, asymptotically efficient estimates of **b** can be obtained in the context of the following single equation:

$$y_t = \phi y_{t-1} + \delta_0 x_t + \delta_1 x_{t-1} + \mu_t \qquad \mu_t \sim NIID(0, \sigma_\mu^2) \tag{8}$$

where

$$\phi = \left[a_{11} - a_{21}\omega_{12} / \omega_{22}\right] \tag{8a}$$

$$\delta_0 = \left[\omega_{12} / \omega_{22}\right] \tag{8b}$$

$$\delta_1 = \left[a_{12} - a_{22}\omega_{12} / \omega_{22}\right] \tag{8c}$$

The relationships (8a) to (8c) are obtained by making use of known properties of the bivariate normal distribution.

Cointegration implies the existence of a single error correction representation for y_t, and hence that the coefficient ϕ on the lagged dependent variable is less than one, i.e. ϕ<1. The standardized cointegration vector is therefore equal to $\bar{b} = \left[1, \bar{b}_x\right]$' where $\bar{b}_x = (\delta_0 + \delta_1)/(1-\phi)$. Asymptotically, there will be no difference between drawing inference on the cointegrationg vector in the context of the full system or the conditional model (8).

Now assume that the investigator decides to measure the persistence of y_t instead. This implies a change in the definition of the parameters of interest. Assuming that the true data generation process is the bivariate VAR(1) model (1), then the parameters of interest, say ν, are the two eigenvalues of the matrix A, i.e. $\nu = \left[\lambda_1, \lambda_2\right]$'. For this new choice of parameters of interest x_t is no longer weakly exogenous, since these parameters cannot be expressed as a function of the parameters of the conditional model (8). Therefore, any inference on persistence based on the conditional model (8) will be misleading. This is often ignored in empirical applications: once a single equation model such as (8) has been formulated, where x_t has been assigned a weak exogeneity status with respect to the long –run parameters, then inference on persistence is usually based on the coefficient ϕ of the lagged dependent variable. The shift of interest from the cointegrating vector to the eigenvalues of the system is overlooked, and the fact that ϕ is less than one will erroneously lead to the conclusion that persistence is small.

On the other hand, if economic theory suggest unidirectional causality running from x_t to y_t, then the true data generation process (DGP) is likely to be approximated by the following model:

1. Caporale.p65 6 16/02/03, 03:54 p.m.

$$y_t = \phi' y_{t-1} + \delta_0' x_t + \delta_1' x_{t-1} + u_{1t} \tag{9a}$$

$$x_t = x_{t-1} + u_{2t} \tag{9b}$$

with $u_t = [u_{1t}, u_{2t}]' \sim NIID[0, \Sigma]$ with $\Sigma = [\sigma_{ij}], \quad i, j = 1,2 \;\; and \;\; \phi < 1$

By substituting (9b) into (9a) and exploiting the properties of the bivariate normal distribution we have:

$$y_t = \phi' y_{t-1} + (\delta_0 + (\sigma_{12} / \sigma_{22}))x_t + (\delta_1 - (\sigma_{12} / \sigma_{22}))x_{t-1} + \varepsilon_t \tag{10}$$

Here it is natural to define ϕ' as the parameter of interest, and to adopt it as a measure of persistence. Therefore, in the context of model (10) the conclusion that the degree of persistence of y_t is small is correct.

The above discussion has two important implications for empirical estimate of persistence:

1) The researcher should be clear at the outset about the parameters of interest –a set of variables may be weakly exogenous with respect to a particular set of parameters of interest but not with respect to another. In the latter case, inference has to be drawn in the context of the full model.

2) The selection of the "full" model, which describes the underlying DGP and with respect to which the parameters of interest are defined, is of equal importance. If the full model is a VAR, then a single equation model is a "partial" model obtained through a reduction of the joint distribution given weak exogeneity assumptions. In such a case the parameters of interest have to be defined in terms of the VAR. On the other hand, if the full model is a single equation model, as in (10), the parameters of interest should be defined in terms of the parameters of the single equation itself. In other words, when the VAR is assumed to be the full model, then the parameters of interest are the two eigenvalues. As Watson (1994) puts it: *"But of course there is nothing inherently special or natural about the finite order VAR; it is just one flexible parameterisation of the process"*. When an alternative parameterisation is used, the set of parameters of interest will change accordingly.

Point two is related to the debate on model specification and the "VAR methodology". One possible approach consists in estimating a VAR model as an approximation of the true DGP, with the role of economic theory at the specification stage being restricted to suggesting which variables should enter the VAR. However, once the VAR has been specified, the parameters of interest still need to be defined within the adopted

1. Caporale.p65 7 16/02/03, 03:54 p.m.

framework. Partial or conditional models should then be employed for the set of variables which are weakly exogenous with respect to the parameters of interest.

The new "structure" obtained from the conditional model is based on purely statistical grounds, namely by testing whether the restrictions that led to this structure are supported by the data. For example, in the context of the first-order bivariate VAR model analysed above, weak exogeneity of x_t for the cointegrating vector is equivalent to the notion of Granger non-causality running from y_t to x_t (see Johansen (1992)). The validity of this restriction points to uni-directional (Granger) causality from x_t to y_t, which simply means that past values of x_t are useful in predicting future values of y_t and not the other way round. Here, economic theory was not exploited in any way to determine the direction of causality. On the contrary, in the context of model (8) the assumption that x_t causes y_t was imposed *a priori*, presumably on the basis of economic theory. As we saw, this led to defining as an appropriate measure of persistence the parameter ϕ instead of λ_1 and λ_2.

The preceding discussion does not imply in any way that model (10) is less "statistical" than the VAR model, but simply that "more" economic theory has been used in the former case, which has enabled the researcher to "correctly" define and measure the degree of persistence. To put it differently, the decision whether to treat variables as endogenous or exogenous, and hence whether to examine persistence in the context of a VAR, or, alternatively, of a dynamic single equation, should be based mainly on economic theory. In the case of persistence, therefore, the more traditional concept of *model exogeneity* appears to be more useful (see Geweke (1990)).

IV. CONCLUSIONS

The persistence of macroeconomic time series is an issue which has attracted a lot of attention in recent years, but is far from being resolved, despite the numerous studies carried out to date (see, e.g., Cochrane (1988), and Campbell and Mankiw (1987a, 1987b)). This paper has argued that this is not very surprising, as measures of persistence are *model dependent,* and are therefore a function of the maintained model which economic theory suggests. Furthermore, statistical inference requires defining the set of *parameters of interest.* For instance, in a simple bivariate case, shifting the focus from the cointegrating parameters to the degree of persistence is equivalent to a change in the parameters of interest. This implies that *model exogeneity* becomes relevant, as weak exogeneity cannot hold in this case – once again economic theory considerations can be seen to play an essential role in obtaining "correct" estimates of persistence.

However, imposing more restrictions derived from economic theory also means that empirical evidence will be conditional upon the adopted theoretical framework. Then the implied answer to the question "How persistent is GDP?" (or some other macroeconomic variable) might be meaningful only in that particular context. Consequently, empirical analysis might not be able to settle conclusively issues such as the effectiveness of stabilisation policies and the relative merits of Keynesian versus RBC theories.

REFERENCES

Campbell, J. and N. Mankiw (1987a). "Are output fluctuations transitory?". *Quarterly Journal of Economics* 102, 857-880.

Campbell, J. and N. Mankiw (1987b). "Permanent and transitory components in macroeconomic fluctuations". *American Economic Review Papers and Proceedings* 77, 111-117.

Christiano, L. and M. Eichenbaum (1989), "Unit roots in GNP: do we know and do we care?". *Carnegie-Rochester Conference Series on Public Policy* 32, 7-62.

Clark, P. (1987). "The cyclical component of US economic activity". *Quarterly Journal of Economics* 102, 857-880.

Cochrane, J. (1988). "How big is the random walk in GNP?". *Journal of Political Economy* 96, 893-920.

Cochrane, J. (1991). "Comment on `Pitfalls and opportunities: what macroeconomists should know about unit roots`". *NBER Macroeconomics Annual*, 201-210.

Cochrane, J. (1994). "Permanent and transitory components of GNP and stock prices". *Quarterly Journal of Economics* 109, 1, 241-265.

Cochrane, J. and A. Sbordone (1988). "Multivariate estimates of the permanent component of GNP and stock prices". *Journal of Economic Dynamics and Control* 12, 255-296.

Engle, R., Hendry, D. And J. Richard (1983). "Exogeneity". *Econometrica* 51, 2, 277-304.

Evans, G. And L. Reichlin (1994). "Information, forecasts, and measurement of the business cycle". *Journal of Monetary Economics* 33, 233-254.

Geweke, J. (1990). "Endogeneity and exogeneity". In Eatwell, J., M. Milgate and P. Newman (eds.). *The New Palgrave – Econometrics*. MacMillan, London.

Johansen, S. (1988), "Statistical analysis of cointegration vectors". *Journal of Economic Dynamics and Control* 12, 231-254.

Johansen, S. (1991). "Estimation and hypothesis testing of cointegration". *Econometrica* 59, 1551-1581.

Johansen, S. (1992). "Cointegration in partial systems and the efficiency if single equation analysis". *Journal of Econometrics* 52, 389-402.

1. Caporale.p65 9 16/02/03, 03:54 p.m.

Lippi, M. and L. Reichlin (1992). "On persistence of shocks to economic variables – A common misconception". *Journal of Monetary Economics* 29, 87-93.

Lo, A. (1991). "Long-term memory in stock prices". *Econometrica* 59, 1279-1313.

Lupi, C. (1993). "Persistence in macroeconomic time series: some general results". Paper presented at the Econometric Society European Meeting. Uppsala University, Uppsala, Sweden.

Lutkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Springer-Verlag. Berlin.

Nelson, C. and C. Plosser (1982). "Trends and random walks in macroeconomic time series: some evidence and implications". *Journal of Monetary Economics* 10, 139-162.

Priestley, M. (1981) . *Spectral Analysis and Timer Series*. London, Academic Press.

Reinsel, G. (1991). *Elements of Multivariate Timer Series Analysis*. Springer-Verlag, Berlin.

Watson, M. (1986). "Univariate deternding methods with stochastic trends". *Journal of Monetary Economcis* 18, 49-75.

Watson, M. (1994). "Vector autoregressions and cointegration". In Engle, R. and McFadden D. (eds), *Handbook of Econometrics.* Vol IV, 2843-2915. Elsevier, Amsterdam.

1. Caporale.p65 10 16/02/03, 03:54 p.m.