Advanced Monetary Economics
Notes on cointegration and VAR models

F.C. Bagliano

March 2009

1 Cointegration in a VAR: the bivariate case

To introduce cointegration analysis in a multivariate framework and the related *vector error correction (VECM)* form, we start from a simple bivariate example (following Engle and Granger, *Econometrica*, 1987). Consider two $I(1)$ variables, y and x, generated by the following model:

$$y_t - \gamma\, x_t = u_{1t} \qquad u_{1t} = \rho\, u_{1t-1} + \varepsilon_{1t} \tag{1}$$

$$\delta\, y_t + x_t = \theta_1\, y_{t-1} + \theta_2\, x_{t-1} + u_{2t} \qquad u_{2t} = u_{2t-1} + \varepsilon_{2t} \tag{2}$$

where ε_{1t} and ε_{2t} are *white noise* processes. Imposing $|\ \rho\ | < 1$ makes the two variables cointegrated of order (1,1) since equation (1) describes a stationary linear combination of y and x with cointegrating vector $(1, -\gamma)$. By differencing and substituting for u_{1t} into (1)[1] the following reduced form is obtained, where a term in the *levels* of the variables captures the tendency of the system to "correct" any deviation from the long-run, cointegrating

[1] Differencing (1) we obtain on the right-hand side

$$\Delta u_{1t} = (\rho - 1)\, u_{1t-1} + \varepsilon_{1t}$$

and, lagging (1) by one period:

$$\Delta u_{1t} = (\rho - 1)\, (y_{t-1} - \gamma\, x_{t-1}) + \varepsilon_{1t}$$

relationship:

$$\begin{aligned}
\Delta y_t &= -\frac{1-\rho}{1+\gamma\delta}(y_{t-1}-\gamma x_{t-1}) + \frac{\gamma\theta_1}{1+\gamma\delta}\Delta y_{t-1} + \frac{\gamma\theta_2}{1+\gamma\delta}\Delta x_{t-1} \\
&\quad + \frac{1}{1+\gamma\delta}(\varepsilon_{1t}+\gamma\varepsilon_{2t}) \\
\Delta x_t &= \delta\frac{1-\rho}{1+\gamma\delta}(y_{t-1}-\gamma x_{t-1}) + \frac{\theta_1}{1+\gamma\delta}\Delta y_{t-1} + \frac{\theta_2}{1+\gamma\delta}\Delta x_{t-1} \\
&\quad + \frac{1}{1+\gamma\delta}(\varepsilon_{2t}-\delta\varepsilon_{1t})
\end{aligned}$$

The existence of this *error-correction* representation of the system in (1) and (2) crucially depends on the magnitude of ρ: if $\rho = 1$, y and x are not cointegrated and the error-correction terms vanish (the reduced form is simply a $VAR(1)$ for the variables expressed in differenced form). In matrix notation we have:

$$\begin{pmatrix} \Delta y_t \\ \Delta x_t \end{pmatrix} = \mathbf{\Pi} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \mathbf{\Gamma} \begin{pmatrix} \Delta y_{t-1} \\ \Delta x_{t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

This correspondence between cointegrating vectors and error-correction representations (one proposition of the "Granger representation theorem") is the basis of the simple two-step Engle-Granger estimation procedure, whereby the long-run equilibrium and the short-run dynamics are modelled sequentially. In the first step, after pre-testing the variables entering the cointegrating relation in order to ensure that they are of the same order of integration, an estimate of the cointegrating vector is obtained by means of a static OLS regression. Under the hypothesis of cointegration, such regression with the variables in levels yields *superconsistent* estimates of the cointegrating vector, with the parameters rapidly converging to their true values (Stock, *Econometrica*, 1987). The intuition behind this result is that, since in general a linear combination of $I(1)$ variables is also $I(1)$, almost all vectors obtained in static levels regressions will yield a residual series with asymptotically infinite variance. The exception will be any cointegrating vector. Since OLS estimation minimizes the residual variance, the estimated vector derived from a static OLS regression should yield a very good approximation to a true cointegrating vector, if it exists. In the second step of the procedure, the residuals from the cointegrating regression are used as an error-correction term in a dynamic model for the differenced variables in order to model the short-run adjustment dynamics.

However, when more than two variables are involved in the analysis, *multiple cointegrating vectors* may exist; in this case, the first-step Engle-Granger static equation yields an (obviously stationary) linear combination of the cointegrating vectors with no means to separate them. Appropriate tools to address the issue of multiple cointegrating vectors are needed.

2 Cointegration in a VAR: the multivariate case

To illustrate the case of multiple cointegrating vectors (analysed by Johansen, *Journal of Economics Dynamics and Control 1988, Econometrica* 1991), consider the *VAR* process for $\mathbf{y}_t$, a vector of n non-stationary ($I(1)$) variables (with only two lags for simplicity):

$$\mathbf{y}_t = \boldsymbol{\delta} + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + \mathbf{u}_t$$

(where $\boldsymbol{\delta}$ is a vector of constant terms) rewritten as a reduced-form error-correction model as follows

$$\begin{aligned} \Delta\mathbf{y}_t &= \boldsymbol{\delta} + (\mathbf{A}_1 + \mathbf{A}_2 - \mathbf{I})\,\mathbf{y}_{t-1} - \mathbf{A}_2\,\Delta\mathbf{y}_{t-1} + \mathbf{u}_t \\ \Rightarrow \Delta\mathbf{y}_t &= \boldsymbol{\delta} + \boldsymbol{\Pi}\,\mathbf{y}_{t-1} + \boldsymbol{\Gamma}\Delta\,\mathbf{y}_{t-1} + \mathbf{u}_t \end{aligned} \tag{3}$$

The elements of matrix $\boldsymbol{\Pi}$ contain the long-run relations among the levels of the variables in $\mathbf{y}$ (obtained by setting $\Delta\mathbf{y}_t = \Delta\mathbf{y}_{t-1} = \mathbf{0}$ and $\mathbf{u}_t = \mathbf{0}$). Each row of $\boldsymbol{\Pi}$ defines a linear combination of the elements in $\mathbf{y}$ that, if stationary, represents a valid long-run relation (i.e. a *cointegrating vector*). With $n > 2$ variables in $\mathbf{y}$ there may exist more than one cointegrating vector capturing the long-run behaviour of the data. The number of linearly independent rows of $\boldsymbol{\Pi}$ (the *rank* of $\boldsymbol{\Pi}$) yields the number of valid cointegrating vectors, that is the number of distinct stationary linear combinations of the n non-stationary variables.

Various cases can arise:

- if $\text{rank}(\boldsymbol{\Pi}) = 0 \Rightarrow \boldsymbol{\Pi} = \mathbf{0}$ and no cointegrating vector exist; there are n stochastic trends in the system, each driving one of the $I(1)$ variables in $\mathbf{y}$. The stationary form of the system is in first differences of all variables, with no term in levels.
- if $\text{rank}(\boldsymbol{\Pi}) = n$ (full rank) $\Rightarrow$ there is a set of n independent restrictions on the long-run values of the elements in $\mathbf{y}$. To derive the long-run

solution, setting $\Delta\mathbf{y}_t = \Delta\mathbf{y}_{t-1} = \mathbf{0}$ and $\mathbf{u}_t = \mathbf{0}$, we get

$$\begin{aligned}
\pi_{11}y_{1t-1} + \pi_{12}y_{2t-1} + ... + \pi_{1n}y_{nt-1} &= -\delta_1 \\
\pi_{21}y_{1t-1} + \pi_{22}y_{2t-1} + ... + \pi_{2n}y_{nt-1} &= -\delta_2 \\
... \quad ... \quad ... & \\
\pi_{n1}y_{1t-1} + \pi_{n2}y_{2t-1} + ... + \pi_{nn}y_{nt-1} &= -\delta_n
\end{aligned}$$

If $\mathbf{\Pi}$ is of full rank this system admits a solution for the long-run values of y_1, y_2, ... y_n, that we call $\bar{y}_1$, $\bar{y}_2$, ... $\bar{y}_n$. In matrix terms:

$$\bar{\mathbf{y}} = -\mathbf{\Pi}^{-1}\boldsymbol{\delta}$$

Vactor $\bar{\mathbf{y}}$ captures the unconditional mean of the n-dimensional stochastic process $\{\mathbf{y}_t\}$: $E(\mathbf{y}_t) = \bar{\mathbf{y}}$. Then , *all variables are stationary* and no stochastic trend is present in the system.

- if rank$(\mathbf{\Pi}) = r < n \Rightarrow$ there are r *cointegrating vectors* (capturing the only r linear combinations of the variables in $\mathbf{y}$ that are stationary) and the number of *common stochastic trends* in the system is $n - r$.[2] In this case the correct representation of the system is the **vector error-correction mechanism** ($VECM$) in (3), including the term in the

[2] (To clarify this point) **Example** for $n = 3$ and $r = 2$. Consider 3 non-stationary variables generated by random walk plus noise processes:

$$y_t^i = \mu_t^i + \varepsilon_t^i \qquad \text{with } i = 1, 2, 3$$

where μ_t^i $(i = 1, 2, 3)$ are the stochastic trends and ε_t^i $(i = 1, 2, 3)$ are independent $I(0)$ processes. If there exist two cointegrating vectors among the three variables $(r = 2)$, then the coefficients of each vector eliminate the linear combination of the stochastic trends in the y variables. Denoting with β_{i1} and β_{i2} $(i = 1, 2, 3)$ the coefficients of the two cointegrating vectors, we have:

$$\begin{aligned}
\beta_{11}\mu_t^1 + \beta_{21}\mu_t^2 + \beta_{31}\mu_t^3 &= 0 \\
\beta_{12}\mu_t^1 + \beta_{22}\mu_t^2 + \beta_{32}\mu_t^3 &= 0
\end{aligned}$$

It is possible to express two stochastic trends in terms of the third, for example as:

$$\begin{aligned}
\mu_t^1 &= \frac{\beta_{21}\beta_{32} - \beta_{31}\beta_{22}}{\beta_{11}\beta_{22} - \beta_{21}\beta_{12}}\,\mu_t^3 \\
\mu_t^2 &= \frac{\beta_{31}\beta_{12} - \beta_{11}\beta_{32}}{\beta_{11}\beta_{22} - \beta_{21}\beta_{12}}\,\mu_t^3
\end{aligned}$$

Therefore, in the 3-variable system, there exist only 1 stochastic trend (for example, μ_t^3) which is common to all 3 non-stationary variables (μ_t^1 and μ_t^2 being simply proportional to μ_t^3).

levels of the variables. If the $n \times n$ matrix $\mathbf{\Pi}$ has reduced rank ($r < n$), it can be written as the product of two $n \times r$ matrices, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, as follows:

$$
\begin{aligned}
\mathbf{\Pi} &= \underbrace{\boldsymbol{\alpha}}_{n \times r} \underbrace{\boldsymbol{\beta}'}_{r \times n} \\
&= (\boldsymbol{\alpha}_1 \ ... \ \boldsymbol{\alpha}_r) \begin{pmatrix} \boldsymbol{\beta}_1' \\ ... \\ \boldsymbol{\beta}_r' \end{pmatrix}
\end{aligned}
$$

where $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ ($i = 1, ...r$) denote the columns of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ respectively. Matrix $\boldsymbol{\beta}$ contains the cointegrating vectors (one in each column), whereas matrix $\boldsymbol{\alpha}$ contains the weights (*loadings*) with which each cointegrating vector enters the n equations in the VAR. Since $\boldsymbol{\beta}'\mathbf{y}_{t-1}$ represents the deviations of the variables from the set of r long-run equilibrium relations, the coefficients in $\boldsymbol{\alpha}$ measure the adjustment of $\Delta\mathbf{y}_t$ to the system's long-run equilibrium. As an example, consider a system of $n = 4$ non-stationary variables with $r = 2$ cointegrating vectors. In the $VECM$ representation of the system the term in levels, $\mathbf{\Pi y}_{t-1}$ is expressed as

$$
\begin{aligned}
\mathbf{\Pi}\,\mathbf{y}_{t-1} &= \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \\ \alpha_{41} & \alpha_{42} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} \\ \beta_{12} & \beta_{22} & \beta_{32} & \beta_{42} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \\ y_{4t-1} \end{pmatrix} \\
&= \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \alpha_{41} \end{pmatrix} \boldsymbol{\beta}_1' \mathbf{y}_{t-1} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \alpha_{32} \\ \alpha_{42} \end{pmatrix} \boldsymbol{\beta}_2' \mathbf{y}_{t-1}
\end{aligned}
$$

whereby the previous period deviation of the variables from the first (second) long-run equilibrium relation, given by the first (second) cointegrating vector, enters the equations of the $VECM$ with weights given by the elements of the first (second) column of $\boldsymbol{\alpha}$.

3 A note on cointegration tests

When more than two variables are involved in the analysis, the possibility arises of multiple cointegrating vectors. In this case, the Engle-Granger estimation procedure is not appropriate anymore, since a static cointegrating regression would yield a linear combination of all the valid cointegrating vectors. To test for the existence and number of cointegrating vectors, the Johansen (maximum likelihood) procedure is usually employed.

The Johansen approach tests hypotheses on the rank of $\mathbf{\Pi}$, which corresponds to the number of cointegrating vectors in the system. The Johansen test exploits the fact that, if there is cointegration and therefore $\mathbf{\Pi}$ has not full rank, then

$$\det(\mathbf{\Pi}) = \lambda_1\, \lambda_2 \ldots \lambda_n = 0$$

where the λ_i are the n eigenvalues (characteristic roots) of $\mathbf{\Pi}$, and the number of valid cointegrating vectors is equal to the number of non-zero eigenvalues.[3] The statistics proposed by Johansen (so called λ_{TRACE} and λ_{MAX}) test for the number of eigenvalues λ that are significantly different from zero (with slightly different null and alternative hypotheses):

$$\begin{aligned} \lambda_{TRACE} &= -T \sum_{i=k+1}^{n} \ln(1-\hat{\lambda}_i) \qquad \text{with } H_0: \; r \leq k \; \text{ and } \; H_A: \; r > k \\ \lambda_{MAX} &= -T \, \ln(1-\hat{\lambda}_k) \qquad \text{with } H_0: \; r = k-1 \; \text{ and } \; H_A: \; r = k \end{aligned}$$

The Johansen's maximum-likelihood procedure leads also to estimation of the elements of Π with the reduced rank restriction imposed (i.e. imposing that $\mathbf{\Pi}$ has a rank equal to the number of valid cointegrating vectors detected by the λ_{TRACE} and λ_{MAX} tests); note that, since the reduced rank of $\mathbf{\Pi}$ implies cross-equation restrictions on the VAR system, OLS estimation (equation by equation) is not implementable here, and a system procedure is needed. The estimated $(n \times n)$ matrix $\mathbf{\Pi}$ is then expressed as

$$\mathbf{\Pi} = \boldsymbol{\alpha}\, \boldsymbol{\beta}'$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $n \times r$ matrices.

However, the estimated coefficients of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ delivered by the Johansen's procedure suffer from a fundamental identification problem: they

---

[3] The eigenvalues (characteristic roots) of $\mathbf{\Pi}$ are scalars λ such that

$$\mathbf{\Pi}\, \mathbf{x} = \lambda\, \mathbf{x} \qquad (*)$$

where $\mathbf{x}$ is a non-null vector. (*) can be rewritten as

$$(\mathbf{\Pi} - \lambda\, \mathbf{I})\, \mathbf{x} = \mathbf{0}$$

where $\mathbf{I}$ is the identity matrix. Since $\mathbf{x} \neq \mathbf{0}$, the matrix $\mathbf{\Pi} - \lambda\, \mathbf{I}$ must not be of full rank (its rows must be linearly dependent), implying that

$$\det(\mathbf{\Pi} - \lambda\, \mathbf{I}) = 0$$

The characteristic roots of $\mathbf{\Pi}$ are the values λ that satisfy this (polynomial) equation.

are obtained from the estimated reduced rank $\mathbf{\Pi}$ by imposing an arbitrary normalization which rules out an immediate economic interpretation for the estimated cointegrating vectors. To see this point note that for any non-singular $r \times r$ matrix $\mathbf{\Theta}$ we can express the same long-run matrix $\mathbf{\Pi}$ as

$$\Pi = (\boldsymbol{\alpha}\mathbf{\Theta})\left(\mathbf{\Theta}^{-1}\boldsymbol{\beta}'\right)$$

The cointegrating vectors yielded by the Johansen procedure might then be linear combinations (obviously stationary) of the r valid cointegrating vectors of economic interest. In order to identify the long-run economic relationships of interest, structural long-run hypotheses must be imposed and tested on the elements of $\boldsymbol{\beta}$.

The idea of the test is that only r linear combinations of the variables are stationary (i.e. the valid cointegrating vectors), whereas all different combinations are non-stationary. If the imposed restrictions on the elements of $\boldsymbol{\beta}$ define linear combinations of the variables that are "very far" from the stationary ones, then the number of estimated cointegrating vectors (i.e. non-zero eigenvalues) of the restricted system should be reduced. In this case, a test comparing the eigenvalues of the unrestricted and restricted systems should reject the long-run restrictions. The opposite occurs when the imposed restrictions are not binding.

4 An example of cointegrated VAR: Cochrane (1994)

Cochrane (1994) studies the dynamic effects and relative importance of permanent and transitory components in the behaviour of some macroeconomic (GNP and aggregate consumption) and financial (stock prices and dividends) series using a bivariate cointegrated VAR.

Focusing on the **GNP-consumption** case, a bivariate VAR system with two lags of the rates of change of consumption and GNP (Δc_t and Δy_t) can be specified as (costant terms omitted):

$$\begin{pmatrix} \Delta c_t \\ \Delta y_t \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} c_{t-1} \\ y_{t-1} \end{pmatrix} + \mathbf{\Gamma}_1 \begin{pmatrix} \Delta c_{t-1} \\ \Delta y_{t-1} \end{pmatrix} + \mathbf{\Gamma}_2 \begin{pmatrix} \Delta c_{t-2} \\ \Delta y_{t-2} \end{pmatrix} + \begin{pmatrix} u_t^c \\ u_t^y \end{pmatrix}$$

where c and y are the logs of consumption and GNP, both $I(1)$ series, and u_t^c and u_t^y are the (reduced form) VAR innovations in consumption and GNP respectively. Cointegration between c and y with cointegrating vector $(1, -1)$ (given the strong evidence of stationarity of the consumption/GNP ratio) is imposed on the system, yielding the following $VECM$ model to be estimated:

$$\begin{pmatrix} \Delta c_t \\ \Delta y_t \end{pmatrix} = \begin{pmatrix} \alpha_c \\ \alpha_y \end{pmatrix} (c_{t-1} - y_{t-1}) + \mathbf{\Gamma}_1 \begin{pmatrix} \Delta c_{t-1} \\ \Delta y_{t-1} \end{pmatrix} + \mathbf{\Gamma}_2 \begin{pmatrix} \Delta c_{t-2} \\ \Delta y_{t-2} \end{pmatrix} + \begin{pmatrix} u_t^c \\ u_t^y \end{pmatrix}$$

Estimation of this system (Table I in the paper) yields $\hat{\alpha}_y = 0.08$ (t-stat. $= 3.45$) pointing to the importance of lagged consumption/GDP ratio in predicting future GNP movements, whereas $\hat{\alpha}_c$ is not statistically significantly different from zero. Moreover, consumption is nearly a random walk.

Impulse response functions (Figure I) are derived after imposing the following recursive (Choleski) identification scheme:

$$\begin{aligned} u_t^c &= v_t^c \\ u_t^y &= a\, u_t^c + v_t^y \end{aligned}$$

where v^c and v^y are "structural disturbances". Within this framework, a disturbance v_t^y that affects y without changing c contemporaneously has no long-run effect on c (since, for a random walk, the contemporaneous response to a shock is the long-run response as well); and, given cointegration between c and y, it has no long-run effect on y. Therefore, the economic interpretation of v_t^y is that of a *transitory* shock, with no long-run effect on GNP and consumption. On the other hand, v_t^c has contemporaneous effects on c and y and also long-run effects on both variables; its interpretation is then that of a *permanent* shock, driving GNP and consumption in the long-run.

This identification scheme has a rationale within the theroretical framework of the "*permanent income model*" of consumption (with rational expectations), according to which consumption should follow a random walk and should be cointegrated with income. A simple formalization of this idea that generates a version (with no lags) of the VAR system estimated by Cochrane is given by the following three equations:

$$\begin{aligned} y_t &= y_t^P + \eta_t \qquad\qquad \eta_t = \rho\, \eta_{t-1} + \nu_t \quad (0 < \rho < 1) && (4) \\ y_t^P &= y_{t-1}^P + \mu + \varepsilon_t && (5) \\ c_t &= y_t^P && (6) \end{aligned}$$

Equation (4) defines observed GNP as the sum of two components: the unobserved "permanent income" component (y_t^P) and a stationary $AR(1)$ process (η_t), interpreted as "transitory income"; v_t is a white noise innovation to transitory income. Permanent income in (5) is generated by a random walk (with drift μ) plus noise stochastic process, where ε_t is a white noise uncorrelated with η_t. Finally, according to (6), consumption is equal to permanent income and is not affected by the transitory component of GNP.[4] First, note that in this model consumption and income are cointegrated with cointegrating vector $(1,-1)$ since $c_t - y_t = -\eta_t$ with $\eta_t \sim I(0)$. Therefore, a $VECM$ representation exists and may be derived by taking first differences of (4)-(6) and using the fact that $\Delta\eta_{t-1} = (1-\rho)(c_{t-1} - y_{t-1}) + v_t$. The following cointegrated VAR system is obtained

$$\begin{pmatrix} \Delta c_t \\ \Delta y_t \end{pmatrix} = \begin{pmatrix} \mu \\ \mu \end{pmatrix} + \begin{pmatrix} 0 \\ 1-\rho \end{pmatrix} (c_{t-1} - y_{t-1}) + \begin{pmatrix} u_t^c \\ u_t^y \end{pmatrix}$$

where the innovations are given by:

$$\begin{aligned} u_t^c &= \varepsilon_t \\ u_t^y &= \varepsilon_t + v_t \end{aligned}$$

Therefore, the model above generates a recursive structure in the relation between the VAR innovations and the structural disturbances ε_t and v_t: the innovation in consumption growth captures only disturbances to permanent income, that also affect contemporaneously the GNP growth rate, whereas transitory income shocks have a contemporaneous impact only on Δy_t.

After this identification is imposed, innovation accounting is carried out with some notable results:

- *the long-run responses of c and y to the two structural disturbances are the same*. Intuitively, since the consumption/GNP ratio is stationary, in the long-run the two variables have to show the same response to any shock to restore the ratio. This point can be proved formally using a simplified version of Cochrane's $VECM$ (no constant terms and no lags), in which c follows a random walk exactly, i.e.

$$\begin{aligned} \Delta c_t &= u_t^c \\ \Delta y_t &= \alpha_y (c_{t-1} - y_{t-1}) + u_t^y \end{aligned}$$

[4] It can be shown that a standard version of the "permanent income model" implies that consumption is set in each period t at the level given by income expected in the long-run (less deterministic growth), that is

$$c_t = \lim_{h\to\infty} (E_t\, y_{t+h} - h\,\mu) \tag{*}$$

Using (4) and (5) above we have

$$E_t\, y_{t+h} = E_t\, y_{t+h}^P + \underbrace{E_t\, v_{t+h}}_{0} = y_t^P + h\,\mu$$

which, substituted into (*), yields

$$c_t = y_t^P$$

as in (6).

Subtracting the second equation from the first we get

$$\begin{aligned}\Delta c_t - \Delta y_t &= -\alpha_y (c_{t-1} - y_{t-1}) + (u_t^c - u_t^y) \\ \Rightarrow c_t - y_t = & \quad (1-\alpha_y)(c_{t-1} - y_{t-1}) + (u_t^c - u_t^y)\end{aligned}$$

which is a (stationary) $AR(1)$ process for $c - y$. To derive the VMA (vector moving average) representation for Δc_t and Δy_t we first express $c - y$ as

$$\begin{aligned}[1-(1-\alpha_y)L]\,(c_t - y_t) &= u_t^c - u_t^y \\ \Rightarrow c_t - y_t = & \quad \frac{1}{1-(1-\alpha_y)L}(u_t^c - u_t^y)\end{aligned}$$

and then substitute it into the equation for Δy_t, obtaining

$$\begin{aligned}\Delta y_t &= \alpha_y \frac{1}{1-(1-\alpha_y)L} L\,(u_t^c - u_t^y) + u_t^y \\ &= \frac{\alpha_y}{1-(1-\alpha_y)L} L\,u_t^c + \left(1 - \frac{\alpha_y}{1-(1-\alpha_y)L} L\right) u_t^y\end{aligned}$$

The $VMA(\infty)$ representation of the bivariate system is then

$$\begin{pmatrix} \Delta c_t \\ \Delta y_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\alpha_y}{1-(1-\alpha_y)L} L & 1 - \frac{\alpha_y}{1-(1-\alpha_y)L} L \end{pmatrix} \begin{pmatrix} u_t^c \\ u_t^y \end{pmatrix}$$

The long-run response of c and y to the two innovations is found simply by taking $L = 1$:

$$\begin{pmatrix} \Delta c \\ \Delta y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u^c \\ u^y \end{pmatrix}$$

showing the same long-run responses.

- *the transitory component of GNP is quantitatively important* (accounting for 70% of the long-run forecast error variance of GNP growth).

Exercise.
Consider the following alternative "permanent income" model in which:

(i) permanent income y^P is directly *observable* (by assumption);

(ii) consumption adjusts *gradually* to permanent income.

Formally, the model is:

$$\begin{aligned} c_t &= y_t^P + \zeta_t \quad \text{with} \quad \zeta_t = \rho\,\zeta_{t-1} + \eta_t \quad (0 < \rho < 1) \\ y_t^P &= y_{t-1}^P + \mu + \varepsilon_t \end{aligned}$$

where η_t and ε_t are uncorrelated.

1. Find the cointegrated $VECM$ representation of the system formed by Δc_t and Δy_t^P;

2. find the (recursive) identification scheme appropriate to recover the impulse response functions of c and y^P to permanent and transitory disturbances.

Answer.

1. First differencing and using the process for ζ_t we get:

$$\begin{aligned} \Delta c_t &= \Delta y_t^P + \Delta \zeta_t \\ &= \Delta y_t^P + (\rho - 1)\, \zeta_{t-1} + \eta_t \\ &= \Delta y_t^P + (\rho - 1)\,(c_{t-1} - y_{t-1}^P) + \eta_t \end{aligned}$$

   and the $VECM$ form is

$$\begin{pmatrix} \Delta c_t \\ \Delta y_t^P \end{pmatrix} = \begin{pmatrix} \mu \\ \mu \end{pmatrix} + \begin{pmatrix} -(1-\rho) \\ 0 \end{pmatrix} (c_{t-1} - y_{t-1}^P) + \begin{pmatrix} u_t^c \\ u_t^{y^P} \end{pmatrix}$$

   with $u_t^c = \eta_t + \varepsilon_t$ and $u_t^{y^P} = \varepsilon_t$.

2. The appropriate recursive structure to achieve identification of the permanent and transitory shocks is:

$$\begin{aligned} u_t^{y^P} &= \varepsilon_t \\ u_t^c &= \eta_t + \varepsilon_t \end{aligned}$$

   implying an ordering of the variables in the VAR with Δy_t^P first and Δc_t second. Now, the shock to permanent income ε_t affects consumption contemporaneously, whereas η_t, affecting only consumption and not y_t^P, is due to deviations from the long-run (cointegrating) relation with permanent income and has a purely transitory nature. In this alternative model, consumption is the series which contains a transitory component and therefore adjusts to the long-run equilibrium relation with permanent income (in fact, the cointegrating vector enters the equation for Δc_t only).