



A systematic framework for analyzing the dynamic effects of permanent and transitory shocks[☆]

Jesús Gonzalo^a, Serena Ng^{b,*}

^a*Department of Statistics and Economics, Universidad Carlos III de Madrid, Spain*

^b*Department of Economics, Boston College, Chestnut Hill, MA 02467, USA*

Received 6 March 1998; accepted 30 August 1999

Abstract

This paper proposes a systematic framework for analyzing the dynamic effects of permanent and transitory shocks on a system of n economic variables. We consider a two-step orthogonalization on the residuals of a VECM with r cointegrating vectors. The first step separates the permanent from the transitory shocks, and the second step isolates $n - r$ mutually uncorrelated permanent shocks and r transitory shocks. The decomposition is computationally straightforward and entails only a minor modification to the Choleski decomposition commonly used in the literature. We then show how impulse response functions can be constructed to trace out the propagating mechanism of shocks distinguished by their degree of persistence. In an empirical example, the dynamic responses to the identified permanent shocks have properties similar to shocks to productivity, the real interest rate, and money growth, even though no economic

* Corresponding author. Tel.: + 1-617-552-2182; fax: + 1-617-552-2308.

E-mail address: serena.Ng@bc.edu (S. Ng).

[☆]This paper was presented at the University of British Columbia, Boston College, University of California (Riverside), University of Toronto, and the winter meeting of the Econometric Society in San Francisco. We thank an anonymous referee for helpful comments. We also thank the seminar participants for helpful discussion and Paul Beaudry and Pierre Perron for comments on an earlier draft. The first author thanks the Foundation of Caja Madrid and the Spanish Secretary of Education (PB 950298) for financial support. The second author acknowledges grants from the Social Science and Humanities Research Council of Canada (SSHRC) and the Fonds de la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR).

theory was used to achieve the identification. We highlight two numerical issues that could affect any identification of permanent and transitory shocks. © 2001 Elsevier Science B.V. All rights reserved.

JEL classification: C2; C3; C5

Keywords: VAR; Impulse responses; Cointegration; Permanent and transitory shocks

1. Introduction

In this paper, we propose a simple and coherent framework for isolating the permanent and the transitory shocks from a system of integrated variables, making explicit the relationship between the common trends and the innovations underlying the reduced form model. We show how dynamic impulse response functions can be constructed to trace out the propagating mechanism of the permanent and the transitory shocks. The analysis is conducted using a VECM (Vector Error Correction Model), i.e. a vector autoregression (VAR) that incorporates cointegration restrictions. The procedure consists of two steps. The first step distinguishes innovations that have permanent effects from those that have transitory effects only. This is accomplished by a transformation of the residuals using information that are readily available from the VECM. The second step uses Choleski decomposition to obtain a set of permanent and transitory shocks that are mutually orthogonal. The advantage of the method is its simplicity, since it operates in much the same way a stationary VAR is used to do impulse response analysis. The only difference is that instead of applying Choleski decomposition to the residuals of the VAR, we apply it to a set of transformed residuals.

In conventional VAR analysis, the identified shocks are usually viewed as innovations to the variables in the system. Because identification is typically based on a priori assumptions, it is only in a limited sense the data reveal the source of the shock.¹ As forcefully argued in Cochrane (1994b), after decades of analysis, we still know very little and perhaps will never know enough about the origin of shocks. A key feature of cointegrated systems is that the variables move together at low frequencies. Given this coherence, our ability to identify the source of the shock is even more limited. But we can exploit the low frequency comovements to identify shocks according to whether their effects are

¹ For example, the so-called monetary policy shocks we back out of a VAR with money, prices, interest rate and output will depend on our view about the monetary policy transmission mechanism. See Bernanke and Mihov (1995).

permanent or transitory. Accordingly, in our analysis, the shocks are distinguished by their degree of persistence, rather than their origin. However, in general, the low frequency movements alone are not sufficient to identify permanent and transitory shocks that are mutually uncorrelated. We show how this can be achieved in a simple framework, and in the process, clarify the limits of cointegration in identifying the permanent and transitory shocks.

The plan of this paper is as follows. The econometric framework used to isolate the permanent and transitory shocks is presented in the next section. We then put into context our decomposition with related work in the literature. Section 3 presents simulated and empirical examples. Pitfalls of analysis on permanent and transitory components are discussed in Section 4. A conclusion completes the analysis.

2. The econometric framework

The objective of this section is to present a framework which systematically isolates the permanent and the transitory shocks from a VECM.

2.1. Preliminaries

Let the $(n \times 1)$ vector $\tilde{X}_t = Z_t + X_t$ be the sum of deterministic components Z_t (such as polynomials in time) and a $(n \times 1)$ vector of $I(1)$ time series, X_t . Our focus is on the detrended series X_t , which has a multivariate moving-average representation

$$\Delta X_t = C(L)e_t, \tag{1}$$

where $\Delta = 1 - L$, $Le_t = e_{t-1}$, and e_t is a $n \times 1$ vector satisfying $E(e_t) = 0$ and

$$E[e_t e'_s] = \begin{cases} 0 & \text{if } t \neq s, \\ \Omega & \text{otherwise.} \end{cases}$$

The matrix polynomial $C(L) = C(1) + (1 - L)C^*(L)$ has the property that $C_0 = I_n$, $C(z)$ is 1-summable, and $C^*(z)$ is full rank everywhere on $|z| \leq 1$.²

By the Granger Representation Theorem (see Engle and Granger, 1987), the vector X_t is said to be cointegrated with rank r if $C(1)$ is of rank $(n - r)$, and there exist two $n \times r$ matrices, α and γ , both of rank r , such that $\alpha' C(1) = 0$ and

²This assumption is necessary for the short-run dynamics to be fundamental (i.e. recoverable through a unique orthogonalization). See Lippi and Reichlin (1993) and Blanchard and Quah (1993) for discussions on the issue.

$C(1)\gamma = 0$.³ The columns of α are the cointegrating vectors of X_t . Furthermore, X_t has a VECM representation of infinite order which can be approximated by

$$\Delta X_t = \gamma\alpha'X_{t-1} + \Gamma(L)\Delta X_{t-1} + e_t, \tag{2}$$

where $\Gamma(L)$ is of finite order $K - 1$. The VECM can be used to deduce a restricted VAR representation:

$$A(L)X_t = e_t, \tag{3}$$

the restriction being $A(1) = \gamma\alpha'$. The rank of $A(1)$ is r . We are interested in expressing ΔX_t in terms of a set of permanent and transitory shocks. These are defined as follows:

Definition 1. Let X_t be a difference-stationary sequence whose VECM is given by (2). Let E_t denote the conditional expectation taken with respect to the information set in period t . The $(n - r) \times 1$ vector of shocks $\tilde{\eta}_t^p$ is said to have permanent effects on the level of X_t if $\lim_{h \rightarrow \infty} \partial E_t(X_{t+h}) / \partial \tilde{\eta}_t^{p'} \neq 0$. Analogously, the $r \times 1$ vector of shocks $\tilde{\eta}_t^T$ is said to have transitory effects on the level of X_t if $\lim_{h \rightarrow \infty} \partial E_t(X_{t+h}) / \partial \tilde{\eta}_t^{T'} = 0$.

Since the variables in X_t are $I(1)$ by definition, some of its innovations must have permanent effects on the levels of X_t . Therefore, ΔX_t also has the following representation:

$$\Delta X_t = \tilde{D}(L)\tilde{\eta}_t, \tag{4}$$

where $\tilde{D}(L) = \tilde{D}_0 + \tilde{D}_1L + \tilde{D}_2L^2 + \dots$, and the covariance matrix of $\tilde{\eta}$ is $\Sigma_{\tilde{\eta}}$. We assume that $\tilde{\eta}_t$ is a $n \times 1$ vector, so that the number of shocks equals the number of variables in the system,⁴ but only a subset has permanent effects. Without loss of generality, the $\tilde{\eta}_t$'s are ordered such that the first $n - r \neq 0$ of them have permanent effects.

2.2. The permanent and transitory (P-T) decomposition: Unorthogonalized shocks

Our objective is to recover $\tilde{\eta}_t$ from information available in the VECM. To accomplish this, we first find a transformation such that the data can be

³Note that because we work with X_t instead of \tilde{X}_t , deterministic components lie outside the cointegrating space in the sense of Johansen (1991).

⁴This ensures that $\tilde{D}(L)$ is invertible. See Watson (1994) for a discussion of invertibility.

expressed in terms of a set of ‘unorthogonalized’ permanent and transitory shocks, which we denote by u_t . We refer to this as the P–T decomposition.

Proposition 1. (The P–T decomposition). Let X_t be a $n \times 1$ vector of $I(1)$ processes with a Wold moving-average representation $\Delta X_t = C(L)e_t$. Suppose there are r cointegrating vectors. Let

$$G = \begin{bmatrix} \gamma'_\perp \\ \alpha' \end{bmatrix} \begin{matrix} (n-r) \times n \\ r \times n \end{matrix}$$

with $\gamma'_\perp \gamma = 0$. Then the $(n-r) \times 1$ vector $u_t^P = \gamma'_\perp e_t$ and the $r \times 1$ vector $u_t^T = \alpha' e_t$ are the permanent and transitory shocks, respectively. The P–T decomposition exists provided (γ_\perp, α') is non-singular.

The proof of the proposition is given in the Appendix. The choice of G is motivated by the Granger Representation Theorem which suggests that γ'_\perp and α' will ‘knock out’ the appropriate terms in the VECM and moving-average representations of ΔX_t to isolate those components with the desired degree of integration. The shocks u_t^P and u_t^T are, respectively, the innovations associated with the permanent and the transitory components of X_t as defined in Gonzalo and Granger (1995). Therefore, the G matrix completely identifies the trend and cycle of X_t as well as their underlying innovations.⁵ An implication of the P–T decomposition is that

$$\Delta X_t = C(L)G^{-1}Ge_t = D(L)u_t = \begin{bmatrix} D_{11}(L) & D_{12}(L) \\ D_{21}(L) & D_{22}(L) \end{bmatrix} \begin{bmatrix} u_t^P \\ u_t^T \end{bmatrix} \tag{5}$$

must have $D_{12}(1) = 0_{(n-r) \times r}$, and $D_{22}(1) = 0_{r \times r}$. This is because the last r columns of the polynomial matrix $D(L)$ are the responses of ΔX_t to the transitory shocks, and by definition, they have no effects on the first difference or the level of X_t in the long term.

2.3. The P–P and the T–T decompositions: Orthogonalized shocks

The G matrix rotates e_t so that it can be decomposed into u_t^P and u_t^T . But these shocks are not mutually uncorrelated. We now seek a transformation from

⁵ With a slight abuse of terminology, we use the terms permanent and transitory interchangeably with trends and cycles. Hence, white noise series will also be referred to as cycles. The trend component in the Granger–Gonzalo decomposition is an integrated process with possibly serially correlated noise and differs from the multivariate Beveridge–Nelson decomposition, considered in Stock and Watson (1988) and Evans and Reichlin (1994), in which the trend component is a random walk. In the working paper version of this paper, we show that the effect of the permanent innovations will have the same long run effect on the level of X_t regardless of how the permanent–transitory decomposition of X_t is attained.

$\Delta X_t = D(L)u_t$ to $\Delta X_t = \tilde{D}(L)\tilde{\eta}_t$ such that $\tilde{\eta}_t$ are mutually uncorrelated. From $\Omega = \tilde{D}_0 \Sigma_{\tilde{\eta}} \tilde{D}_0'$, it is immediate that we have $n(n + 1)/2$ equations to solve for $n^2 + n(n + 1)/2$ unknowns which is possible only with additional restrictions imposed. Comparing the two representations, we have the following:

$$\begin{aligned} \text{(R1)} \quad & D_0 u_t = \tilde{D}_0 \tilde{\eta}_t, \\ \text{(R2)} \quad & D(1) \Sigma_u D(1)' = \tilde{D}(1) \Sigma_{\tilde{\eta}} \tilde{D}(1)'. \end{aligned}$$

We now introduce the following assumption:

Assumption 1. $\tilde{\eta}_t = (\tilde{\eta}_t^P, \tilde{\eta}_t^T)'$ are serially and mutually uncorrelated and have unit variance.

The assumption imposes $n(n - 1)/2$ zero restrictions on the off-diagonals of $\text{cov}(\tilde{\eta})$. The unit variance property is a normalization assumption and puts n restrictions on the diagonal elements of $\text{cov}(\tilde{\eta})$. This leaves $n(n - 1)/2$ number of restrictions to be imposed on \tilde{D}_0 .

Consider a matrix H satisfying $HH' = \Sigma_u$. We now exploit the restrictions imposed by cointegration. First note that $D(1)$ and $\tilde{D}(1)$ must have the last r columns equal to zero since transitory shocks cannot have permanent effects on X_t . But (R1) and (R2) together imply that $D(1)D_0^{-1}\tilde{D}_0 = \tilde{D}(1)$. Therefore, if $\tilde{D}(1)$ and $D(1)$ were to have the last r columns equal to zero, $D_0^{-1}\tilde{D}_0$ must be a lower block triangular matrix. But by (R1) and Assumption 1, $\tilde{D}_0 = D_0 H$. Hence, H also has to be lower block triangular, thereby determining $r(n - r)$ coefficients in \tilde{D}_0 . This, however, is as far cointegration can guide us. After zero restrictions on H_{12} , we still need $(n - r)(n - r - 1)/2 + r(r - 1)/2$ restrictions on $\tilde{D}(1)$ and \tilde{D}_0 . The shocks are exactly identified by the P–T decomposition alone only when if there is exactly one permanent ($n - r = 1$) and one transitory ($r = 1$) shock. To complete the decomposition, we use the following assumption:

A practical rule. Let H be the Choleski decomposition of $\text{cov}(u)$, where $u_t = Ge_t$. Then $H^{-1}u_t = \tilde{\eta}_t$ achieves the P–P and the T–T decompositions.

Since the Choleski decomposition produces a lower triangular matrix, it puts the exact number of zero restrictions and we have a set of mutually uncorrelated shocks that also satisfy Assumption 1 by construction. Essentially, the submatrix H_{11} orthogonalizes \tilde{u}_t^P by putting $(n - r)(n - r - 1)/2$ zero restrictions, and the submatrix H_{22} orthogonalizes \tilde{u}_t^T through $r(r - 1)/2$ zero restrictions. The terms P–P and T–T decompositions are motivated by the consideration that the two decompositions can in fact be implemented separately.

Although $\tilde{\eta}_t$ is not unique because there is more than one way to choose H , the lower-block triangularity of H severely restricts the number of alternatives. The choice of the Choleski decomposition arises as a matter of convenience, but suffers from the usual limitation that the ordering of the variables will matter

(except in the bivariate case when $n - r = r = 1$). Nonetheless, the recursive structure is not as rigid as when Choleski decomposition is applied to $\text{cov}(e)$ because $D_0 = G^{-1}$ is *not* an identity matrix. Hence, the residuals of the first equation can have a non-zero weight in the second permanent shock via γ'_\perp . Furthermore, X_i can respond contemporaneously to permanent shock j even if $j > i$.

The complete P–T decomposition can now be summarized as follows:

$$\Delta X_t = C(L)G^{-1}HH^{-1}Ge_t = D(L)HH^{-1}u_t = \tilde{D}(L)\tilde{\eta}_t \tag{6}$$

with $\tilde{D}_{12}(1) = \tilde{D}_{22}(1) = 0$. To obtain this decomposition,

1. Decide the number of cointegrating vectors, r , and estimate a VECM with $\Gamma(L)$ of order $K - 1$ incorporating the cointegrating relationships. This yields consistent estimates of α and γ , denoted $\hat{\alpha}$ and $\hat{\gamma}$, from which one can construct $\hat{\gamma}'_\perp$;
2. Construct $\hat{G} = (\hat{\gamma}'_\perp \ \hat{\alpha}')$ and the set of permanent and transitory shocks, $\hat{G}\hat{e}_t$;
3. Obtain a lower block triangular matrix \hat{H} , such as by applying Choleski decomposition to $\text{cov}(\hat{G}\hat{e})$. The orthogonalized permanent and transitory shocks are $\hat{\eta}_t = \hat{H}^{-1}\hat{G}\hat{e}_t$. These have unit variance and are mutually uncorrelated.
4. Post-multiply $\hat{D}(L) = \hat{C}(L)\hat{G}^{-1}$ by \hat{H} to obtain an estimate of $\tilde{D}(L)$.

Thus, $\hat{D}(L)\hat{H}$ is the sample analog of the $\tilde{D}(L)$ matrix, and $\hat{\eta}_t$ is the sample analog of $\tilde{\eta}_t$. Impulse response functions and the decomposition of variances can then be constructed by canned routines in software packages, except that Choleski decomposition is applied to $\text{cov}(\hat{G}\hat{e})$ rather than to $\text{cov}(\hat{e})$.

The impulse response functions depend on the parameters of the model in a complex way and is of little use in constructing confidence bands, but we can assess sampling variability by bootstrapping. In our applications, we first condition on the number of cointegrating vectors, $\hat{\gamma}$ and $\hat{\alpha}$ are then estimated, and subsequently $\hat{\Gamma}$. There is little in the literature on bootstrapping that focuses on all three aspects of the procedure. We opted for a variant of the method discussed in Runkle (1987). It is implemented as follows: first, estimate α , and conditional on it, estimate the remaining parameters of the VECM to obtain the fitted residuals \hat{e}_t . A new sample of data is constructed (using the initial estimates of $\hat{\alpha}$, $\hat{\gamma}$, $\hat{\Gamma}(L)$) by random sampling of \hat{e}_t with replacement. Given a new sample of data, all the parameters are re-estimated holding the number of cointegrating vectors fixed, and the impulse response functions stored. This is repeated N times. We then evaluate the empirical standard errors from the N samples of the bootstrapped impulse response functions. In practice, N is set to 1000. Admittedly, the procedure is time consuming especially when n (the dimension of the model) and/or K (the lag length) is large.

2.4. Other approaches and limitations

Our two step decomposition gives $\tilde{D}(L)$, where $\tilde{D}_0 = D_0H$ and $\tilde{D}(1) = D(1)H$. But suppose we seek to obtain a specific $\tilde{D}(1)$. In our framework, this would require solving H , given $\tilde{D}(1)$ from economic theory and $D(1)$ from the P–T decomposition, such that $\tilde{D}(1) = D(1)H$, and subject to the condition that H must be lower block triangular. This is in the spirit of structural identification proposed by Bernanke (1986), and adopted in King et al. (1991) [hereafter, KPSW].⁶ The permanent shocks in KPSW are also determined by γ_{\perp} . This is not surprising since the permanent shocks are dictated by cointegration analysis. KPSW also invoked Assumption 1 and choose H_{11} to be lower triangular. There appears no way around a subjective choice on H_{11} because, as discussed earlier, cointegration restrictions are not sufficient for exact identification of the shocks.

Apart from the fact that the proposed decomposition is computationally simpler, our analysis differs from KPSW in two ways. First, KPSW used economic theory to pin down α and γ , whereas as presented, we impose long-run restrictions that are implied by the data rather than economic theory. However, our analysis still goes through, if we were to impose restrictions on α , and efficiency gains can always be obtained by imposing economic restrictions on the cointegrating vectors if the restrictions are correct. The second difference is that KPSW did not study the dynamic effects of the transitory shocks. In our setting, this can be accommodated by solving for H_{11} , given $\tilde{D}(1)$ from theory and $D(1)$ from the data, such that $\tilde{D}(1)_{n \times (n-r)} = D(1)_{n \times (n-r)}H_{11}$.

Another possibility, developed by Warne (1991), is to use $\gamma'\Omega^{-1}$ (instead of α') to define the transitory shocks. The advantage of Warne's approach is perhaps a sense of consistency since the permanent and transitory shocks both depend on γ . The advantage of our approach is that the two-steps necessary to obtain the decomposition are made explicit. In Warne's implementation, the identification is determined in one step. It is thus much less transparent when subjective assumptions are made, and when restrictions are based on long-run properties of the data.

Using a bivariate VECM, Cochrane (1994a) showed that shocks to GNP holding consumption constant are transitory and that shocks to consumption have persistent effects. But Cochrane obtained this result by directly applying Choleski decomposition to the residuals of the VECM. In other words, he accomplished the P–T decomposition in one step. To see why this is a very special case, let $X_t = (c_t, gnp_t)$ denote the logarithmic transforms of consumption and GNP and consider the estimates obtained in Cochrane (1994a). Using the notation of the previous section, $\hat{\alpha}'$ is $(1, -1)$ and $\hat{\gamma}'$ is $(-0.02, 0.08)$. Since

⁶ Another application using the same approach is Koray et al. (1995).

the $\hat{\gamma}$ in the consumption equation is not statistically significant and is constrained to zero, $\hat{\gamma}'_{\perp}$ is of the form $(x,0)$ for some $|x| > 0$. This implies GNP has no weight in the trend component, and all permanent shocks are due to consumption. In this bivariate model, the ordering of the variables is such that G is already lower triangular. If the order of X_t is reversed, γ'_{\perp} would have been $(0, x)$ and G would not have been lower triangular. Direct application of Choleski decomposition to $\text{cov}(\hat{e})$ would then be different from Choleski decomposition applied to $\text{cov}(\hat{u})$. But $u_t = Ge_t$ is invariant to the order of the variables in our P–T decomposition of two variables.

3. Simulations and examples

To see that the P–T decomposition functions well in practice, we provide two simple simulated examples. We assume that the rank of the cointegrating matrix is known to focus on orthogonalization issues. We use reduced rank regressions with two lags to obtain $\hat{\alpha}$.⁷ Conditional on $\hat{\alpha}$, unrestricted estimates of γ are obtained from the VECM (2). For each equation in the VECM, $\hat{\gamma}$ is constrained to zero and the equation is re-estimated if the unconstrained estimate of γ is not statistically significant at the two-tailed 5% level. This is important for the precision of the results and will be discussed in more detail in the next section. The null space of $\hat{\gamma}$ is spanned by the $r + 1$ through n left singular vectors of $\hat{\gamma}$. The sample size is 200 and there are 1000 simulations. Monte-carlo standard errors are also computed. The code is written in Gauss 3.21.

Example (DGP 1). The first DGP we considered is based on the following triangular representation:

$$\begin{aligned} \Delta x_t &= u_{1t}, \\ -x_t + y_t - z_t &= u_{2t}, \\ 0.5x_t + 0.5y_t + z_t &= u_{3t}, \end{aligned} \tag{7}$$

where u_{1t} , u_{2t} , and u_{3t} are $N(0,1)$ random errors that are mutually and serially uncorrelated. There are two cointegrated vectors and one common unit root. The dynamics are deliberately made simple so that the long and short run responses to the shocks can be easily verified.

⁷The cointegrating vectors are identified by solving the eigenvalue problem $|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}| = 0$. Let V be the matrix of r eigenvectors associated with the r largest eigenvalues. The identifying restrictions are that $V'S_{11}V = I$, and V diagonalize $S_{10}S_{00}^{-1}S_{01}$. The S_{ij} matrices are the sample moments of the residual cross correlation from projections of ΔX_t and X_{t-1} on $\Gamma(L)\Delta X_t$. Details are given in Johansen (1995).

Table 1

DGP 1 ($r = 2$), $\Delta x_t = u_{1t}$, $-x_t + y_t - z_t = u_{2t}$, $0.5x_t + 0.5y_t + z_t = u_{3t}$

Period	P Shock			T Shock 1			T Shock 2		
	x	y	z	x	y	z	x	y	z
<i>Impulse response functions</i>									
1	0.882	0.291	-0.665	0.118	0.035	0.656	-0.005	0.324	0.064
2	0.894	0.295	-0.597	-0.010	0.031	-0.046	-0.005	0.032	-0.019
3	0.893	0.298	-0.602	0.000	0.001	0.002	-0.000	0.002	-0.000
4	0.893	0.298	-0.602	-0.000	0.000	-0.000	-0.000	0.000	-0.000
5	0.893	0.298	-0.602	0.000	0.000	0.000	-0.000	0.000	-0.000
6	0.893	0.298	-0.602	-0.000	0.000	-0.000	-0.000	0.000	-0.000
SE_{\max}	0.000	0.011	0.011	0.006	0.071	0.111	0.004	0.034	0.052
<i>Decomposition of variances</i>									
1	0.987	0.112	0.569	0.013	0.001	0.397	0.000	0.886	0.033
2	0.993	0.202	0.703	0.006	0.002	0.273	0.000	0.796	0.025
3	0.996	0.278	0.774	0.004	0.002	0.207	0.000	0.720	0.019
4	0.997	0.341	0.818	0.003	0.002	0.167	0.000	0.658	0.015
5	0.997	0.393	0.848	0.003	0.001	0.140	0.000	0.605	0.013
6	0.998	0.438	0.869	0.002	0.001	0.120	0.000	0.561	0.011

The results are given in Table 1. The response of y_t and z_t to the permanent innovation driving x_t is close to the theoretical value of $1/3$ and $-2/3$ respectively. The variable x_t is uncorrelated with u_{2t} and u_{3t} , and the simulations reveal this property. The decomposition of variance correctly assesses the relative importance of the shocks.

Example (DGP 2). Data for the second DGP is generated as follows:

$$\begin{aligned}
 x_t &= y_t + 2z_t + u_{1t}, \\
 \Delta y_t &= u_{2t}, \\
 \Delta z_t &= u_{3t},
 \end{aligned} \tag{8}$$

where u_t is 3×1 vector of $N(0,1)$ innovations that are mutually uncorrelated. This example has one cointegrating vector and hence two permanent shocks. The results are presented in Table 2. The two permanent shocks are correctly identified, and the transitory shock has no impact on y_t and z_t as should be the case. In this DGP, 20% of the variance of x_t is due to the first permanent shock and 80% to the second. The error decomposition suggests a split of 20% and 76%, close to the true values.

Table 2
 DGP 2 ($r = 1$), $x_t = y_t + 2z_t + u_{1t}$, $\Delta y_t = u_{2t}$, $\Delta z_t = u_{3t}$

Period	P Shock 1			P Shock 2			T Shock		
	x	y	z	x	y	z	x	y	z
<i>Impulse response functions</i>									
1	1.049	1.000	0.094	2.047	0.000	1.000	1.000	0.000	0.000
2	1.136	1.000	0.094	2.031	0.000	1.000	0.363	0.000	0.000
3	1.168	1.000	0.094	2.026	0.000	1.000	0.132	0.000	0.000
4	1.179	1.000	0.094	2.024	0.000	1.000	0.048	0.000	0.000
5	1.184	1.000	0.094	2.023	0.000	1.000	0.017	0.000	0.000
6	1.185	1.000	0.094	2.023	0.000	1.000	0.006	0.000	0.000
SE_{max}	0.065	0.000	0.005	0.049	0.000	0.000	0.002	0.001	0.000
<i>Decomposition of variance</i>									
1	0.153	1.000	0.007	0.697	0.000	0.993	0.149	0.000	0.000
2	0.176	1.000	0.007	0.734	0.000	0.993	0.090	0.000	0.000
3	0.189	1.000	0.007	0.749	0.000	0.993	0.062	0.000	0.000
4	0.197	1.000	0.007	0.756	0.000	0.993	0.047	0.000	0.000
5	0.202	1.000	0.007	0.760	0.000	0.993	0.038	0.000	0.000
6	0.205	1.000	0.007	0.763	0.000	0.993	0.032	0.000	0.000

3.1. Is money superneutral?

There is a vast literature in empirical macroeconomics questioning the effects of monetary policy variables on real variables. As discussed in Fisher and Seater (1993), the ability to interpret coefficients purporting to test money neutrality crucially depends on the order of integration of the variables. In the data, $M2$ and price are $I(2)$ processes, while interest rates, output, and consumption are $I(1)$.⁸ A model consisting of the level of money, prices, interest rates, and output is an ‘unbalanced’ model in the sense that the variables are integrated of different orders. Nevertheless, one can assess whether money is superneutral. That is, if a change in the growth rate of money leaves the real variables unchanged. The issue was analyzed theoretically in Sidrauski (1967), and empirically in Geweke (1986), among others.

⁸ For the sample 1959Q1–1994Q1, the $t_\rho(k)$ statistic of Said and Dickey (1984) with the truncation lag selected as discussed in Ng and Perron (1997) cannot reject a unit root in either money growth or inflation (of the GDP deflator). The statistics, based upon $k_{max} = 10$ are -2.14 and -1.88 respectively. For FYGM3 (3 month t-bill rate) and FFYF (fed-funds rate), the unit root tests are -1.68 and -1.88 , respectively.

In this example, we present a P–T view to the issue of superneutrality using four variables: the growth rate of $M2$ (Δm_2), the Fed Funds rate (FF), the log of total consumption, and inflation (π , defined in terms of the GDP deflator). The estimation is over the sample 1959Q2 to 1994Q2. A constant and four lags are used in the VECM. The Trace and Max- λ statistics for the null hypothesis of $r = 0$ are 46.66 and 43.95, and are 20.21 and 9.34 for testing $r = 1$.⁹ The cointegration tests suggest one cointegrating vector and it is estimated by reduced rank regressions. The whole analysis is deliberately made atheoretical. Our objective is to see if we can still use economic reasoning to interpret the impulse response functions.

The cointegrating vector for $X_t = (\Delta m_2, FF, c, \pi)$ is $(1, -0.378, 0.045, -0.805)$. The coefficient on consumption is numerically small and not significant when four lags are used, but is sometimes statistically significant (at the 10% level) when we vary the lag length of the VECM. This coefficient is therefore left unconstrained, and uncertainty around it is to be resolved through the standard errors of the impulse response functions. The vector $\hat{\gamma}$ is $(-0.346, 0.090, 0.081, 0.16)'$ with t statistics of $-3.44, 2.09, 1.71, \text{ and } 2.59$. The three columns of $\hat{\gamma}_\perp$ are $[(0.225, 0.973, -0.024, -0.049)', (0.203, -0.024, 0.978, -0.044)', (0.407, -0.049, -0.044, 0.911)']$. Accordingly, none of the permanent shocks come exclusively from one variable alone.

The impulse response functions are presented in Fig. 1. The corresponding standard errors are small and not reported to conserve space. A unit increase in the first permanent shock raises the interest rate and inflation, but increases the latter by less than in proportion. The result is a permanent increase in the real interest rate. The dynamic responses are akin to those to the real interest rate shock identified by King et al. (1991), except that we have not imposed balanced growth or neutrality assumptions on the VECM. Consumption eventually falls by 1.2%, showing a large semi-elasticity of consumption to the real interest rate. Note that the short-run response of consumption to this increase in the real interest rate is small, consistent with the small interest elasticity of consumption found in estimations of Euler equations.

The second permanent shock leads to a permanent increase in consumption that is similar in shape and magnitude as that found in Cochrane's consumption–output analysis. A one percent shock raises consumption by 1.4% similar to what one commonly finds from a productivity shock. Note that money growth also increases temporarily; this can be interpreted as monetary policy accommodation to higher output induced by the productivity shock, consistent with the passive role of money suggested by some real business cycle model proponents.

⁹The critical values for cointegration tests are taken from Osterwald-Lenum (1992). The 10% values for $r = 0$ are 43.95 and 24.73 for the Trace and Max- λ statistics. For $r = 1$, the critical values are 26.79 and 18.6. These assume there is an unrestricted constant in the VECM.

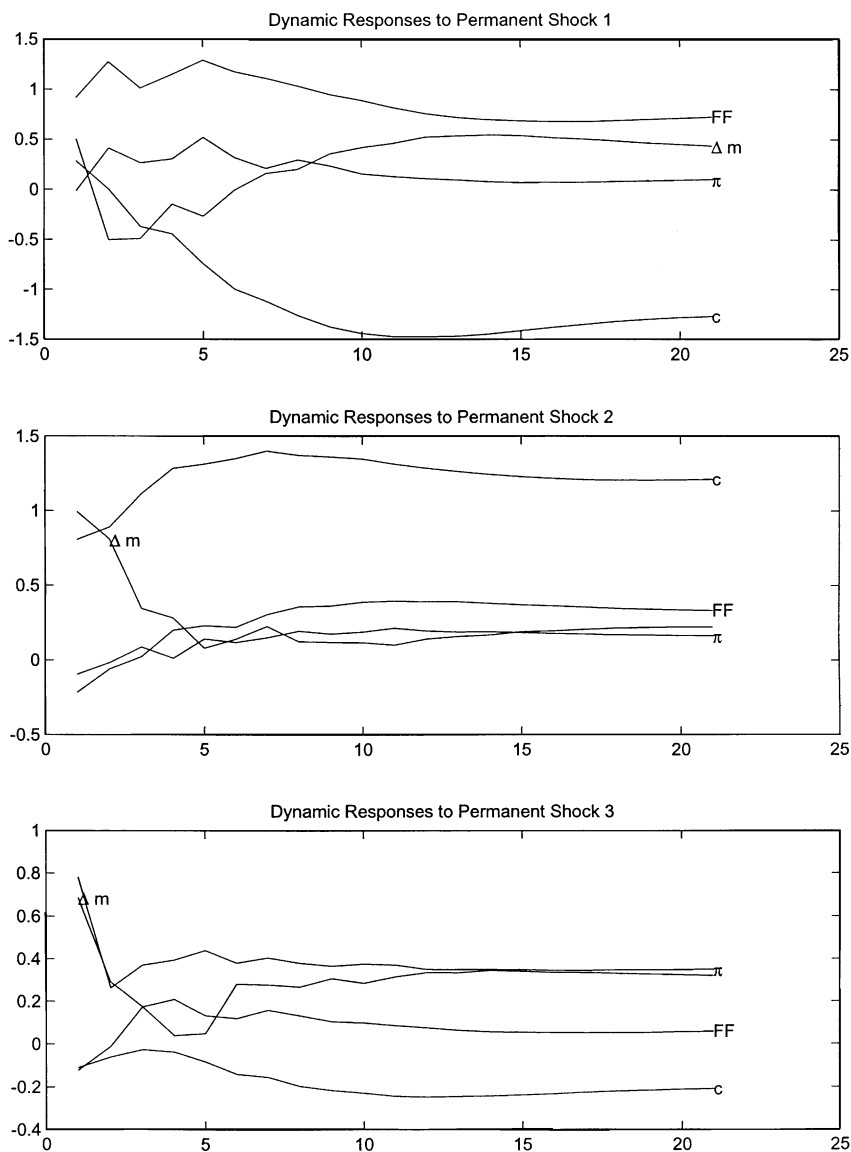


Fig. 1.

The third permanent shock raises money growth and inflation by roughly the same proportion. These responses are what one would expect of a money growth shock. The shock leads to a mild reduction in consumption of -0.2 of one percent with a maximum standard error of 0.09. Therefore, if we were to

Table 3
 Δm_2 , FF , c , π

Period	<i>P-Shock 1</i>				<i>P-Shock 2</i>			
	Δm_2	FF	c	π	Δm_2	FF	c	π
<i>Decomposition of variance</i>								
1	0.040	0.720	0.057	0.000	0.268	0.068	0.801	0.006
2	0.054	0.798	0.028	0.057	0.302	0.028	0.877	0.005
3	0.073	0.782	0.043	0.068	0.298	0.020	0.902	0.008
4	0.073	0.791	0.051	0.081	0.301	0.026	0.913	0.007
5	0.079	0.816	0.083	0.119	0.299	0.032	0.891	0.013
6	0.077	0.829	0.123	0.125	0.294	0.035	0.853	0.016
7	0.077	0.829	0.155	0.120	0.293	0.045	0.821	0.020
8	0.078	0.824	0.187	0.122	0.287	0.058	0.788	0.028
9	0.086	0.820	0.216	0.121	0.279	0.070	0.758	0.033
10	0.097	0.813	0.240	0.116	0.270	0.083	0.731	0.039
11	0.109	0.805	0.261	0.111	0.260	0.095	0.709	0.046
12	0.123	0.798	0.278	0.107	0.250	0.107	0.691	0.051
13	0.137	0.791	0.292	0.102	0.241	0.117	0.677	0.055
14	0.150	0.785	0.303	0.098	0.233	0.126	0.665	0.059
15	0.162	0.780	0.312	0.095	0.226	0.134	0.656	0.062
16	0.171	0.775	0.319	0.091	0.221	0.141	0.649	0.065
17	0.179	0.772	0.324	0.088	0.217	0.147	0.643	0.067
18	0.185	0.769	0.329	0.085	0.213	0.152	0.639	0.069
19	0.191	0.766	0.332	0.083	0.211	0.157	0.636	0.070
Period	<i>P-Shock 3</i>				<i>T-Shock</i>			
	Δm_2	FF	c	π	Δm_2	FF	c	π
1	0.228	0.039	0.026	0.736	0.464	0.172	0.116	0.258
2	0.183	0.015	0.017	0.693	0.461	0.159	0.078	0.245
3	0.177	0.032	0.010	0.710	0.451	0.167	0.045	0.214
4	0.172	0.046	0.007	0.726	0.454	0.138	0.029	0.186
5	0.170	0.042	0.006	0.711	0.452	0.110	0.020	0.157
6	0.189	0.040	0.008	0.717	0.441	0.096	0.016	0.143
7	0.204	0.042	0.010	0.730	0.427	0.084	0.014	0.129
8	0.218	0.042	0.013	0.731	0.417	0.075	0.013	0.118
9	0.234	0.041	0.015	0.736	0.401	0.069	0.012	0.110
10	0.246	0.040	0.017	0.743	0.388	0.064	0.011	0.102
11	0.259	0.039	0.019	0.747	0.373	0.060	0.010	0.096
12	0.271	0.038	0.021	0.751	0.355	0.057	0.009	0.091
13	0.282	0.037	0.023	0.755	0.339	0.055	0.008	0.087
14	0.293	0.036	0.024	0.759	0.323	0.053	0.008	0.083
15	0.303	0.035	0.025	0.764	0.309	0.051	0.007	0.080
16	0.312	0.035	0.026	0.767	0.296	0.049	0.007	0.077
17	0.320	0.034	0.026	0.771	0.285	0.047	0.006	0.074
18	0.327	0.033	0.027	0.775	0.274	0.046	0.006	0.071
19	0.334	0.033	0.027	0.778	0.265	0.044	0.006	0.068

rationalize this permanent shock as a shock to money growth, money is approximately superneutral.

Although we have not explicitly used economic theory to identify the cointegrating vectors, the results can still be rationalized by a monetary growth model with multiple shocks. The decomposition of variances reported in Table 3 suggests that one-third of the variations in consumption is due to the real interest rate shock, and two-thirds are due to innovations in productivity, with only small contributions from the money growth shock. Results using the 3 month treasury bill rate are similar.¹⁰

4. Two caveats

The procedure outlined above necessitates estimates of r , α , and γ_{\perp} . One can use the reduced rank analysis of Johansen (1988) to obtain r and α . An alternative is to use the common trend statistic of Stock and Watson (1988) to determine r , and then estimate α by fully efficient estimators. Which method (or combination of methods) to use is at the user's discretion. In this section, we stress the strong dependence of any P–T decomposition on precise and consistent estimation of these parameters.

4.1. The cointegrating rank

In practice, all permanent–transitory decompositions assume the cointegrating rank is correctly specified. Suppose the true number of cointegrating vectors is r^* , and the practitioner imposes $r \neq r^*$ cointegrating relationships. Consider a VECM with no lags, and hence $A(L)$ is of order one. The assumed $A(1)$, denoted $\tilde{A}(1)$, is related to the truth by $\tilde{A}(1) = A(1) - (I_n + \gamma\alpha' - \tilde{\gamma}\tilde{\alpha}')$, where $\tilde{\alpha}$ and $\tilde{\gamma}$ are of rank r rather than r^* . But when $r \neq r^*$, the identified shocks will not, in general, be innovations.¹¹ There will be little systematic relationship between the genuine and the identified shocks. Whether the misspecification arises from imposing too many or too few transitory shocks, incorrect dynamic response functions will result.

Consider the two simulated examples once again. Suppose instead of two cointegrating vectors in Example 1, we assume there is only one. Likewise,

¹⁰ Additional empirical examples can be found in the working paper, Gonzalo and Ng (1996).

¹¹ The problem is most easily seen in a univariate framework. Suppose $X_t = e_t$ is white noise, so that $r^* = 1$. If we assume $r = 0$, the VECM is $\Delta X_t = u_t$ but $u_t = (1 - L)e_t$ is not an innovation. Suppose, on the other hand, $\Delta X_t = e_t$ so that $r^* = 0$. If we assume $r = 1$, the VECM is $\Delta X_t = \rho X_{t-1} + u_t$, where $u_t = e_t - \rho X_{t-1}$. Although u_t is a transitory shock, it is not an innovation.

Table 4

DGP 1 ($r = 1, r^* = 2$); $\Delta x_t = u_{1t}$, $-x_t + y_t - z_t = u_{2t}$, $0.5x_t + 0.5y_t + z_t = u_{3t}$; decomposition of variances

Period	P Shock 1			P Shock 2			T Shock 1		
	x	y	z	x	y	z	x	y	z
1	1.000	0.066	0.269	0.000	0.236	0.731	0.000	0.698	0.000
2	1.000	0.099	0.269	0.000	0.432	0.731	0.000	0.469	0.000
3	1.000	0.116	0.269	0.000	0.528	0.731	0.000	0.356	0.000
4	1.000	0.126	0.269	0.000	0.587	0.731	0.000	0.287	0.000
5	1.000	0.133	0.269	0.000	0.626	0.731	0.000	0.241	0.000
6	1.000	0.138	0.269	0.000	0.655	0.731	0.000	0.207	0.000

Table 5

DGP 2 ($r = 2, r^* = 1$); $x_t = y_t + 2z_t + u_{1t}$, $\Delta y_t = u_{2t}$, $\Delta z_t = u_{3t}$; decomposition of variances

Period	P Shock 1			T Shock 1			T Shock 2		
	x	y	z	x	y	z	x	y	z
1	0.749	0.007	1.000	0.116	0.017	0.000	0.134	0.975	0.000
2	0.791	0.007	1.000	0.066	0.017	0.000	0.142	0.975	0.000
3	0.809	0.007	1.000	0.045	0.017	0.000	0.146	0.975	0.000
4	0.818	0.007	1.000	0.035	0.017	0.000	0.148	0.975	0.000
5	0.823	0.007	1.000	0.028	0.017	0.000	0.149	0.975	0.000
6	0.826	0.007	1.000	0.024	0.017	0.000	0.150	0.975	0.000

instead of two permanent shocks in Example 2, we assume one. The decomposition of variances for the misspecified models are given in Tables 4 and 5. Misspecifying the cointegrating rank apparently gives very misleading assessments of the importance of the shocks (compare with Tables 1 and 2). Not imposing the cointegrating restrictions can likewise mix up the permanent and the transitory components of the model. Essentially, the number of permanent shocks will always be $n - r$, where r is what we assume. All analysis of this nature are subject to potential misspecification of the cointegration rank.

4.2. *The estimation of γ and γ_{\perp}*

Conditional on $\hat{\alpha}$, the VECM provides an estimate of $\hat{\gamma}$. Various methods are then available for calculating γ_{\perp} . Among these (1) the eigenvectors associated with the $n - r$ smallest eigenvalues of the matrix $\gamma\gamma'$, (2) the eigenvectors associated with the smallest $n - r$ eigenvalues of the matrix $\gamma(\gamma'\gamma)^{-1}\gamma'$, and (3) the

generalized eigenvectors associated with the $n - r$ smallest eigenvalues of the matrix $S_{01}S_{11}^{-1}S_{10}$ with respect to S_{00} as proposed in Gonzalo and Granger (1995). We have experimented with all three methods. Although the non-uniqueness of γ_{\perp} makes it difficult to compare the various methods, our experience is that Method 3 is numerically less precise unless the sample size is very large, in the sense that $\gamma'_{\perp}\gamma$ is closer to zero with Methods 1 and 2. From a practical standpoint, Method 2 could be unstable when $\gamma'\gamma$ is closer to singular, a case which we cannot rule out in practice. We have therefore opted for Method 1 in our empirical analyses.

The role of γ in any P-T decomposition is crucial since γ_{\perp} defines the permanent shocks. But as discussed in Podivinsky (1992), $\hat{\gamma}$ can have poor finite sample properties. To the extent that $\hat{\gamma}_{\perp}$ depends directly on $\hat{\gamma}$, we always adopt the strategy of constraining insignificant estimates of $\hat{\gamma}$ to zero before constructing $\hat{\gamma}_{\perp}$. The rationale for this is most clearly seen from the following three variable example. Suppose the true value of γ is $(1,0,0)'$ as in DGP2 of (8). Then

$$\gamma = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \gamma'_{\perp} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This implies that there are two permanent shocks, one due solely to y_t , and one due solely to z_t . Now suppose estimation of γ yields

$$\hat{\gamma} = \begin{bmatrix} 1 \\ 0.1 \\ 0.1 \end{bmatrix} \Rightarrow \hat{\gamma}'_{\perp} = \begin{bmatrix} 0.197 & -0.985 & -0.985 \\ 3.43 & -16.68 & -17.68 \end{bmatrix}.$$

Although 0.1 seems not too far from the true value of 0, the permanent shocks now attribute rather misleading weights to the second and the third variable. The problem arises because the orthogonal complement of a matrix, say, z , is not continuous in small perturbations in z . This problem persists irrespective of how $\hat{\gamma}_{\perp}$ is constructed. Thus, $D(1)$ (which depends on $\hat{\gamma}_{\perp}$) can be very sensitive to small variations in γ .

Finally, note that the choice of the truncation order (K) for the VECM is also important. If the model is underparameterized, e_t will be serially correlated which will also affect the estimation of r and γ .

5. Conclusion

Vector autoregressions is a valuable framework for dynamic economic analyses. When some variables share common stochastic trends, the system of

variables is bind together by cointegrating restrictions. This paper shows that information on these linear relationships can be used to decompose shocks into permanent and transitory components. The analysis (i) presents the two steps necessary to obtain a permanent–transitory decomposition, and (ii) clarifies that cointegration restrictions are used in G and that it is the block in the lower triangular matrix H that is at the discretion of the practitioner. The main advantage of the proposed procedure is its simplicity. Once the permanent and transitory shocks are isolated, standard VAR identification tools can be used to make them mutually uncorrelated. However, all P–T decompositions are sensitive to assumptions on the cointegrating rank and estimation of γ_{\perp} . These are issues that practitioners should take seriously when conducting permanent–transitory decompositions of any kind.

Appendix

Proof of Proposition 1. Without loss of generality, we assume $\Gamma(L) = \Gamma_1$, so that $A(L) = I_n - A_1L - A_2L^2$, $A_1 = \Pi + I_n + \Gamma_1$, and $A_2 = -\Gamma_1$. Let $f_t = \gamma'_{\perp}X_t$ and $z_t = \alpha'X_t$. Using the result in Gonzalo and Granger (1995) that $\theta_1 = \alpha'_{\perp}(\gamma'_{\perp}\alpha_{\perp})^{-1}$ and $\theta_2 = \gamma'(\alpha'\gamma')^{-1}$, we can have

$$X_t = \theta_1f_t + \theta_2z_t = P_t + T_t. \tag{9}$$

Write the AR representation of $(\Delta f_t, z_t)$ as

$$\begin{bmatrix} I_n - (\gamma'_{\perp}\Gamma_1A_1)L & (-\gamma'_{\perp}\Gamma_1A_2)L(1-L) \\ (-\alpha'\Gamma_1A_1)L & I_n - (\alpha'\gamma - I_n + \alpha'\Gamma_1A_2)L + (\alpha'\Gamma_1A_2)L^2 \end{bmatrix} \begin{bmatrix} \Delta f_t \\ z_t \end{bmatrix} = \begin{bmatrix} \gamma'_{\perp}e_t \\ \alpha'e_t \end{bmatrix} = \begin{bmatrix} u_t^p \\ u_t^T \end{bmatrix},$$

which we write more compactly as

$$\begin{bmatrix} F_{11}(L) & F_{12}(L) \\ F_{21}(L) & F_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta f_t \\ z_t \end{bmatrix} = \begin{bmatrix} u_t^p \\ u_t^T \end{bmatrix}. \tag{10}$$

Note that $F_{12}(1) = 0$. Inverting (10),

$$\begin{bmatrix} \Delta f_t \\ z_t \end{bmatrix} = \begin{bmatrix} F^{11}(L) & F^{12}(L) \\ F^{21}(L) & F^{22}(L) \end{bmatrix} \begin{bmatrix} u_t^p \\ u_t^T \end{bmatrix}, \tag{11}$$

where the elements F^{ij} are assumed to exist and are determined by partitioned inverse. Of note is that

$$F^{12}(L) = F_{11}^{-1}(L)F_{12}(L)[F_{21}(L)F_{11}^{-1}(L)F_{12}(L) - F_{22}(L)]^{-1}.$$

Since $F_{12}(1) = 0$, we have $F^{12}(1) = 0$. Using the definition of X_t , we have the MA representation of ΔX_t as

$$\begin{bmatrix} \theta_1 & 0 \\ 0 & \theta_2(1-L) \end{bmatrix} \begin{bmatrix} \Delta f_t \\ z_t \end{bmatrix} = \begin{bmatrix} \theta_1 & 0 \\ 0 & \theta_2(1-L) \end{bmatrix} \begin{bmatrix} F^{11}(L) & F^{12}(L) \\ F^{21}(L) & F^{22}(L) \end{bmatrix} \begin{bmatrix} u_t^p \\ u_t^T \end{bmatrix},$$

and hence

$$\begin{bmatrix} \Delta P_t \\ \Delta T_t \end{bmatrix} = \begin{bmatrix} \theta_1 F^{11}(L) & \theta_1 F^{12}(L) \\ \theta_2(1-L)F^{21}(L) & \theta_2(1-L)F^{22}(L) \end{bmatrix} \begin{bmatrix} u_t^p \\ u_t^T \end{bmatrix}.$$

Thus,

$$\begin{aligned} \Delta X_t &= \Delta P_t + \Delta T_t \\ &= [\theta_1 F^{11}(L)\theta_2(1-L)F^{21}(L) + \theta_1 F^{12}(L) + \theta_2(1-L)F^{22}(L)] \begin{bmatrix} u_t^p \\ u_t^T \end{bmatrix} \end{aligned} \tag{12}$$

$$= \begin{bmatrix} D_{11}(L) & D_{12}(L) \\ D_{21}(L) & D_{22}(L) \end{bmatrix} \begin{bmatrix} u_t^p \\ u_t^T \end{bmatrix}. \tag{13}$$

and note that $D_{12}(1) = D_{22}(1) = 0$, because $\theta_1 F^{12}(1) + \theta_2 \times 0 \times F^{22}(1) = 0$.

References

- Bernanke, B., 1986. Alternative explanations of the money-income correlation. Carnegie Rochester Conference on Public Policy 25, 49–100.
- Bernanke, B.S., Mihov, I., 1995. Measuring Monetary Policy. Princeton University, Princeton, NJ.
- Blanchard, O., Quah, D., 1993. The dynamic effects of aggregate demand and supply: reply. American Economic Review 83, 653–658.
- Cochrane, J.H., 1994a. Permanent and transitory components of GNP and stock prices. Quarterly Journal of Economics CIX, 241–266.
- Cochrane, J.H., 1994b. Shocks. NBER working paper No. 4689.
- Engle, R.F., Granger, C., 1987. Cointegration and error-correction: representation, estimation, and testing. Econometrica 55, 251–276.
- Evans, G., Reichlin, L., 1994. Information, forecasts, and measurement of the business cycle. Journal of Monetary Economics 33, 233–254.
- Fisher, M., Seater, J., 1993. Long-run neutrality and superneutrality in an ARIMA framework. American Economic Review 83, 402–415.

- Geweke, J., 1986. The superneutrality of money in the United States; An interpretation of the evidence. *Econometrica* 54, 54–121.
- Gonzalo, J., Granger, C., 1995. Estimation of common long-memory components in cointegrated systems. *Journal of Business and Economic Statistics* 13, 27–35.
- Gonzalo, J., Ng, S., 1996. A systematic framework for analyzing the dynamic effects of permanent and transitory shocks, C.R.D.E. working paper 0396, University of Montreal.
- 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 vectors in Gaussian vector autoregressive models. *Econometrica* 59, 1551–1580.
- Johansen, S., 1995. Identifying restrictions of linear equations with applications to simultaneous equations and cointegration. *Journal of Econometrics* 69, 111–132.
- King, R.G., Plosser, C., Stock, J., Watson, M., 1991. Stochastic trends and economic fluctuations. *American Economic Review* 81 (4), 819–840.
- Koray, F., Lee, T.H., Palivos, T., 1995. Stochastic trends and fluctuations in national income, wages, and profits. Mimeo, Louisiana State University.
- Lippi, M., Reichlin, L., 1993. The dynamic effects of aggregate demand and supply disturbances: comment. *American Economic Review* 83, 644–652.
- Ng, S., Perron, P., 1997. Estimation and inference in nearly unbalanced and nearly cointegrated systems. *Journal of Econometrics* 79, 54–81.
- Osterwald-Lenum, M., 1992. A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics* 54, 461–472.
- Podivinsky, J.M., 1992. Small sample properties of tests of linear restrictions on cointegrating vectors and their weights. *Economic Letters* 39, 13–18.
- Runkle, D.E., 1987. Vector autoregressions and reality. *Journal of Business and Economic Statistics* 5, 437–443.
- Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71, 599–607.
- Sidrauski, M., 1967. Rational choice and patterns of growth in a monetary economy. *American Economic Review Papers and Proceedings*, pp. 534–544.
- Stock, J.H., Watson, M.W., 1988. Testing for common trends. *Journal of the American Statistical Association* 83, 1097–1107.
- Warne, A., 1991. A common trends model: identification, estimation and asymptotics. Mimeo, Stockholm School of Economics.
- Watson, M.W., 1994. Vector autoregressions and cointegration. North-Holland, Amsterdam, *Handbook of Econometrics*, Vol. 4. pp. 2844–2910.