

CAHIER 9603

**A SYSTEMATIC FRAMEWORK FOR ANALYZING THE DYNAMIC EFFECTS OF
PERMANENT AND TRANSITORY SHOCKS**

Jesus GONZALO¹ and Serena NG²

¹ Boston University and Universidad Carlos III de Madrid

² Département de sciences économiques and Centre de recherche et développement
en économique (C.R.D.E.), Université de Montréal

March 1996

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 the seminar participants for helpful discussions and Paul Beaudry and Pierre Perron for comments on an earlier draft. The first author thanks the Fundación Caja de Madrid for financial support. The second author acknowledges grants from the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Fonds pour la formation de chercheurs et l'aide à la recherche (FCAR) of Québec.

C.P. 6128, succursale Centre-ville
Montréal (Québec) H3C 3J7

Télécopieur (FAX) : (514) 343-5831
Courrier électronique (E-Mail) : econo@ere.umontreal.ca

RÉSUMÉ

Dans cet article, nous proposons une approche systématique pour l'analyse des effets dynamiques des chocs transitoires et permanents sur un système de n variables économiques. Nous considérons une méthode d'orthogonalisation en deux étapes des résidus d'un modèle vectoriel à correction d'erreur (VECM) avec r vecteurs de cointégration : la première sépare les chocs permanents des chocs transitoires et la seconde isole $n - r$ chocs permanents mutuellement indépendants et r chocs transitoires. L'approche proposée exploite la relation de cointégration suggérée par les données et ne tient pas compte des restrictions théoriques sous-jacentes. Nous considérons que ces dernières ne sont pas nécessaires, même si nous pouvons les utiliser. Nous montrons aussi comment les fonctions de réaction peuvent être construites pour mettre en évidence le mécanisme de propagation des chocs en fonction de leur degré de persistance. Cette méthode, en distinguant les chocs selon leur origine, diffère de l'approche standard et peut être considérée comme une voie complémentaire pour analyser les dynamiques macroéconomiques.

Mots clés : VAR, fonctions de réaction, cointégration

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 independent permanent shocks and r transitory shocks. The approach exploits the cointegrating relationships in the data. Although theoretical restrictions can be used, they are not necessary. We also show how impulse response functions can be constructed to trace out the propagating mechanism of shocks distinguished by their degree of persistence. This differs from the common approach of distinguishing shocks by their origin, and hence offers a complementary way of analyzing macroeconomic dynamics.

Key words : VAR, impulse responses, cointegration

Ce cahier a également été publié au Centre de recherche et développement en économique (C.R.D.E.) (publication no 0396).

Dépôt légal - 1996
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

ISSN 0708-9231

1 Introduction

Many issues of economic interest evolve around two themes: - are shocks to the system permanent or transitory, and are some shocks common to the variables under investigation. Issues of such nature include, do asset prices share a common trend; are fluctuations across economies the result of common shocks; does consumption respond to increases in permanent and transitory income differently, will there be convergence of economies at different stages of economic development; are variations in macroeconomic variables due to transitory or transitional dynamics, and how long do adjustments to these shocks take. The two themes are evidently not unrelated, as permanent shocks are necessarily associated with a trend component, and the existence of common trends implies that there must be permanent shocks which affect macroeconomic aggregates simultaneously.

In this paper, we propose a simple and coherent framework for isolating the transitory and the permanent shocks from a system of integrated variables, making explicit the relationship between the common trends and the innovations underlying the reduced form model. We then discuss 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. Our framework is set up so that the permanent and transitory shocks can be isolated in an atheoretical way. Although structural assumptions can be used, they are not necessary.

Much has been written about the ability of cointegrating relationships in enhancing our understanding of long-run economic relationships. In this paper, we consider a framework which makes use of the cointegration restrictions to provide a description of the data that highlights the permanent and the transitory nature of the shocks. The procedure consists of two steps. The first 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 (though not explicit) from the VECM. The analysis then proceeds in the same way as a stationary VAR, using methods such as Choleski decomposition to complete the exercise of innovation accounting. The advantage of the method is its simplicity, since all the required ingredients are contained in the VECM. We use examples to illustrate the information that can be revealed by the permanent-transitory decomposition.

Numerous studies have devised ways to orthogonalize shocks into permanent and transitory components, and to decompose the level of a series into trends and cycles, but the two issues are often treated in separate contexts.¹ Permanent and transitory shocks, sometimes labelled supply

¹With a slight abuse of terminology, we shall use the terms permanent and transitory interchangeably with trends and cycles. Hence, white noise series will also be referred to as cycles.

and demand shocks, are usually identified by imposing structural assumptions about the degree of persistence of the innovations on certain variables. In most cases, explicit derivations of the trend functions are bypassed. The work of Blanchard and Quah (1989), Shapiro and Watson (1988), and Quah and Valley (1995) falls into this category. Studies which focus on the trend-cycle decomposition usually take the reduced form of the data as the starting point, and as such, do not have to be explicit about the economic structure that generates the data. The work of Stock and Watson (1988) and Gonzalo and Granger (1995) falls into this category. Identifying the permanent shocks without taking into account the presence of common trends could imply misleading dynamic responses, and information is not fully exploited if we identify the common trends without analyzing the dynamic responses to the innovations which drive those trends. The framework used in this paper takes into account that the two issues are intertwined. The permanent shocks that we back out are explicitly tied to the (possibly common) trends underlying the data.

A VAR in first differences is misspecified if there are non-zero cointegrating relationships, and a VAR in level form is inefficient relative to a VECM if it ignores the long-run relationships. The advantages of imposing the cointegrating restrictions from an estimation standpoint are by now well known. Lesser known are the advantages and the disadvantages of imposing cointegrating restrictions on a VAR from the viewpoint of impulse response analysis. The relative merits are not immediately obvious as we need to entertain the possibility that the cointegration restrictions might not be correctly estimated. Our analysis also sheds light on this issue.

In VAR analyses, it is conventional to give the orthogonalized residuals (achieved by Choleski or structural decomposition) labels that define the source of the shock. This necessitates an implicit or explicit view about the structure underlying the variables under investigation. In consequence, the properties of the shocks we identify are sensitive to assumptions we impose. A case in point is a VAR consisting of money, prices, interest rate and output. As clearly explicated in Bernanke and Mihov (1995), the monetary policy shocks we back out will depend on our view about the monetary policy transmission mechanism. The problem arises because economic variables tend to move together, and we do not have a "big bang" type of theory that tells us decisively which variables move first. As forcefully argued in Cochrane (1994b), after decades of analysis, we still know very little and perhaps will never know enough about the source of the shock.

A key feature of cointegrated systems is that the variables move together at low frequencies. Given this coherence, the ability to identify the shocks by names is even more limited. However, it is also by exploiting these comovements that allows us to decompose the regression residuals into shocks that have permanent effects from those that have transitory effects only. Thus, in our analysis, shocks are distinguished by their degree of persistence, rather than with names of the

variables in the system. This permanent-transitory decomposition can be accomplished without appeal to economic theory. When this is the case, the decomposition should be seen as an atheoretical toolkit for describing the data. In order to give tight economic interpretation to the results, we would need to be specific about the structure of the model to be identified. As will be discussed below, such structural restrictions can be easily adapted in our framework. However, because our decomposition permits but does not require us to take a stand on the economic structure, the results should be interpreted in the proper context.

The plan of this paper is as follows. The econometric framework used to isolate the permanent and transitory shocks are presented in the next section. Section 3 focuses on the construction of impulse response functions. Section 4 puts into context our decomposition with related work in the literature, and applies our method to two artificial examples. Section 5 applies the method to three empirical examples. Practical issues are deferred till Section 6, where we consider the sensitivity of results to the cointegrating rank and to other parameter estimates. 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 . Throughout, we shall focus on the detrended series X_t , which has a multivariate moving-average representation

$$\Delta X_t = C(L)e_t \quad (1)$$

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

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

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

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.

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 $C(1)\gamma = 0$. The columns of α are the cointegrating vectors of X_t . Furthermore, X_t has a VECM representation:

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

where $\Gamma(L)$ is a polynomial of finite order $K-1$. The VECM can be taken to be the true DGP for ΔX_t , or a finite order approximation to the true model. The term $\alpha' X_{t-1}$ is the equilibrium error which generates the "correction" necessary to ensure that the variables will return to the desired long-run levels. Note that because we work with X_t instead of \bar{X}_t , deterministic components lie outside the cointegrating space in the sense of Johansen (1991). In other words, (2) expressed in terms of \bar{X}_t would include a set of (differenced) deterministic terms with unrestricted coefficients. If it is desirable to include a constant (and/or a trend) in the cointegrating space, we should work with \bar{X}_t , and augment the error-correction term by deterministic components with appropriate adjustments to the dimensions of α and γ .

The VECM can be used to deduce a restricted VAR representation:

$$A(L)X_t = \varepsilon_t, \quad (3)$$

the restriction being $A(1) = \gamma \alpha'$. The rank of $A(1)$ is r , which may or may not coincide with that corresponding to an unrestricted VAR which ignores the cointegrating relationships.

2.2 The Permanent and Transitory (P-T) Decomposition

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 sequence of shocks $\tilde{\eta}_t^T$ is said to have permanent effects on the level of X_t if $\lim_{h \rightarrow \infty} \partial E_t(X_{t+h}) / \partial \tilde{\eta}_t^T \neq 0$. Analogously, the sequence of shocks $\tilde{\eta}_t^T$ is said to have transitory effects on the level of X_t if $\lim_{h \rightarrow \infty} \partial E_t(X_{t+h}) / \partial \tilde{\eta}_t^T = 0$.

The elements in the vector ε_t in (1) to (3) are the innovations to X_t . Since the variables in X_t are $I(1)$ by definition, some of these innovations, or combinations of them, must have a lasting effect on the level of X_t . This suggests ΔX_t should have the following representation:

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

where $\bar{D}(L) = \bar{D}_0 + \bar{D}_1 L + \bar{D}_2 L^2 \dots$. The covariance matrix of $\tilde{\eta}_t$ is $\Sigma_{\tilde{\eta}_t}$, and only a subset of $\tilde{\eta}_t$ has permanent effects. Without loss of generality, the $\tilde{\eta}_t$'s are ordered such that the first $n-r$ of them have permanent effects. We also assume that $\tilde{\eta}_t$ is an $n \times 1$ vector, so that the number of shocks equals the number of variables in the system.³ Together with the assumption that $n \neq r$ and $r \neq 0$, we essentially assume that there is at least one permanent shock and one transitory shock. We therefore rule out cases where all shocks are permanent and yet cointegration exists.

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 expressed in terms of a set of "nonorthogonalized" permanent and transitory shocks, which we denote u_t . We shall refer to this problem as the P-T decomposition.

It is natural to assume that shocks that have permanent effects are those associated with the trend components, and shocks that have transitory effects are associated with the stationary components. Suppose the vector X_t is comprised of stationary and non-stationary latent factors. Let z_t be the stationary factors, and suppose that the common permanent factors, f_t , are linear in X_t . Then Gonzalo and Granger (1996) showed that a trend-cycle decomposition for X_t exists, and is given by

$$X_t = \theta_1 f_t + \theta_2 z_t, \quad (5)$$

$$f_t = \gamma_L' X_t, \quad z_t = \alpha' X_t, \quad (6)$$

where γ_L is $(n-r) \times n$, $\gamma_L' \gamma = 0$. The factor loading matrices are $\theta_1 = \alpha_L (\gamma_L' \alpha_L)^{-1}$ and $\theta_2 = \gamma (\alpha' \gamma)^{-1}$. These are evidently functions of α, γ , their orthogonal complements and no other parameters. The existence of r cointegrating vectors implies that there are $(n-r)$ unit roots amongst X_t . In this context, f_t are the $(n-r)$ processes in the system with stochastic trends.

An alternative trend-cycle decomposition is due to Stock and Watson (1988). Recursive substitution of (1) and using the factorization $C(L) = C(1) + (1-L)C^*(L)$ gives

$$X_t = X_0 + C(1) \sum_{s=1}^t \varepsilon_s + C^*(L) \varepsilon_t. \quad (7)$$

Because $C(1)$ has reduced rank, $C(1) \sum_{s=1}^t \varepsilon_s$ represents the $n-r$ common trends amongst X_t . The above representation is the multivariate extension of the Beveridge and Nelson (1981) decomposition.⁴

While the permanent components from the two decompositions have different dynamics, the following Lemma states that they are driven by the same random walks. In fact, the result applies to any other P-T decompositions.

³This ensures that $\bar{D}(L)$ is invertible. See Watson (1994) for a discussion of invertibility.

⁴Warne (1991) and Proietti (1995) discussed other ways of rewriting this common trend representation.

Lemma 1 *The permanent components identified from one permanent-transitory decomposition of X_t are driven by random walks that are linearly dependent of the random walks underlying the permanent components of alternative decompositions.*

The proof, given in the Appendix, essentially uses the argument that two sets of random walks purporting to explain the same variables must, by definition, be cointegrated. This in turn restricts the random walks to be linearly dependent. Thus, the innovations which induce the permanent components must also be linearly related. In other words, the permanent shocks which underlie different decompositions of X_t must span the same space. The effect of a given shock from one decomposition will be the same, whether it is analyzed in the common trend or the Gonzalo-Granger representation of X_t . It follows that either can be used as the basis for backing out the permanent innovations. Our motivation for using the Granger-Gonzalo decomposition as the starting point is based on the practical consideration that all the information necessary for the P-T decomposition can be obtained from estimation of a VECM. The method is simple and is made explicit in the following Proposition.

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 such that $C(1)$ is of rank $(n-r)$. Let*

$$G = \begin{bmatrix} \gamma'_1 \\ \alpha' \end{bmatrix} \begin{matrix} (n-r) \times n \\ r \times n \end{matrix}$$

with $\gamma'_1 \gamma = 0$. Then the permanent and transitory shocks to the n variate system are defined by Ge_t , where for each t , the $(n-r) \times 1$ vector $u_t^p = \gamma'_1 e_t$ 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 $(\gamma'_1 \alpha')$ is non-singular.

The proof to the Proposition is given in the Appendix. The decomposition fails to exist only if the rows of γ'_1 are linearly dependent of the rows of α' . However, such a case is not of economic interest as it implies that one or more permanent shocks are scalar multiples of the transitory shocks. Ruling these cases out does not restrict the scope of the analysis.

The choice of the matrix G defined in Proposition 1 is due to two results from the properties of the VECM [see (2)] and the Wold decomposition [see (1)] of ΔX_t as implied by the Granger Representation Theorem. In effect, γ'_1 and α' "knock out" terms in the different representations of ΔX_t . Indeed, the interpretation of u_t^p and u_t^t as the permanent and the transitory shocks is immediate in view of the stationary and non-stationary components defined by (6). The first differences of the $(n-r)$ unit roots f_t are innovations to the trend components, hence the permanent

shocks. By contrast, the first differences of the stationary variates z_t are overdifferentenced. These have no effects on ΔX_t in the long-run, and are therefore the transitory shocks to the system. This leads to the following:

Corollary 1 *Suppose there are r cointegrating vectors in a system of n $I(1)$ variables, and let $G = (\gamma'_1 \alpha')$ with $\gamma'_1 \gamma = 0$. Let $u_t^p = \gamma'_1 e_t$ and $u_t^t = \alpha' e_t$ define the permanent and transitory innovations respectively. Then the moving-average representation of ΔX_t in terms of $u_t = (u_t^p u_t^t)'$*

$$\Delta X_t = C(L)G^{-1}Ge_t = D(L)u_t = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \cdot \begin{bmatrix} u_t^p \\ u_t^t \end{bmatrix} \quad (8)$$

with $D_{12}(1) = 0_{(n-r) \times r}$, and $D_{22}(1) = 0_{r \times r}$.

The last r columns of the $n \times n$ polynomial matrix $D(L)$ are coefficients pertaining to the response by ΔX_t to the transitory shocks. The requirement that $D_{12}(1) = D_{22}(1) = 0$ puts in mathematical terms that the transitory shocks have no effects on the first difference or the level of X_t in the long-term.

If there is no cointegration, the matrix G has no role for α . All shocks are therefore permanent and are determined by $\gamma'_1 e_t$, where $\gamma'_1 = I_n$. When $r = n$, all variables are stationary. In this case, there is no role for γ'_1 . All the shocks are transitory. Since $\pi = \gamma\alpha'$ is full rank, any π will make πX_{t-1} stationary, the VECM (2) reduces to an unrestricted VAR in level form. Thus, a P-T decomposition can trivially be obtained in the degenerate cases which we have ruled out, i.e. when there are n stationary variables or n distinct unit roots.

2.3 The P-P and the T-T Decompositions

As is well known, the innovations associated with a stationary VAR are linear combinations of the structural shocks in the system, and therefore do not bear meaningful economic interpretation. This has motivated the use of triangular and structural decompositions to obtain the orthogonalized shocks. The G matrix rotates e_t so that it can be decomposed into u_t^p and u_t^t . However, these shocks are not mutually uncorrelated. The question we are faced with is, given these shocks and their moving average representation $\Delta X_t = D(L)u_t$, can we transform the model to $\Delta X_t = \bar{D}(L)\bar{\eta}_t$ such that we can interpret the dynamic response to $\bar{\eta}_t$. Comparing the two representations, we have the following:

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

Evidently, the data contain fewer pieces of information than the number of unknowns in the model to be recovered. We therefore need to impose restrictions on $\tilde{D}_0, \tilde{D}(1)$, and $\tilde{\eta}$ in order to have a unique relationship between the two models. We now introduce the following assumption:

Assumption 1 *The $\tilde{\eta} = (\tilde{\eta}'_i)'$ are mutually uncorrelated and have unit variances.*

The assumption essentially imposes $n(n-1)/2$ zero restrictions on the off-diagonals and n restrictions on the diagonal elements of $\text{cov}(\tilde{\eta})$. This leaves $n(n-1)/2$ number of restrictions on \tilde{D}_0 and $\tilde{D}(1)$. The following Proposition suggests that many of these restrictions are dictated by the distinct properties of the permanent and the transitory shocks.

Proposition 2 *The matrix $D_0^{-1}\tilde{D}_0$ must be lower block triangular. Let H be the square root matrix of Σ_u satisfying $HH' = \Sigma_u$. Then H is also lower block triangular. Furthermore, we have*

$$R1': \quad \tilde{D}_0 = D_0H,$$

$$R2': \quad \tilde{D}(1) = D(1)H \Rightarrow \begin{bmatrix} \tilde{D}_{11}(1) \\ \tilde{D}_{12}(1) \end{bmatrix} = \begin{bmatrix} D_{11}(1) \\ D_{12}(1) \end{bmatrix} H_{11}.$$

From (R1) and (R2), we have $D(1)D_0^{-1}\tilde{D}_0 = \tilde{D}(1)$. Note that $D(1)$ and $\tilde{D}(1)$ must have the last r columns equal to zero for the transitory shocks not to have permanent effects. This in turn restricts $D_0^{-1}\tilde{D}_0$ to be a lower block triangular matrix, proving the first part of the proposition. From (R1), $D_0HH'D_0 = \tilde{D}_0\tilde{D}_0$ implies $\tilde{D}_0 = D_0H$. Hence $D_0^{-1}\tilde{D}_0 = H$ is also a lower block triangular matrix, and is alternatively represented by (R1'). Condition (R2') follows from the specific properties of $D(1)$ and $\tilde{D}(1)$, and it is clear that the submatrix H_{11} relates the long-run properties of $D(L)$ to that of $\tilde{D}(L)$. The lower block-triangularity of H in turn allows us to solve for an additional $r(n-r)$ coefficients in \tilde{D}_0 and $\tilde{D}(1)$.

Note that when there is only one permanent and one transitory shock, the two shocks are automatically identified by the P-T decomposition alone. In general, identification of the permanent and transitory shocks requires $(n-r)(n-r-1)/2 + r(r-1)/2$ restrictions on $\tilde{D}(1)$ and \tilde{D}_0 . When these cases are encountered in our applications, we identify the shocks as follows:

A Practical Rule 1 *Let H be a lower triangular orthogonal matrix, or 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 as required. Trivially, Proposition 2 which requires lower block triangularity is also satisfied. The practical appeal of the lower triangularity of H is that, by standard arguments of error variance decomposition, the residuals $u_{jt} - E[u_{jt}|u_{1t}, \dots, u_{j-1,t}]$ are orthogonal to u_{it} , $i < j$.

The Choleski decomposition produces a set of these residuals that are mutually uncorrelated and hence satisfy Assumption 1 by construction. The structure induced by this practical rule is easily recovered from $R1'$ and $R2'$, and will be application specific.

To gain more insight into the role of H , suppose we have a set of residuals \tilde{u}_t that are uncorrelated between blocks. This could be obtained, say, by letting $\tilde{u}_t^p = u_t^p$ and \tilde{u}_t^t be the residuals from a projection of u_t^t on u_t^p . The remaining task is to make \tilde{u}_t^p mutually uncorrelated, and likewise for \tilde{u}_t^t . It is then clear that the lower triangularity of H_{11} orthogonalizes \tilde{u}_t^p by putting $(n-r)(n-r-1)/2$ restrictions, and H_{22} orthogonalizes \tilde{u}_t^t through $r(r-1)/2$ restrictions. The terminology P-P and T-T decomposition is motivated by the consideration that the two decompositions can in fact be implemented independently.

Choosing an H that is lower triangular implies that $\tilde{\eta}_t$ will be solved recursively from u_t according to the ordering of u_t . Since $u_t = Ge_t$, the ordering of the variables in the system will also affect the $\tilde{\eta}_t$ that are being identified (except in the bivariate case when $n-r = r = 1$). However, this does not imply that the residuals from the second equation, say, will not influence the variable ordered before it. The reason is that $D_0 = G^{-1}$ is not an identity matrix, which is usually the case in VARs. Hence, the residual of the first equation ϵ_{1t} can have a non-zero weight in the second permanent shock via γ_{12} . Furthermore, there could be a contemporaneous response by variable X_i to permanent shock j , where $j > i$. The lower triangularity of H does not impose a recursive causal structure on the system as in a standard VAR.

In other cases when economic theory imposes a particular structure for $\tilde{D}(1)$, we would need to solve for the square root matrix H from $\tilde{D}(1) = D(1)H$. This amounts to putting restrictions on both $D(1)$ and H , subject to the condition that H must be lower block triangular. This would be in the spirit of structural identification proposed by Bernanke (1986), and is essentially implemented in King, Plosser, Stock and Watson (1991) [hereafter, KPSW] and Koxay, Lee and Paliivos (1995). In both applications, H_{11} is chosen to be lower block triangular.

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

$$\begin{aligned} \Delta X_t &= C(L)G^{-1}HH^{-1}Ge_t \\ &= D(L)HH^{-1}u_t \\ &= D^*(L)\eta_t \end{aligned} \tag{9}$$

with $D_{12}^*(1) = D_{22}^*(1) = 0$. 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 $\hat{\gamma}_L$.

where we recall that Ω is $\text{cov}(\hat{\epsilon})$, $\hat{\epsilon}$ is the vector of residuals from estimating the VECM. Thus, $\text{MSE}(h)$ is the mean squared error matrix of the optimal h -step ahead forecast of X_t .

3.1 Standard Errors of the Impulse Response Functions

Commonly reported in economic analyses is the cumulated impulse response functions, Φ , and it is necessary to judge the precision of these estimates. Let \hat{A} be the estimates of the VECM in autoregressive form, and let $\sigma = \text{vec}(\Sigma_u)$, with $\Sigma_u = T^{-1} \sum_{i=1}^T \hat{u}_i \hat{u}_i'$. As shown in Engle and Granger (1987), $\sqrt{T}(\hat{A} - A) \rightarrow N(0, \Sigma_A)$, and $\sqrt{T}(\hat{\sigma} - \sigma) \rightarrow N(0, \Sigma_\sigma)$. Extending formula (3.7.9) in Lutkepohl (1993), we have

$$\sqrt{T}(\text{vec}(\hat{\Theta}_1) - \text{vec}(\Theta_1)) \rightarrow N(0, B_n \Sigma_A B_n' + \hat{B}_n \Sigma_\sigma \hat{B}_n'), \quad (13)$$

where $B_n = (P' \otimes I_n) F_1$, $F_1 = \partial \text{vec}(\Phi_1) / \partial \text{vec}(A)'$, $\hat{B}_n = (I_n \otimes (\Phi_1)) \partial \text{vec}(H) / \partial \sigma$, and H is the lower triangular Choleski factorization of Σ_u .

The standard errors for the impulse response functions based on one-step orthogonalization of Ω are given in Lutkepohl and Reimers (1992). These are analytically complex expressions because both the long run relationships and the short-run dynamics affect the impulse response functions. As seen from (13), the standard errors of the impulse responses following the P-T decomposition are functions of $\Sigma_u = G\Omega G'$, and its lower triangular factorization H . Variability in $\hat{\alpha}$ and $\hat{\gamma}$ therefore has additional effects on the standard errors via the first step orthogonalization. The resulting expression, as given in (13), is analytically complex, and we therefore resort to numerical methods in practice.

There are at least two ways of obtaining standard errors for the impulse response functions [see Section 11.7 of Hamilton (1994)]. One is to bootstrap from the asymptotic distribution for the parameters of the VECM. This method does not seem suited for our purpose because the parameters $\hat{\gamma}$ and $\hat{\Gamma}(L)$ are conditional on $\hat{\alpha}$. Even though $\hat{\alpha}$ is superconsistent, bootstrapping the former parameters will ignore the finite sample variations induced by $\hat{\alpha}$. This consideration precludes use of the (Bayesian) method as implemented in RATS and variants of it discussed in some detail in Sims and Zha (1994).

We opted for 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 $\hat{\epsilon}_t$. A new sample of data is constructed (using the initial estimates of $\hat{\alpha}$, $\hat{\gamma}$, $\hat{\Gamma}(L)$) by random sampling of $\hat{\epsilon}_t$ 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 error from the N

2. Construct $G = (\hat{\gamma}_1' \hat{\alpha}')'$ and the set of permanent and transitory shocks, $G\epsilon_t$;

3. Obtain a lower block triangular matrix H , such as by applying Choleski-Decomposition to $\text{cov}(G\epsilon)$. The orthogonalized permanent and transitory shocks are $\eta_t = H^{-1}G\epsilon_t$. These have unit variances and are mutually uncorrelated.

4. Post-multiply $D(L) = C(L)G^{-1}$ by H to obtain $D^*(L)$.

The resulting $D^*(L)$ matrix is the sample analog of the $\tilde{D}(L)$ matrix, and η_t is the sample analog of $\tilde{\eta}_t$. This completes the P-T decomposition.

3 Impulse Response Functions and the Decomposition of Variance

The impulse response of ΔX_t to the shocks are implied by $D^*(L)$, which in turn depend on $C(L)$. Since $A(L)C(L) = (I_n - L)$, a recursion formula can be used to obtain $C(L)$ given estimates of $A(L)$ as implied by the VECM in (2) of order $K - 1$. However, it is often more convenient to interpret changes in the levels of X_t in response to the shocks. This can be easily constructed from the partial sums of the impulse responses. We denote these cumulated impulse responses by Φ_t . It can be shown that [see, for example, Lutkepohl and Reimers (1992)]:

$$\Phi_t = (\Phi_{ij,t}) = \sum_{j=0}^t C_j = \sum_{j=1}^t \Phi_{t-j} A_j, \quad t = 1, 2, \dots, \quad (10)$$

where $\Phi_0 = I_n$, $A_j = 0$ for $j > k$. Thus, $\Phi_{ij,t}$ is the change in the level of X_i in response to a unit increase in the j^{th} element of ϵ_t , t periods hence. Once we have $\Phi(L)$, the impulse responses to the orthogonalized permanent and transitory shocks can be defined as $\Theta_t = \Phi_t G^{-1} H$. This gives $\sum_{j=0}^t D_j^*$. Its first $(n - r)$ columns are the impulse responses of X_t to a standard deviation increase in the permanent shocks, and the final r columns are the impulse responses to a standard deviation increase in the transitory shocks. Dividing these impulse response functions by the diagonal elements of H give the impulse response functions to shocks of one unit.

The decomposition of variance provides an assessment of the relative importance of permanent and transitory shocks in h step ahead forecasts. This is given by

$$\omega_{k_j,h} = \sum_{i=0}^{h-1} \theta_{k_j,i}^2 / \text{MSE}_k(h), \quad h = 1, 2, \dots \quad (11)$$

where $\theta_{k_j,i}$ is the k_j^{th} element of Θ_i , the orthogonalized impulse response of variable k to shock j , and $\text{MSE}_k(h)$ is the k^{th} diagonal element of

$$\text{MSE}(h) = \Omega + \sum_{i=1}^{h-1} \Phi_i \Omega \Phi_i', \quad (12)$$

samples of the bootstrapped impulse response functions. In practice, N is set to 1000. Admittedly, the procedure is time consuming especially when the n (the dimension of the model) and/or K (the lag length) is large. Development of an analytically and numerically more efficient method for obtaining IRFs in such contexts is worthy of further investigation.

4 Comparison with Alternative Procedures and Simulated Examples

Structural assumptions have on occasions been imposed on VARs to identify permanent and transitory shocks. For example, demand shocks are assumed to have no lasting effects on the real variables in Blanchard and Quah (1989). Other examples include the work of Shapiro and Watson (1988), and Quah and Vahey (1995). Starting with the moving-average representation $C(L)e_t$ for two variables, these studies can be seen as isolating $n - r$ unorthogonalized permanent shocks by finding a matrix, say, Q , such that $C(1)Q$ has a north-east element equal to zero. The unorthogonalized permanent and transitory shocks are therefore $Q^{-1}e_t$. Since only one of the two variables is non-stationary in their analysis, the issue of cointegration is not relevant. However, if their analysis was extended to include more variables and some of which are cointegrated, they would need to find a Q matrix such that $C(1)Q$ has the last r columns equal to zero. Then our $D(1)$ matrix is analogous to the $C(1)Q$ matrix of Blanchard and Quah.

The analysis by KPSW also provides a way of isolating the permanent shocks. Their procedure can be cast in terms of the common-trend representation (7). They find the matrix $Q = (Q_1, Q_2)$ such that X_t can be decomposed into $X_t = X_0 + A\tau_t + a_t$, where $A = C(1)Q_1$, and $\tau_t = \sum_{s=1}^t \epsilon_s$ contains the $(n - r)$ permanent components of X_t . The permanent shocks are the innovations underlying $\gamma_L \tau_t$, which are then made mutually uncorrelated upon imposing a lower triangular structure on the matrix of long-run multipliers. Thus, our $D(1)$ matrix plays the same role as their $C(1)Q_1$ matrix, and lower triangularity of our H matrix incorporates the same identifying restrictions as in KPSW. While KPSW make a priori assumptions on the number of permanent shocks, we pretest for the cointegrating rank.

Since $C(1) = \alpha_L(\gamma_L \psi \alpha_L)^{-1} \gamma_L$, the permanent innovations in KPSW's analysis evidently lie in the space spanned by γ_L , the same as our permanent innovations. By Lemma 1, this means that the long-run elasticities with respect to the orthogonalized permanent shocks from the two decompositions will be identical provided we use the same identifying restrictions. This is the case because both KPSW and our decomposition assume H_{11} is lower triangular. The fundamental difference between the approach of KPSW and ours is that their matrix Q is based upon economic theory. For example, KPSW assume balanced growth between consumption and investment, and that the inflation shock has no real effects in the long-run. These assumptions implicitly impose

restrictions on α and zero restrictions on γ . The G matrix which induces $D(1)$ in our analysis is based on long-run properties of the data as determined by the estimates of α and γ_L (which may or may not accord with theory). Our method can be seen as an atheoretical of finding Q . The two approaches should coincide if the economic restrictions underlying Q are "credible".

The focus of KPSW's analysis was on the permanent shocks. Warne (1991) extends KPSW's analysis to allow for simultaneous identification of the permanent and transitory shocks. Warne's analysis is therefore in the same spirit as ours. There are nevertheless important differences. Whereas our transitory shocks lie in the space spanned by α' , Warne's lie in the space spanned by γ . Essentially, what identify the transitory shocks are linear combinations of e_t that are statistically orthogonal to $\gamma_L e_t$. Since $\alpha' C(1) = C(1)\gamma = 0$, both methods are valid. Apart from maintaining a parallel with the Gonzalo-Granger decomposition, there is admittedly no reason why γ cannot be used in our analysis, though the reverse can be said of Warne's not choosing α' . The other difference lies in the implementation. Warne's method is achieved in a sequence of steps which has no apparent relation to standard VAR analyses. We accomplish the P-T decomposition in two steps, through the G and the H matrix respectively. In doing so, we keep the orthogonalized and the unorthogonalized shocks distinct. This separation makes it apparent, in our view, that a dynamic analysis of P-T shocks can be carried out as a simple extension of the way we have been using VARs. Additionally, the two-step exposition makes clear that G is dictated by the long-run properties of the data, and it is the properties of H matrix which is at the practitioner's discretion.

So far we have assumed that the cointegrating vectors are statistically identifiable. There are at least three ways how this can be implemented. The first is to impose restrictions on the eigenvectors of the solution to reduced rank regressions.⁵ The second is to apply linear restrictions based on economic theory when estimating a cointegrated system of simultaneous equations, such as discussed in Johansen (1995) and Beswijk (1995). The third is to estimate the model on an equation by equation basis, and make use of exclusion restrictions to achieve identification. For example, KPSW assume that investment does not enter the demand for money equation to identify one of the cointegrating vectors as the demand for money equation.

With the latter two estimation methods, the economic interpretation of the cointegrating vectors is clear, and it facilitates giving names to the shocks. Although it may seem ambiguous what interpretation one can give to the normalized eigenvectors from the first method, we have chosen to present our illustrative examples using this estimation strategy. Since we choose a Choleski

⁵The cointegrating vectors are identified by solving the eigenvalue problem $|\Delta S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0$. 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 $V(L)\Delta X_t$. Details are given in Johansen (1995)

decomposition for H , and we do not impose theoretical restrictions on $D(1)$, the $\hat{D}(1)$ that we identify is atheoretical. Essentially, we want to see how far the P-T decomposition can take us if we impose minimal economic theory throughout. For this reason, we refrain from giving names to the shocks, and merely refer to them as permanent shock 1, or permanent shock 2. As we will see, we can still use economic reasoning to interpret the impulse response functions. However, efficiency gains can be obtained by imposing economic restrictions on the cointegrating vectors if the restrictions are correct, as is always the case.

It is also useful to put into perspective our method of innovation accounting with the standard method based upon one step orthogonalization. In a bivariate case, it can be shown that the P-T orthogonalization is identical to the one-step Choleski decomposition if G is lower triangular, and if the orthogonalized shocks are assumed to have unit variances. In a multivariate setting, the equivalence holds only if G is lower triangular. A special case is when each row of γ_1' has zero elements in all but one position, which in turn implies that each permanent shock is induced by a different variable in the system. In such case, the source of the shock and the degree of persistence can both be identified.

4.1 Simulated Examples

We now use two examples to illustrate the properties of the proposed P-T decomposition. We shall assume that the rank of the cointegrating matrix is known. This allows us to focus on orthogonalization issues, but we will return with some remarks on the cointegrating rank. Throughout, we use reduced rank regressions with two lags to obtain $\hat{\alpha}$. Conditional on $\hat{\alpha}$, unrestricted estimates of γ are obtained from the VECM (2). In all cases, $\hat{\gamma}$ is constrained to zero and the VECM is re-estimated if the unconstrained estimate of γ is not statistically significant at the two-tailed 5% level. As will be discussed later, this is important for the precision of the results. 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 running under a 66mhz PC.

Example: DGP 1:

The first DGP we considered is based on the following triangular representation:

$$\begin{aligned} \Delta x_t &= u_{1t} \\ -x_t + y_t - z_t &= u_{2t} \\ .5x_t + .5y_t + z_t &= u_{3t}, \end{aligned} \tag{14}$$

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.

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 weakly exogenous to 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:

$$\begin{aligned} x_t &= y_t + 2z_t + u_{1t} \\ \Delta y_t &= u_{2t} \\ \Delta z_t &= u_{3t} \end{aligned} \tag{15}$$

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.

5 Three Empirical Examples

The methodology outlined in Section 2 is well suited for testing "convergence" in growth across countries, and models which predict differential response to permanent and transitory shocks. Indeed, use of our P-T decomposition in dynamic analysis was implicit in Cochrane (1994a). In that paper, Cochrane used a bivariate VECM to show that shocks to GNP holding consumption constant are transitory and that shocks to consumption have persistent effects. Consumption is defined as the sum of non-durables and services (CND S_t) and all data are in per-capita terms.

Our P-T decomposition can be used to obtain further intuition to Cochrane's results. Let $X_t = (gry_t, cnds_t)$ with lower case letters denoting logarithmic transforms. Using the notation of the previous section, $\hat{\alpha}'$ is $(1, -1)$ and $\hat{\gamma}'$ is $(-0.08, 0.02)$ for the sample 1947Q1-1989Q4. Since the $\hat{\gamma}$ in the consumption equation is not statistically significant, the coefficient is subsequently constrained to zero. The equilibrium error therefore does not enter the consumption equation. This particular structure of $\hat{\gamma}$ implies a $\hat{\gamma}'_1$ of $(0, x)$ for some $|x| > 0$. This implies that GNP has no weight in the trend component, from which it follows that all permanent shocks are due to consumption.

We reconsider Cochrane's analysis for the extended sample 1947Q1-1993Q4. A reduced-rank regression with 4 lags is used to estimate the cointegrating vector, and it is found to be $(1, -.886)'$. A constant is added to the VECM and left outside of the cointegration space. As in Cochrane, the error-correction term is not significant in the consumption equation. The adjustment coefficient in the GNP equation is $-.118$, slightly larger than Cochrane's. The impulse response functions based on the P-T decomposition are presented in the Figure 1 with standard errors in Table 3a. Consumption and GNP respond roughly in proportion to the permanent shock in the long-run, and the GNP response to the transitory shock is larger than that by consumption. The decomposition of variance reported in Table 3b confirms that transitory shocks are important for the short-run dynamics of GNP.

In the consumption-GNP example above, $\hat{\gamma}_1$ has the particular structure that gives a zero weight to GNP. In Cochrane's analysis, his variables are ordered with consumption first, so that G has a zero above the diagonal. As discussed earlier, the P-T and the one step (choleski) decomposition coincides when G is a lower block triangular matrix. However, this also suggests that had Cochrane reversed the ordering of the variables, the permanent and transitory shocks identified by his one step decomposition will be quite different. However, our P-T decomposition is robust to the ordering of the variables in this bivariate example.

The simple and specific structure of G in Cochrane's analysis is likely to be an exception rather than the rule. We now apply the P-T decomposition to other examples. In what follows, the critical values for cointegration tests are taken from Osterwald-Lennu (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. When, in addition, there is a trend in the VECM, the critical values are 31.4 and 21.5 for $r = 0$, and 16.06 and 14.84 for $r = 1$. We continue to use reduced-rank regressions. Economic restrictions are therefore not imposed. However, as we will see, the impulse response functions appear to bear meaningful economic interpretation.

5.1 Example 1: Consumption, GNP and Government Spending

Consider a three variable model with private sector output (GDPQ-GGEEQ), government expenditures (GGEEQ), and the consumption of non-durables and services (C=GCSQ+GCNQ). All variables are in per-capita terms and expressed in logs. Much interests have evolved around whether government spending is a substitute for private spending, and whether an increase in government spending crowds out private sector output in the sense of raising interest rates to meet financing requirements. Evidently, the results depend on whether increases in government spending are

permanent or transitory. We use the P-T decomposition to shed some light on these issues.

The Trace and Max- λ statistics based on a VECM with four lags are 34.96 and 23.17 for testing $r = 0$, and are 11.83 and 8.34 for the testing $r = 1$. We therefore conclude that $r = 1$. The cointegrating vector for $X_t = (y_t, c_t, g_t)$ is $(1, -315.2, -299.7)$, strongly indicating that the cointegration is between consumption and private sector output. The adjustment coefficient is significant only in the output equation with a value of .0006 and a t statistic of 4.37.

A cointegrating rank of one implies the presence of two permanent shocks. The implied $\hat{\gamma}_1$ is $[(0, 1, 0)', (1, 0, 0)']$. This suggests one permanent shock is due to consumption, and one is associated with government expenditures. The impulse response functions are presented in Figure 2 with standard errors in Table 4a. The dynamics of output and consumption in response to the first permanent shock are almost identical to that found in the bivariate model for consumption and output presented earlier.

Of more interest is the second permanent shock, which has a small and negative effect on private sector output in both the short and the long run. Output falls by .2% over the simulation horizon with a maximum standard error of is .09. Putting the analysis in the context of a stochastic growth model with a government sector, the first permanent shock can be seen as a productivity shock and the second as a government expenditure shock. Our results suggest that the effects of government spending are small, but nevertheless statistically significant. Quantitatively, our results contrast the findings of Baxter and King (1993), who report that permanent government shocks have large effects on private sector behavior in a simulation model. It would be interesting to see if labor supply and intertemporal substitution effects, which are not accounted for in our VECM, can explain the differences.

The decomposition of variance reported in Table 4b reinforces the finding that consumption is dominated by permanent shocks, and that permanent government shocks have little long run effects on private sector behavior. As well, the transitory shock is found to account for one-third to three-quarters of the output variations in the first four periods, suggesting once again that output variations have a large transitory component.

5.2 Example 2: A Two Country Model

The second example we present is based on real GNP of Japan (JP), the U.S. (US), and the real exchange rate (REX) between the two countries. All variables are expressed in logs. Since the nominal exchange rate is Yen per US dollar, a higher real exchange rate means that the US dollar is worth more. There are evidently important structural differences between the two economies, and it is interesting to see if the two countries respond to permanent and transitory shocks differently.

Our estimation is based on data from 1961Q1 to 1994Q4.⁶ The three variable model has a VECM representation with four lags, a constant and a deterministic trend outside the cointegration space. This implies that the level of the variables could have quadratic trends. This is intended to pick up the fact that Japan's average growth rate is noticeably slower after 1973. The Trace and Max- λ statistics are 36.55 and 20.996 for the null hypothesis that $r = 0$, and are 15.55 and 11.92 for the null hypothesis that $r = 1$. The statistics therefore suggest one cointegrating vector and two unit roots. The cointegrating vector normalized for $X = (us, jp, rex)$ is $(1, -.194 \text{ and } .199)$. The λ' vector is $(-.11, .14, 0)$. The coefficient is significant (with t statistics close to 3) in the two output equations but not in the real exchange rate equation. The implied $\hat{\gamma}_1$ suggests that the first permanent shock is due to innovations in the two output equations (with weights .70 and .60), and that the second permanent shock is exclusively a real exchange rate shock.

The impulse response functions are shown in Figure 3 with standard errors in Table 5a. A unit increase in the first permanent shock has a positive impact on output in the two countries, but the long-run impact of the shock is substantially more favorable for Japan. During the transition from one equilibrium to another, the real exchange rate appears to depreciate through this effect is not well determined. For the second permanent shock, an appreciation in the real value of the U.S. dollar vis-à-vis the Yen (i.e. an increase in the real exchange rate) increases Japan's output and decreases output in the U.S. by a proportionally larger amount. The responses to the transitory shock are generally not well determined, though they suggest that transitory shocks are important for the dynamics of U.S. output, in accordance with the closed-economy models.

Results from the decomposition of variance confirm that over the short horizon, output in both countries are affected by both permanent shocks. However, the sources of output variations in the two countries differ over longer horizons. With a two-country growth model as the backdrop, the first permanent shock behaves like a productivity shock and the second like a terms of trade shock. Using this interpretation, the results suggest that while long term variations in Japan's output are due exclusively to the permanent productivity shock, shocks to the real exchange rate explain close to 40% of the variations in output in the U.S.. In other words productivity shocks are important but are not the only source of fluctuations in U.S. output.

Our P-T decomposition suggests a permanent component in the real exchange rate. Campbell and Clarida (1987) used a completely different methodology and arrived at a similar conclusion. However, our results contrast a recent finding of Clarida and Gali (1994) that there is a large transitory component in the dollar-yen real exchange rate. Although Clarida and Gali proceeded

⁶The data are taken from the BIS (Bank of International Settlements) The authors acknowledge the Bank of Canada for supplying this data.

with a multivariate Beveridge-Nelson decomposition of the real exchange rate, and is in the spirit of our P-T decomposition, cointegration restrictions were not taken into account in their analysis. This could account for the drastically different implications.

5.3 Example 3: Is Money Superneutral?

There is a vast literature in empirical macroeconomics questioning the effects of monetary policy variables on real variables. As discussed in Cochrane (1994b), the evidence based on VARs is quite sensitive to the size of the model, lag length selection, and the assumption on trends. However, M2 and price are $I(2)$ processes, but interest rates, output, and consumption are $I(1)$.⁷ 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.

As discussed in Fisher and Seater (1993), the ability to interpret coefficients purporting to test money neutrality depends on the order of integration of the variables. If money is $I(2)$, it is not informative to ask if money has a neutral effect on output because the latter is one order of integration lower than money. Nevertheless, it is appropriate to ask if 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 Sidranski (1967), and empirically in Geweke (1986), among others.

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. Unlike standard VARs which make explicit assumption about the causal ordering of the variables, our orthogonalization allows for a concurrent change in money growth and the Fed Funds rate. Thus, we entertain the possibility of a monetary policy variable comprising of two instruments, in so far as money growth and the interest rate have non-zero weights in $\hat{\gamma}_1$.

The cointegration tests suggest one cointegrating vector. 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 $r = 1$. The cointegrating vector for $X_t = (\Delta m_2, FF, c, \pi)$ is $(1, -.378, .045, -.805)$. 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

⁷For the sample 1959Q1-1994Q1, the $I_t(k)$ statistic of Said and Dickey (1984) with the truncation lag selected as discussed in Ng and Perron (1986) cannot reject a unit root in either money growth or inflation (of the GDP deflator). The statistics, based upon $kmxz = 10$ are -2.14 and -1.88 respectively. For FYGM3 (3 month t-bill rate) and FFVF (fed-funds rate), the unit root tests are -1.68 and -1.88 respectively.

errors of the impulse response functions.

The vector $\hat{\gamma}$ is (-.346, .090, .081, and .16)' with t statistics of -3.44, 2.09, 1.71, and 2.59. The three columns of $\hat{\gamma}_1$ are [(-.225, .973, -.024, -.049)', (.203, -.024, .978, -.044)', (.407, -.049, -.044, .911)']. Accordingly, none of the permanent shocks come exclusively from one variable alone. The innovations in the four I(1) processes each contribute to more than one permanent shock. It is precisely under these conditions that the impulse response functions based on the P-T decomposition will reveal dynamics that will be different from a one-step decomposition.

The impulse response functions are presented in Figure 4 with standard errors in Table 6a. 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. Such a real interest rate shock was also identified by King et al. (1991), but unlike that study, 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 witnessed a permanent increase in consumption similar in shape and magnitude as that found in the consumption-output example. A one percent shock raises consumption by 1.4 percent, 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.

The third permanent shock can be viewed as a money growth shock that raises the inflation rate by roughly the same proportion. The question of whether money is superneutral is therefore best analyzed in this context. The impulse responses suggest a permanent reduction in the real interest rate and a mild reduction in consumption of -.2 of one percent with a maximum standard error of 0.09. Note that the Mundell-Tobin effect can be evaluated by extending the model to include output. If output does not fall by more than consumption, saving would increase, leading to a higher level of capital. As it is, the data suggest mild super non-neutrality of money growth.

The transitory shock is evidently a temporary money growth shock in which the interest rate falls, but there is also evidence of the price puzzle with inflation falling in response to an increase in money growth. This raises the real interest rate, and causes a reduction in consumption. The decomposition of variance is reported in Table 6b. Although we have not used economic theory to identify the cointegrating vectors, the results can still be rationalized by a monetary growth model with multiple shocks. According to the results, 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 small contributions from the money growth shock. Results using the 3 month treasury bill rate are similar.

6 Practical Issues

In this section, we discuss two practical issues that apply to the P-T decomposition: the cointegrating restrictions, and the estimation of γ and γ_1 .

6.1 Importance of the Cointegrating Restrictions

Suppose there are $r^* < n$ cointegrating vectors, and we estimate an unrestricted VAR in level form. The efficiency loss on the estimates of $A(L)$ from ignoring the cointegrating restrictions are by now well known. Lesser known are the implications of the restrictions from the point of view of innovation accounting. In a recent paper, Phillips (1995) showed that the long horizon impulse responses estimated by an unrestricted VAR can be inconsistent. The reason is that the true impulse responses of an unstable system does not die out as the sample size increases, and inherit the effects of the unit roots at all horizons. Errors from estimating the unit roots therefore cumulate as the simulation horizon increases. In consequence, the impulse responses will be distorted regardless of how we orthogonalize the shocks.

Ignoring cointegrating restrictions can have two additional implications which applies only to the P-T decomposition. The first is that γ is not an explicitly identifiable parameter in an unrestricted VAR, and therefore values of γ_1 are not available for the P-T decomposition. The second problem is that the number of permanent shocks eventually depends on the rank of $C(1)$ associated with the unrestricted VAR, which may not coincide with $n - r^*$, the true number of permanent shocks. Equivalently, the issue is whether $A(1)$ will have rank r^* absent cointegrating restrictions. As discussed in Engle and Yoo (1987), the rank of $A(1)$ will be exactly r^* when cointegration restrictions are imposed. However, when the restrictions are not imposed, this rank will generally be $r \neq r^*$ in finite samples except for an event with probability measure zero.

The problem of not imposing cointegrating restrictions can easily be seen from a VECM with no lags, and hence $A(L)$ is of order one. The unrestricted estimates \hat{A}_1 and the restricted estimate \bar{A}_1 are related by

$$\bar{A}_1 = \hat{A}_1 - (I_n + \hat{\gamma}\hat{\alpha}' - \hat{A}_1).$$

The rank of $\bar{A}(1)$ evidently depends on how far are the unrestricted estimates from $I_n + \gamma\alpha'$, since $\hat{\alpha}$ and $\hat{\gamma}$ converges to α and γ in probability.

Conditional on $\hat{\alpha}$, the VECM provides an estimate of $\hat{\gamma}$, from which one constructs $\hat{\gamma}_\perp$. Methods for finding orthogonal vectors can be found in matrix textbooks [e.g. Golub and Loan (1984)]. One can take the appropriate eigenvectors from the matrix $\hat{\gamma}(\hat{\gamma}'\hat{\gamma})^{-1}\hat{\gamma}'$, or one can take the $r+1$ columns of the left singular vectors of $\hat{\gamma}$.⁸ All methods will satisfy the restriction that $\hat{\gamma}_\perp'\hat{\gamma} = 0_{(n-r) \times r}$. However, as discussed in Podivinsky (1992), $\hat{\gamma}$ can have poor finite sample properties. To the extent that the $\hat{\gamma}_\perp$ 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 examples, where we take the left eigenvectors of $I_3 - \hat{\gamma}(\hat{\gamma}')^{-1}\hat{\gamma}'$ as $\hat{\gamma}_\perp$.

Consider a three variable (x_t, y_t, z_t) system. Suppose the true value of γ is $(1, 0, 0)'$ as in DGP2 of (15). 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 the γ in the second equation is not statistically significant but has a numerically small value of .05. We have

$$\hat{\gamma} = \begin{bmatrix} 1 \\ .05 \\ 0 \end{bmatrix} \Rightarrow \hat{\gamma}_\perp = \begin{bmatrix} .05 & .09 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This implies that x_t has a non-zero weight in the first permanent shock, which is apparently inconsistent with the DGP. If we further let

$$\hat{\gamma} = \begin{bmatrix} 1 \\ .1 \\ .1 \end{bmatrix} \Rightarrow \hat{\gamma}_\perp = \begin{bmatrix} .197 & -.985 & -.985 \\ 3.43 & -16.68 & -17.68 \end{bmatrix}.$$

Although .1 seems not too far from the true value of 0, the implications for the P-T decomposition can be far reaching. In the last case, $\hat{\gamma}_\perp$ gives y_t and z_t equal importance in the two permanent shocks instead of giving an exclusive weight of 1 to one variable. The results are similar if we estimate $\hat{\gamma}_\perp$ by singular value decomposition. The problem arises because the orthogonal complement of a matrix, say, z , is not continuous in small perturbations in z . For this reason, $C(1)$ (which depends on $\hat{\alpha}_\perp$ and $\hat{\gamma}_\perp$) and $D(1)$ (which depends on $\hat{\gamma}_\perp$) are very sensitive to small variations in $\hat{\alpha}$ and/or $\hat{\gamma}$. While an estimate of $C(1)$ is not needed in our P-T decomposition, careful estimation of $\hat{\gamma}_\perp$ is still necessary because it affects the precision of $D(1)$ through G .

⁸The singular value decomposition of A is given by $A'Av = \text{diag}(\sigma_1, \dots, \sigma_n)$, where v and v are the left and right singular vectors, and $\sigma_1, \dots, \sigma_n$ are the singular values.

Engle and Yoo conjectured that the rank of $\hat{A}(1)$ will likely be $r > r^*$ (in fact, n), on the ground that downward biases in the autoregressive parameters in finite samples will lead one to conclude too few unit roots. Although our simulations fail to detect a systematic under-estimation of the number of permanent shocks, it is generally the case that $\hat{C}(1)$ has rank $n - r \neq n - r^*$. In other words, we can have too many or too few permanent shocks. Not imposing cointegrating restrictions on a VAR mixes up the permanent and the transitory components of the model, and hence the degree of persistence of the shocks. The implications for analyzing the dynamic effects of permanent and transitory shocks can be far reaching.

The above discussion suggests that the number of permanent and transitory shocks can also be mis-diagnosed if we impose the wrong number of cointegrating restrictions. Suppose the true number of cointegrating vectors is r^* , and the practitioner imposes $r \neq r^*$ cointegrating relationships. By analogy to the unrestricted estimates, we have

$$\hat{A}_1 = \hat{A}_1 - (I_n + \hat{\gamma}\hat{\alpha}' - \hat{\gamma}\hat{\alpha}'),$$

where $\hat{\alpha}$ and $\hat{\gamma}$ are of rank r rather than r^* . Simple arithmetic shows that $\hat{C}(1)$ will have rank $n - r$, and the P-T decomposition will reveal $n - r$ permanent shocks.

When $r > r^*$, the problem is that non-stationary combinations of the variables are included in the VECM, and statistics associated with them will have non-standard distributions. Classical inference may falsely find these spurious regressors to be significant. The consequence of increasing the column rank of $\hat{\gamma}$ beyond r^* is finding too few permanent shocks.

When $r < r^*$, the omitted regressors are stationary and are necessarily correlated with the included regressors. The orthonormalized shocks could be serially correlated, and the so-called orthogonalized shocks could be cross correlated. The problem can be traced to the fact that the residuals of the VECM are not genuine innovations and $\hat{\gamma}_\perp'\gamma \neq 0$. The consequence is over-estimating the rank of $C(1)$ and finding too many permanent shocks.

Summarizing, when $r \neq r^*$ the matrix G will not be consistently estimated. While it is undesirable to ignore cointegrating restrictions, imposing false restrictions could be equally treacherous in impulse response analysis. Careful attention should therefore be paid to determining the rank of the cointegrating matrix before implementing the P-T decomposition.

6.2 The estimation of γ and γ_\perp

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 $\hat{\alpha}$. An alternative is to use the common trend statistic of Stock and Watson (1988) to determine the r , and then estimate $\hat{\alpha}$ by fully efficient estimators. Which method (or combination of methods) to use is at the user's discretion.

An alternative method for estimating $\hat{\gamma}_L$ with a completely different orientation was proposed by Gonzalo and Granger (1995). There, the calculation of $\hat{\gamma}_L$ was treated as a dual to the problem of estimating $\hat{\alpha}$. It essentially solves for the eigenvalues λ and eigenvectors V from the equations $|\lambda S_{00} - S_{01} S_{11}^{-1} S_{10}|$ subject to the constraint that $V' S_{00} V = I$. It follows from this latter normalization that $\hat{\gamma}_L$ satisfies $\hat{\gamma}_L \hat{\Omega} \hat{\gamma}_L' = I$. In other words, the shocks $\hat{\gamma}_L' e_t$ are already made mutually independent by the construction of $\hat{\gamma}_L$. Therefore, one can obtain orthogonalized transitory shocks by projecting the unorthogonalized ones onto the orthogonalized permanent shocks. This could be seen as a modified two-step orthogonalization.

We have experimented with various methods of calculating $\hat{\gamma}_L$. Among these are 1) the eigenvectors associated with the $n - r$ smallest eigenvalues of $I - \hat{\gamma}(\hat{\gamma}')^{-1} \hat{\gamma}'$, 2) the singular value decomposition of $\hat{\gamma}$, and 3) the method discussed in Gonzalo and Granger (1995). Although the non-uniqueness of $\hat{\gamma}_L$ makes for a precise comparison of the various methods difficult, our experience based on a small Monte Carlo experiment is that Method 3 is numerically less precise in the sense that $\hat{\gamma}_L' \hat{\gamma}$ is closer to zero with Methods 1 and 2. In general, Method 3 requires a large sample size to achieve the same level of accuracy as the matrix methods. The weakness of Method 3 is more noticeable when many elements of $\hat{\gamma}$ are zero. From a practical standpoint, Method 1 could be unstable when $\hat{\gamma}' \hat{\gamma}$ is closer to singular, a case which we cannot rule out in practice. We have therefore opted for the approach of Singular Value Decomposition in our empirical analyses.

7 Conclusion

Vector Autoregressions is a valuable framework for dynamic economic analyses. The conventional wisdom is to analyze the impulse response of variables in the system with respect to shocks to these variables. 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. We suggest a two-step orthogonalization which allows the dynamic response of variables to the permanent and transitory shocks to be traced out in a systematic way. The main advantage of the proposed procedure is its simplicity. The framework provides an alternative view to many issues of macroeconomic interests.

Appendix

Proof of Lemma 1

Suppose we have two different permanent-transitory decompositions for the level of X_t , so that

$$X_t = A_1 f_t + x_{1t}, \quad (16)$$

where f_t is $I(1)$ and x_{1t} is $I(0)$, $i = 1, 2$. Suppose Δf_t is serially correlated. We can rewrite $f_t = r_{1t} + \tilde{x}_{1t}$, where $r_{1t} = r_{1,t-1} + a_{1t}$, \tilde{x}_{1t} is $I(0)$. Then we have

$$X_t = A_1 a_{1t} (1-L)^{-1} + x_{1t} + \tilde{x}_{1t} = A_1 a_{1t} (1-L)^{-1} + \tilde{x}_{1t}. \quad (17)$$

Since $A_1 a_{1t} / (1-L)$ and $A_2 a_{2t} / (1-L)$ are now two sets of random walks both explaining X_t , they must be cointegrated. Assuming that $E_{t-1} a_{1t} = 0$, this implies $A_1 a_{1t} = A_2 a_{2t}$, or that the two decompositions have random walks that are linearly related. The implications in terms of their respective innovations follow.

The Stock-Watson trend-cycle decomposition is a special case where \tilde{x}_{1t} is a white noise. To see that the random walk component in the Stock-Watson trend is the same as that in the permanent part of the Granger-Gonzalo decomposition, multiplying (2) by γ_L' . Using (5), it can be shown that

$$\Delta f_t = \sum_{i=1}^{K-1} \gamma_L' \Gamma_i \theta_i \Delta f_{t-i} + \sum_{i=1}^{K-1} \gamma_L' \Gamma_i \theta_2 \Delta x_{t-i} + \gamma_L' e_t. \quad (18)$$

The random walk component of f_t in the Gonzalo-Granger decomposition is therefore

$$(I - \sum_{i=1}^{K-1} \gamma_L' \Gamma_i \theta_i)^{-1} \gamma_L' \sum_{s=1}^t e_s. \quad (19)$$

The permanent component in the Stock-Watson decomposition is $C(1) \gamma_L'$. Since $\alpha_{\perp} (\gamma_L' \psi(\alpha_L))^{-1} \gamma_L' = C(1)$, where $\psi = I_n - \sum_{i=1}^{K-1} \Gamma_i$, we can also write

$$\begin{aligned} C(1) &= \alpha_{\perp} (\gamma_L' [I_n - (\gamma_L' \sum_{i=1}^{K-1} \Gamma_i \alpha_{\perp})^{-1} \gamma_L']^{-1} \gamma_L' = \alpha_{\perp} (\gamma_L' \alpha_{\perp})^{-1} (\gamma_L' \alpha_{\perp}) [\gamma_L' (I_n - \sum_{i=1}^{K-1} \Gamma_i) \alpha_{\perp}]^{-1} \gamma_L' \\ &= \alpha_{\perp} (\gamma_L' \alpha_{\perp})^{-1} [\gamma_L' (I_n - \sum_{i=1}^{K-1} \Gamma_i) \alpha_{\perp} (\gamma_L' \alpha_{\perp})^{-1}]^{-1} \gamma_L' \\ &= \theta_1 [I_n - \gamma_L' \sum_{i=1}^{K-1} \Gamma_i \theta_i]^{-1} \gamma_L'. \end{aligned} \quad (20)$$

Using the fact that $\theta_1 = \alpha_{\perp} (\gamma_L' \alpha_{\perp})^{-1}$. Comparing (20) with (19) it can be seen that the random walk components (in the Beveridge-Nelson sense) in the two representations of X_t are the same.

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 $f_t = \gamma_1' X_t$ and $z_t = \alpha' X_t$. Using the result in Gonzalo and Granger (1995) that $\theta_1 = \alpha_1' (\gamma_1' \alpha_1)^{-1}$ and $\theta_2 = \gamma' (\alpha' \gamma)^{-1}$, we can have

$$X_t = \theta_1 f_t + \theta_2 z_t = P_t + T_t. \quad (21)$$

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

$$\begin{bmatrix} I_n - (\gamma_1' \Gamma_1 A_1) L & (-\gamma_1' \Gamma_1 A_2) L(1-L) \\ (-\alpha' \Gamma_1 A_1) L & I_n - (\alpha' \gamma - I_n + \alpha' \Gamma_1 A_2) L + (\alpha' \Gamma_1 A_2) L^2 \end{bmatrix} \begin{bmatrix} \Delta f_t \\ z_t \end{bmatrix} = \begin{bmatrix} u_t^p \\ u_t^z \end{bmatrix}, \quad (22)$$

which we write more compactly as

$$\begin{bmatrix} F_{11}(L) & F_{12}(L) \\ F_{21}(L) & F_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta f_t \\ z_t \end{bmatrix} = \begin{bmatrix} u_t^p \\ u_t^z \end{bmatrix}. \quad (23)$$

Note that $F_{12}(1) = 0$. Inverting (22),

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

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

$$\begin{bmatrix} \theta_1 & 0 \\ 0 & \theta_2(1-L) \end{bmatrix} \begin{bmatrix} \Delta f_t \\ z_t \end{bmatrix} = \begin{bmatrix} \theta_1 & 0 \\ 0 & \theta_2(1-L) \end{bmatrix} \begin{bmatrix} F^{11}(L) & F^{12}(L) \\ F^{21}(L) & F^{22}(L) \end{bmatrix} \begin{bmatrix} u_t^p \\ u_t^z \end{bmatrix}, \text{ and hence} \\ \begin{bmatrix} \Delta P_t \\ \Delta T_t \end{bmatrix} = \begin{bmatrix} \theta_1 F^{11}(L) & \theta_1 F^{12}(L) \\ \theta_2(1-L) F^{21}(L) & \theta_2(1-L) F^{22}(L) \end{bmatrix} \begin{bmatrix} u_t^p \\ u_t^z \end{bmatrix}.$$

Thus, $\Delta X_t = \Delta P_t + \Delta T_t$

$$= \begin{bmatrix} \theta_1 F^{11}(L) \theta_2(1-L) F^{21}(L) + \theta_1 F^{12}(L) + \theta_2(1-L) F^{21}(L) \\ \theta_2(1-L) F^{22}(L) \end{bmatrix} \begin{bmatrix} u_t^p \\ u_t^z \end{bmatrix} \quad (24)$$

$$= \begin{bmatrix} D_{11}(L) & D_{12}(L) \\ D_{21}(L) & D_{22}(L) \end{bmatrix} \begin{bmatrix} u_t^p \\ u_t^z \end{bmatrix}. \quad (25)$$

and note that $D_{12}(1) = D_{22}(1) = 0$, because $\theta_1 F^{12}(1) + \theta_2 \times 0 \times F^{21}(1) = 0$.

Proof of Corollary 1

Since

$$G = \begin{bmatrix} \gamma_1' \\ \alpha' \end{bmatrix},$$

inverting G , we have

$$G^{-1} = [\theta_1 | \theta_2] = [\alpha_1 (\gamma_1' \alpha_1)^{-1} | \gamma' (\alpha' \gamma)^{-1}].$$

Using the definition of $C(1)$ in (20), we have $C(1)G^{-1} = D(1)$, where the last r columns are zero.

References

- Baxter, M. and King, R. G. (1993), Fiscal Policy in General Equilibrium, *American Economic Review* **83**, 315-334.
- Bernanke, B. (1986), Alternative explanations of the money-income correlator, *Carnegie-Rochester Conference on Public Policy* **25**, 49-100.
- Bernanke, B. S. and Mihov, I. (1995), Measuring Monetary Policy, Princeton University.
- Beveridge, S. and Nelson, C. R. (1981), A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle, *Journal of Monetary Economics*.
- Blaug, O. and Quah, D. (1989), The Dynamic Effects of Aggregate Demand and Supply, *American Economic Review* **79**, 655-73.
- Blanchard, O. and Quah, D. 1993, The Dynamic Effects of Aggregate Demand and Supply: Reply, *American Economic Review* **83**, 653-58.
- Boswijk, H. P. (1995), Efficient Inference on Cointegrating Parameters in Structural Error Correction Models, *Journal of Econometrics* **69**, 131-157.
- Campbell, J. Y. and Clarida, R. (1987), The Dollar and Real Interest Rates: An Empirical Investigation, *Carnegie-Rochester Series on Public Policy* **27**, 103-140.
- Clarida, R. and Gali, J. (1994), Sources of Real Exchange-Rate Fluctuations: How Important are Nominal Shocks, *Carnegie-Rochester Series on Public Policy* **41**, 1-65.
- 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. and Granger, C. (1987), Cointegration and Error-Correction: Representation, Estimation, and Testing, *Econometrica* **55**, 251-276.
- Engle, R. F. and Yoo, B. S. (1987), Forecasting and Testing in Cointegrated Systems, *Journal of Econometrics* **35**, 143-59.
- Fisher, M. and 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.
- Golub, G. H. and Loan, C. F. V. (1984), *Matrix Computations*, John Hopkins University Press, Baltimore.
- Gonzalo, J. and Granger, C. (1995), Estimation of Common Long-Memory Components in Cointegrated Systems, *Journal of Business and Economic Statistics* **13**, 27-35.
- Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, N.J.
- 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**, 1551-80.
- 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. and Watson, M. (1991), Stochastic Trends and Economic Fluctuations, *American Economic Review* **81**:4, 819-40.
- Koray, F., Lee, T. H. and Palivos, T. (1995), Stochastic Trends and Fluctuations in National Income, Wages, and Profits, mimeo, Louisiana State University.
- Lippi, M. and Reichlin, L. 1993, The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment, *American Economic Review* **83**, 644-52.
- Lutkepohl, H. (1993), *Introduction to Multiple Time Series*, Springer Verlag, Berlin.
- Lutkepohl, H. and Reimers, H. E. (1992), Impulse Response Analysis of Cointegrated Systems, *Journal of Economic Dynamics and Control* **16**, 53-78.
- Ng, S. and Perron, P. (1995), Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag, *Journal of the American Statistical Association* **90**, 268-281.
- 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.
- Phillips, P. (1995), Impulse Response and Forecast Error Variance Asymptotics in Nonstationary VAR's, Cowles Foundation Working Paper 1102.
- Podivinsky, J. M. (1992), Small Sample Properties of Tests of Linear Restrictions on Cointegrating Vectors and their Weights, *Economic Letters* **39**, 13-18.
- Proietti, T. (1995), Short Run Dynamics in Cointegrated Systems. mimeo, Università di Perugia.
- Quah, D. and Vahey, S. P. (1995), Measuring Core Inflation, *Economic Journal* **105**, 1130-1144.
- Runkle, D. E. (1987), Vector Autoregressions and Reality, *Journal of Business and Economic Statistics* **5**, 437-43.
- Said, S. E. and Dickey, D. A. (1984), Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, *Biometrika* **71**, 599-607.

Table 1: DGP 1

$$\Delta z_t = u_{1t}$$

$$-z_t + y_t - z_t = u_{2t}$$

$$.5z_t + .5y_t + z_t = u_{3t}$$

Impulse Response Functions

Period	P Shock			T Shock 1			T Shock 2		
	z	y	z	z	y	z	z	y	z
1	0.882	0.291	-0.688	0.118	0.036	0.656	-0.005	0.324	0.064
2	0.894	0.296	-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

Decomposition of Variances

Period	P Shock			T Shock 1			T Shock 2		
	z	y	z	z	y	z	z	y	z
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	0.383	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

Table 2: DGP 2

$$z_t = y_t + 2z_t + u_{1t}$$

$$\Delta y_t = u_{2t}$$

$$\Delta z_t = u_{3t}$$

Impulse Response Functions

Period	P Shock			T Shock 2			T Shock		
	z	y	z	z	y	z	z	y	z
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

Decomposition of Variance

Period	P Shock			T Shock 2			T Shock		
	z	y	z	z	y	z	z	y	z
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

Shapiro, M. and Watson, M. (1989), Sources of Business Cycle Fluctuations, *NBER Macroeconomics Annual*.

Sidrauski, M. (1967), Rational Choice and Patterns of Growth in a Monetary Economy, *American Economic Review Papers and Proceedings* pp. 534-44.

Sims, C. A. and Zha, T. (1994), Error Bands for Impulse Responses, mimeo, Yale University.

Stock, J. H. and 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, *Handbook of Econometrics*, Vol. 4, pp. 2844-2910.

Table 4a: Model G-CNDS-Y:
 $G = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 315.21 & -299.77 \end{bmatrix}$

Period	Standard Errors of the Impulse Response Functions						
	P Shock 1		P Shock 2		T Shock		
	g	y	g	y	g	y	
1	0.135	0.127	0.109	0.035	0.075	0.006	0.001
2	0.202	0.178	0.197	0.046	0.088	0.006	0.001
3	0.230	0.221	0.373	0.068	0.111	0.007	0.001
4	0.301	0.321	0.700	0.094	0.116	0.008	0.001
5	0.404	0.394	1.062	0.107	0.200	0.009	0.001
6	0.476	0.436	1.168	0.123	0.224	0.008	0.001
7	0.563	0.481	1.129	0.142	0.228	0.007	0.001
8	0.644	0.514	1.122	0.160	0.231	0.006	0.001
9	0.713	0.534	1.086	0.176	0.234	0.005	0.001
10	0.774	0.552	1.031	0.191	0.236	0.004	0.001
11	0.832	0.569	0.979	0.204	0.239	0.003	0.001
12	0.881	0.581	0.927	0.214	0.241	0.002	0.001
13	0.924	0.593	0.874	0.224	0.242	0.002	0.001
14	0.963	0.604	0.835	0.232	0.243	0.001	0.000
15	1.000	0.613	0.809	0.239	0.244	0.001	0.000
16	1.029	0.623	0.792	0.245	0.245	0.000	0.000
17	1.054	0.632	0.784	0.251	0.246	0.000	0.000
18	1.074	0.640	0.781	0.255	0.247	0.000	0.000
19	1.090	0.648	0.782	0.259	0.247	0.000	0.000

Period	Standard Errors of the Impulse Response Functions						
	P Shock 1		P Shock 2		T Shock		
	g	y	g	y	g	y	
1	0.002	1.000	0.233	0.988	0.000	0.000	0.767
2	0.001	0.997	0.437	0.999	0.000	0.002	0.573
3	0.002	0.995	0.545	0.996	0.003	0.001	0.454
4	0.002	0.994	0.645	0.987	0.004	0.001	0.354
5	0.002	0.993	0.716	0.988	0.006	0.005	0.278
6	0.002	0.992	0.765	0.988	0.007	0.007	0.228
7	0.003	0.990	0.796	0.989	0.009	0.007	0.196
8	0.004	0.989	0.820	0.989	0.010	0.007	0.173
9	0.006	0.987	0.838	0.988	0.012	0.006	0.156
10	0.008	0.985	0.851	0.987	0.014	0.006	0.143
11	0.010	0.983	0.862	0.985	0.016	0.006	0.133
12	0.012	0.982	0.870	0.983	0.018	0.005	0.124
13	0.015	0.980	0.877	0.981	0.019	0.006	0.118
14	0.017	0.979	0.882	0.979	0.020	0.006	0.112
15	0.019	0.978	0.887	0.977	0.022	0.007	0.106
16	0.022	0.977	0.891	0.976	0.023	0.007	0.102
17	0.024	0.976	0.894	0.974	0.023	0.008	0.098
18	0.026	0.975	0.897	0.972	0.024	0.009	0.094
19	0.027	0.975	0.900	0.970	0.025	0.010	0.090

Table 3a: Model GNP-CNDS
 Standard Errors of the Impulse Response Functions

Period	P Shock		T Shock	
	gnp	cnlds	gnp	cnlds
1	0.070	0.029	0.011	0.015
2	0.096	0.037	0.038	0.017
3	0.116	0.061	0.050	0.020
4	0.140	0.085	0.063	0.018
5	0.149	0.095	0.056	0.018
6	0.155	0.108	0.050	0.015
7	0.155	0.118	0.045	0.012
8	0.153	0.124	0.041	0.009
9	0.148	0.129	0.036	0.007
10	0.140	0.133	0.030	0.005
11	0.132	0.134	0.024	0.003
12	0.123	0.135	0.018	0.002
13	0.116	0.136	0.013	0.002
14	0.110	0.137	0.010	0.001
15	0.106	0.138	0.007	0.001
16	0.104	0.139	0.006	0.001
17	0.103	0.140	0.004	0.000
18	0.103	0.141	0.003	0.000
19	0.104	0.142	0.003	0.000

Table 3b: Decomposition of Variance

Period	P Shock		T Shock	
	gnp	cnlds	gnp	cnlds
1	0.217	1.000	0.783	0.000
2	0.358	0.992	0.642	0.008
3	0.459	0.992	0.541	0.008
4	0.535	0.995	0.465	0.005
5	0.597	0.995	0.403	0.005
6	0.642	0.996	0.358	0.004
7	0.675	0.997	0.325	0.003
8	0.700	0.997	0.300	0.003
9	0.720	0.998	0.280	0.002
10	0.736	0.998	0.264	0.002
11	0.749	0.998	0.251	0.002
12	0.761	0.998	0.239	0.002
13	0.771	0.998	0.229	0.002
14	0.780	0.999	0.220	0.001
15	0.786	0.999	0.212	0.001
16	0.795	0.999	0.205	0.001
17	0.801	0.999	0.199	0.001
18	0.807	0.999	0.193	0.001
19	0.813	0.999	0.187	0.001

Table 5a: Model USA-JAPAN-REX

$$G = \begin{bmatrix} .793 & .609 & 0 \\ 0 & 0 & 1 \\ 1 & -.194 & .199 \end{bmatrix}$$

Standard Errors of the Impulse Response Functions

Period	P Shock 1		P Shock 2		T Shock	
	US	JAPAN	US	JAPAN	US	JAPAN
1	0.176	0.223	0.218	0.103	0.005	0.072
2	0.276	0.243	0.449	0.151	0.010	0.182
3	0.370	0.349	0.605	0.205	0.018	0.267
4	0.360	0.389	0.707	0.193	0.030	0.320
5	0.399	0.414	0.852	0.206	0.044	0.361
6	0.369	0.463	0.893	0.195	0.055	0.392
7	0.359	0.533	0.899	0.194	0.072	0.437
8	0.313	0.609	0.862	0.184	0.086	0.495
9	0.286	0.690	0.828	0.178	0.101	0.547
10	0.249	0.772	0.801	0.170	0.115	0.588
11	0.224	0.849	0.793	0.164	0.128	0.625
12	0.198	0.920	0.784	0.159	0.140	0.655
13	0.182	0.977	0.780	0.157	0.150	0.681
14	0.167	1.025	0.777	0.156	0.158	0.700
15	0.158	1.061	0.778	0.155	0.163	0.715
16	0.150	1.090	0.779	0.154	0.170	0.724
17	0.146	1.109	0.782	0.155	0.174	0.728
18	0.142	1.123	0.784	0.155	0.177	0.730
19	0.140	1.132	0.786	0.156	0.179	0.732

Table 5b: Model USA-JAPAN-REX

Decomposition of Variance

Period	P Shock 1		P Shock 2		T Shock	
	US	JAPAN	US	JAPAN	US	JAPAN
1	0.652	0.557	0.003	0.001	0.001	0.997
2	0.706	0.642	0.005	0.008	0.010	0.994
3	0.705	0.699	0.012	0.050	0.011	0.984
4	0.714	0.737	0.014	0.071	0.018	0.980
5	0.710	0.767	0.016	0.091	0.023	0.977
6	0.702	0.795	0.018	0.114	0.029	0.975
7	0.691	0.819	0.019	0.138	0.029	0.975
8	0.680	0.841	0.019	0.160	0.030	0.975
9	0.668	0.860	0.019	0.181	0.031	0.975
10	0.657	0.876	0.018	0.202	0.031	0.976
11	0.645	0.889	0.018	0.221	0.030	0.977
12	0.634	0.900	0.017	0.239	0.029	0.978
13	0.624	0.909	0.016	0.255	0.028	0.979
14	0.614	0.917	0.015	0.270	0.027	0.981
15	0.604	0.924	0.015	0.284	0.026	0.982
16	0.595	0.930	0.014	0.297	0.025	0.983
17	0.587	0.934	0.013	0.309	0.024	0.983
18	0.579	0.939	0.013	0.320	0.023	0.984
19	0.572	0.942	0.012	0.330	0.022	0.985

Standard Errors of the Impulse Response Functions

Period	P-Shock 1		P-Shock 2		P-Shock 3		T Shock	
	Δm ₂	FF	Δm ₂	FF	Δm ₂	FF	Δm ₂	FF
1	0.287	0.040	0.009	0.021	0.191	0.011	0.030	0.016
2	0.214	0.053	0.019	0.021	0.094	0.017	0.033	0.021
3	0.102	0.043	0.046	0.023	0.056	0.029	0.050	0.021
4	0.076	0.042	0.068	0.026	0.047	0.041	0.075	0.025
5	0.056	0.052	0.096	0.030	0.041	0.056	0.094	0.028
6	0.051	0.056	0.131	0.028	0.030	0.060	0.115	0.025
7	0.049	0.056	0.167	0.032	0.027	0.062	0.138	0.024
8	0.041	0.061	0.193	0.032	0.022	0.067	0.162	0.026
9	0.043	0.064	0.226	0.032	0.023	0.065	0.184	0.024
10	0.044	0.064	0.258	0.034	0.020	0.060	0.205	0.022
11	0.043	0.066	0.279	0.034	0.022	0.058	0.221	0.022
12	0.043	0.066	0.298	0.034	0.025	0.054	0.233	0.021
13	0.043	0.064	0.313	0.034	0.027	0.050	0.243	0.019
14	0.046	0.063	0.322	0.033	0.029	0.048	0.245	0.019
15	0.046	0.059	0.327	0.032	0.030	0.046	0.247	0.018
16	0.045	0.056	0.327	0.031	0.031	0.044	0.247	0.018
17	0.044	0.053	0.324	0.029	0.032	0.043	0.245	0.017
18	0.044	0.049	0.320	0.028	0.032	0.042	0.240	0.017
19	0.043	0.046	0.313	0.027	0.032	0.041	0.240	0.017

Standard Errors of the Impulse Response Functions

$$G = \begin{bmatrix} 0.225 & 0.973 & -0.024 & -0.049 \\ 0.407 & -0.049 & -0.044 & 0.911 \\ 0.203 & -0.024 & 0.978 & -0.044 \\ 1.000 & -0.378 & 0.045 & -0.805 \end{bmatrix}$$

Table 6a: Δm₂-FF-cx

FIGURE 2

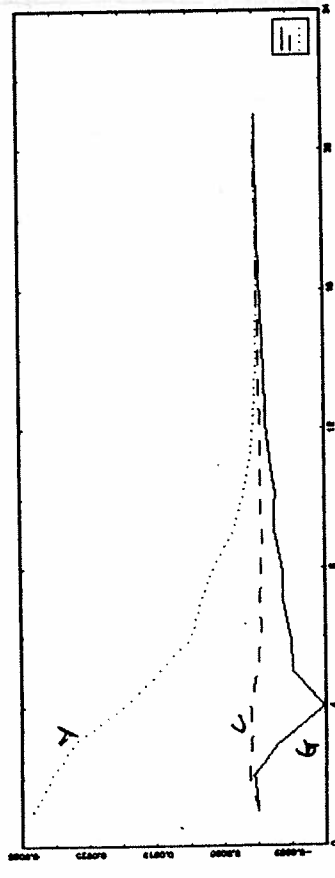
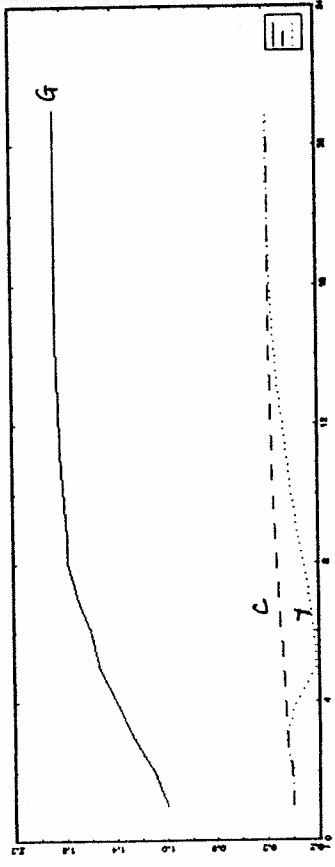
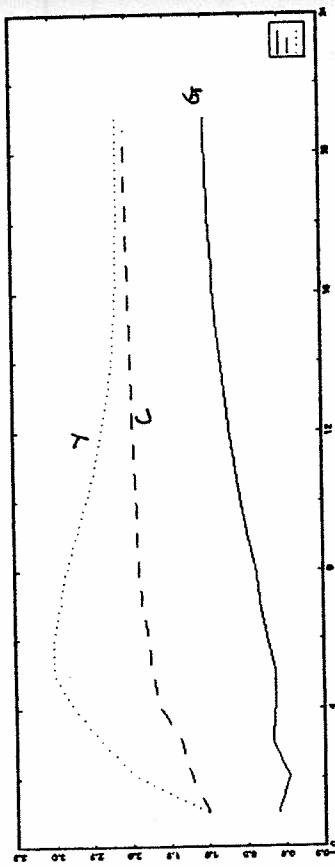
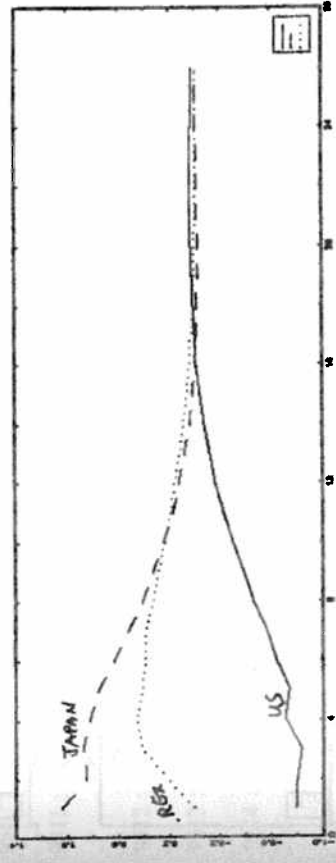
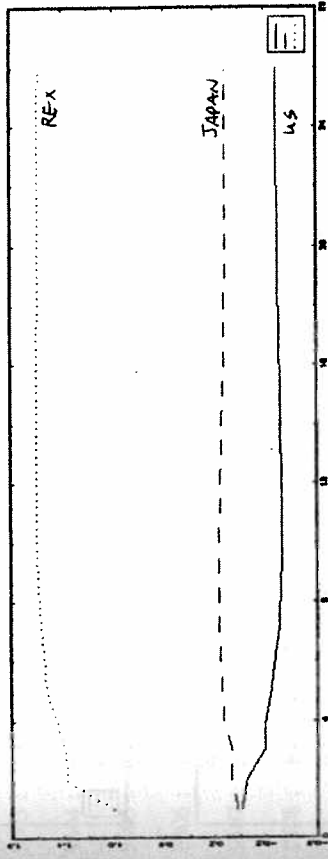
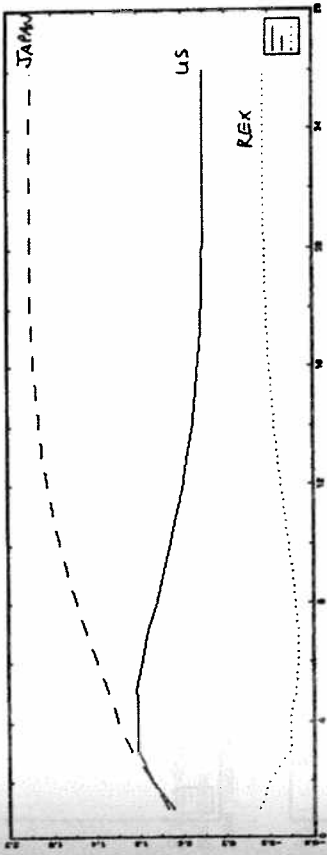


FIGURE 3



Cahiers de recherche (Discussion papers)
 1995 à aujourd'hui (1995 to date)

Si vous désirez obtenir un exemplaire, vous n'avez qu'à faire parvenir votre demande et votre paiement (5 \$ l'unité) à l'adresse ci-haut mentionnée. / To obtain a copy (\$ 5 each), please send your request and prepayment to the above-mentioned address.

- 9501 : Boyer, Marcel et Jean-Jacques Laffont, "Environmental Risks and Bank Liability", janvier 1995, 46 pages.
- 9502 : Margolis, David N., "Firm Heterogeneity and Worker Self-Selection Bias Estimated Returns to Seniority", décembre 1994, 29 pages.
- 9503 : Abowd, John M., Francis Kramarz et David N. Margolis, "High-Wage Workers and High-Wage Firms", janvier 1995, 73 pages
- 9504 : Cardia, Emanuela et Steve Ambler, "Indexation Lags and Heterodox Stabilization Programs", janvier 1995, 29 pages.
- 9505 : Garcia, René et Huntley Schaller, "Are the Effects of Monetary Policy Asymmetric?", février 1995, 42 pages.
- 9506 : Parent, Daniel, "Survot des contributions théoriques et empiriques liées au capital humain", février 1995, 70 pages.
- 9507 : Parent, Daniel, "Wages and Mobility : The Impact of Employer-Provided Training", février 1995, 34 pages.
- 9508 : Parent, Daniel, "Industry-Specific Capital and the Wage Profile : Evidence from the NLSY and the PSID", février 1995, 21 pages.
- 9509 : Parent, Daniel, "Matching, Human Capital, and the Covariance Structure of Earnings", février 1995, 54 pages.
- 9510 : Garcia, René, "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models", mars 1995, 50 pages.
- 9511 : Garcia, René, Annamaria Lusardi et Serena Ng, "Excess Sensitivity and Asymmetries in Consumption : An Empirical Investigation", mars 1995, 26 pages.
- 9512 : Sprumont, Yves, "An Axiomatization of the Pazner-Schmeidler Rules in Large Fair Division Problems", mars 1995, 26 pages.
- 9513 : Ghysels, Eric, Lynda Khalaf et Cosmé Vodounou, "Simulation Based Inference in Moving Average Models", mars 1995, 10 pages.
- 9514 : Ng, Serena, "Looking for Evidence of Speculative Stockholding in Commodity Markets", mars 1995, 25 pages.

FIGURE 4

