Testing for Cobreaking and Super exogeneity in the Presence of Deterministic Shifts

Hans-Martin Krolzig
Institute of Economics and Statistics, and Nuffield College, Oxford.
Juan Toro*
European University Institute, Florence

November 12, 2000

Abstract

We introduce a reduced rank technique for testing for common deterministic shifts. The reduced rank approach is analysed also in the context of super exogenity and an alternative test for super-exogeneity is proposed. One important advantage of this approach is that departing from the unrestricted model we depart from a more general model that does not impose a priori which are the target and the policy variables. This could be useful in case in which there did not existed an exact knowledge about a classification between target and instrument nor about their relationship. Monte Carlo simulations are implemented to investigate the power of this technique.
Keywords: Co-breaking, Super-Exogenity, Reduced Rank Regression, Regime Shifts, Markov Switching, Common Deterministic Shifts.

JEL classification: E32, E37, C32, E24

1 Introduction

Deterministic shifts in the conditional mean of economic variables is a recurrent feature in empirical economics. These shifts happen to affects not just one single economic variable but affect contemporaneously other related economic variables. Furthermore these shifts that repeat themselves in time, might be related linearly and this linear relationship might prevail throughout time. We have here proposed a technique that can be used to analyze such phenomena, and can help to gather important information about how breaks are related thought economic variables and across time. Frequently, deterministic shifts are induced by policy changes. Policy makers move the level of some variables in order to affect some target variables and reach specific goals. When deterministic shifts are induced by policy makers, the relationship between common deterministic shifts and super-exogeneity become apparent. Super exogeneity (see EngleR.F., .D.F. and Richard (1983)) establishes conditions under which the parameters of the partial model are invariant to changes in the parameters of the marginal model. On an economic context, the marginal model can be thought as an instrument that policy-makers can move(say interest rate) in order to achieve some goal. The partial model could be thought as the process for the goal variable (say inflation). Super-exogeneity sets the conditions under which the partial model has invariant parameters and can be used for policy analysis, despite changes in the marginal model. The concept of

*Preliminary: please do not quote without the authors' permission. We are grateful to David Hendry, Soren Johansen, Uwe Hassler and Bent Nielsen for useful comments and discussions. Financial support from the UK Economic and Social Research Council under grant L116251015 is gratefully acknowledged by the first author.

common deterministic shifts are super-exogenity are hence closely related if we limit the set of policy-makers interventions to changes in the conditional mean of the marginal process(say the level of interest rates or the rate of growth of money).

In the next section we define the concept of common deterministic shifts. In section 3 we define the model and introduce a reduced rank technique to estimate and test for common deterministic shifts. In section 4 the size and power of the technique are investigated with a Monte Carlo simulation experiment. Testing for super exogeneity based on the existence of common deterministic shifts is discussed in section 5. Section 6 concludes.

2 The concept of common deterministic shifts:

Engle and Kozicki (1993) have recently proposed the idea of common features in time series. This idea is inspired by the concept of cointegration introduced in Granger (1986) and Engle and Granger (1987). Engle and Kozicki (1993) show that a feature is common to a set of time series if a linear combination of them do not have the feature though each of the series individually have it. Some particular examples of this concept are the idea of common cycle introduced by Engle and Kozicki (1993) and co-breaking introduced by Hendry (1997). The concept of co-breaking is closely related to the idea of cointegration: while cointegration removes unit roots from linear combinations of variables, co-breaking can eliminate the effects of regime shifts by taking linear combinations of variables.

Definition 1. *Consider* $\{x_t\}$ *to be a n dimensional vector process, where is modeled as an* $VAR(k)$*,* $A(L)x_t = \mu_t + \varepsilon_t$ *We say that the equations in the* VAR *are subject to common deterministic shifts(CDS) if shifts taking place across the* n *individual equations are linearly related.*

In the definition of common deterministic shifts we just require that shifts are related across variables and throughout time, which can be expressed as a convenient reduced rank condition in the coefficients of the interventions variables. This concept is milder that the co-breaking concept of Hendry (1997) which requires that linear combinations of variables cancel the shifts in the process itselfshifts. In the following we consider the n-dimensional linear Gaussian VAR(p):

$$x_t = \bar{\mu} + \sum_{i=1}^{p} A_i x_{t-i} + \mu_t + \varepsilon_t \tag{1}$$

where $\varepsilon_t \sim \mathsf{NID}(0, \Sigma)$ and the roots of the vector autoregressive polynomial are within the unit circle, $\left| \mathbf{I} - \sum_{i=1}^{p} A_i z^i \right| = 0 \Longrightarrow |z| > 1$. Thus there are no unit roots in the system and possible non-stationarity are due to the deterministic breaks. This implies that the process possess the infinite vector moving average representation

$$x_t = \left(\mathbf{I} - \sum_{i=1}^{p} A_i \right)^{-1} \bar{\mu} + \sum_{i=0}^{\infty} \Psi_i \mu_{t-i} + \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}$$

where $A(L)\Psi(L) = \mathbf{I}$. Note that in the case of a VAR(1) we have that $\Psi_i = A^i$.

Consider now the $n \times T$ matrix $\mathbf{M} = (\mu_1 \ \mu_2 \ \cdots \ \mu_T)$ where $T > n$. The condition for common deterministic shifts can be written as $\Phi' M_T^1 = \mathbf{0}$. Thus we have that $r = \mathrm{rank}\,[\mathbf{M}] < n$ is necessary and sufficient for $\Phi'\mu_t = 0$ for all $t \in \mathcal{T}$ where $\Phi \neq \mathbf{0}$ is $n \times (n-r)$. The methodology proposed in this paper relies on using appropriate shift dummies for known dates. Consider a set of $n \leq s < T$ dummy variables d_{t_i} each of which is zero except for unity at times $t \in \mathcal{T}_i$, such that $\mu_t = \sum_{i=1}^{s} \boldsymbol{\mu}_i d_{t_i}$, or $\mu_t = \mathcal{M} d_t$, where $\mathcal{M}$ is $n \times s$ and d_t is $s \times 1$. Then $\mathbf{M} = (\mathcal{M}d_1 : \mathcal{M}d_2 : \cdots : \mathcal{M}d_T) = \mathcal{M}D$. It is assumed that the points at which these shifts occur are known which avoids the problem of nuisance parameters.

In order to illustrate the reduced rank approach consider a VAR(1) with shifts in the intercept:

$$x_t = \sum_{i=1}^{p} A_i x_{t-i} + \sum_{j=1}^{s} \boldsymbol{\mu}_j d_{t_j} + \varepsilon_t. \tag{2}$$

Suppose furthermore that the shifts are permanent, then we can use the corresponding shift dummy to model them as $d_{t_j} = I\left(t > t_j\right)$,where $I(\bullet)$ is the indicator function and $1 < t_j < T$.

CDS is at least of order r if there exist r linearly independent vectors satisfying $\phi_i' \mu_t = \phi_i' \mathcal{M} d_t = 0$ such that the $n \times r$ matrix $\Phi = (\phi_1 : \cdots : \phi_r)$ has rank r. Then $\Phi' \mathcal{M} = \mathbf{0}$ so rank$(\mathcal{M}) \leq r$, so the nullity of $\mathcal{M}$ determines the order of CDS. Thus CDS implies that $\mathcal{M}$ is of reduced rank; $\mathcal{M}$ can be decomposed to the product of two matrices of full rank,η and ξ.

$$x_t = \bar{\mu} + \sum_{i=1}^{p} A_i x_{t-i} + \eta \xi' D_t + \varepsilon_t, \tag{3}$$

Furthermore, note that the matrices η and ξ are not unique without suitable normalization, since if H is any $r \times r$ non-singular matrix, then $\mathcal{M} = \eta\xi$ implies that $\eta^* \xi^{*\prime} = (\eta H)(H^{-1}\xi) = \mathcal{M}$ as well. If common deterministic shifts is a particularity of the data, the coefficient matrix of the dummy regressors would have a reduced rank and the vectors that link the shifts across processes would be the outcome of an eigenvalue problem.

3 Estimating CDS vectors by reduced rank regressions

3.1 The reduced-rank regression problem

Maximum likelihood estimation of the CDS$(n-r)-$VAR$_n(p)$ is close to the analysis of the likelihood in cointegrating systems, and both are based in the reduced rank regression technique introduced in Anderson (1958) and Tso (1981). We follow the notation in the Johansen (1995) reduced-rank regression approach to cointegration except for the decomposition of the loading and the linear relationship across breaks which we refer to η and ξ, respectively. The analogy with the cointegration model is straightforward if one bare in mind that the regime-dummies d_t behave like a non-stationary process if there are structural breaks. In this case the matrix $\mathcal{M}$ then determines how the non-stationarity feed into the variables of the systems: the rank r of matrix $\mathcal{M}$ gives the number of *common deterministic breaks*, and the CDS rank $n-r$ gives the dimension of the space whose one-step predictions are free from *deterministic breaks.*

In contrast to the cointegration problem, however, the number of breaks s is not necessarily identical to the number of endogenous variables in the system, such that the matrix $\mathcal{M}$ is $n \times s$ with rank $r \leq \min(n, s)$.

In matrix notation we have:

$$X = \mathcal{B}Z + \mathcal{M}D + E \tag{4}$$

where $X := (x_1 : x_2 : \cdots : x_T)$ is $n \times T$, $Z := (z_1 : z_2 : \cdots : z_T)$ is $(1+p)n \times T$ with $z_t := (\mathbf{1} : x_{t-1}' : \ldots x_{t-p}')'$, $\mathcal{B} : =(\bar{\mu} : A_1, \ldots, A_p)$ is $n \times (1+p)n$, D is $s \times T$, and $\mathcal{M} = \eta\xi'$ is $n \times s$. The log-likelihood function for a sample size T is easily seen to be

$$\ln L = -\frac{nT}{2} \ln 2\pi - \frac{T}{2} \ln |\Sigma| - \frac{1}{2} \text{tr} \left[(X - \mathcal{B}Z - \eta\xi' D)' \Sigma^{-1} (X - \mathcal{B}Z - \eta\xi' D) \right] \tag{5}$$

3.2 Estimation of $\mathcal{B}$ and Σ conditional on $\eta\xi$

Note that for any fixed η and ξ the maximum of $\ln L$ is obtained for

$$\tilde{\mathcal{B}}(\eta\xi') = (X - \eta\xi' D)Z'(ZZ')^{-1}. \tag{6}$$

If we substitute $\mathcal{B}$ by (6) in (5), we get

$$\ln L = -\frac{nT}{2}\ln 2\pi - \frac{T}{2}\ln|\Sigma| - \frac{1}{2}\mathrm{tr}\left[(XH - \eta\xi' DH)'\Sigma^{-1}(XH - \eta\xi' DH)\right] \tag{7}$$

Hence, we just have to maximize this expression with respect to $\eta,\ \xi'$ and Σ.

For given η and ξ, the maximum is obtained if

$$\tilde{\Sigma}(\eta\xi') = T^{-1}(XH - \eta\xi' DH)(XH - \eta\xi' DH)'.$$

is substituted for Σ. Consequently we must maximize:

$$-\frac{T}{2}\ln\left|T^{-1}(XH - \eta\xi' DH)(XH - \eta\xi' DH)'\right| \tag{8}$$

or, equivalently, minimize the determinant with respect to η and ξ (see Lütkepohl (1991)).

3.3 Estimation of η conditional on ξ

Note that in (7) X and D are corrected for Z. Define the corresponding residuals as R_{0t} and R_{1t} :

$$\begin{aligned}
\underset{(T\times T)}{H} &: = (\mathbf{I}_T - Z'(ZZ')^{-1}Z),\\
\underset{(n\times T)}{R_X} &: = XH,\\
\underset{(s\times T)}{R_D} &: = DH,
\end{aligned}$$

and the corresponding moment matrices as:

$$S_{ij} = T^{-1}R_i R_j' \text{ for } i,j = X, D.$$

Then (8) can be rewritten as

$$-\frac{T}{2}\ln\left|T^{-1}(R_x - \eta\xi' R_D)(R_x - \eta\xi' R_D)'\right| \tag{9}$$

For fixed ξ, (8) is maximized with respect to matrix η by regression:

$$\begin{aligned}
\tilde{\eta}(\xi) &= R_X\left(\xi' R_D\right)'\left[(\xi' R_D)(\xi' R_D)'\right]^{-1}\\
&= R_X R_D'\xi\left[\xi'(R_D R_D')\xi\right]^{-1}\\
&= S_{XD}\xi(\xi' S_{DD}\xi)^{-1}
\end{aligned} \tag{10}$$

3.4 Estimation of ξ

Apart from a constant, the concentrated log-likelihood for our reduced rank problem can be shown to be:

$$
\begin{aligned}
&\ln L\left(\tilde{\eta}(\xi), \xi, \tilde{\Sigma}\left(\tilde{\eta}(\xi)\xi'\right)\right) \\
&= -\frac{T}{2}\ln\left|\tilde{\Sigma}\left(\tilde{\eta}(\xi)\xi'\right)\right| \\
&= -\frac{T}{2}\ln\left|\frac{1}{T}(R_X - R_X R_D'\xi\left[\xi'(R_D R_D')\xi\right]^{-1}\xi' R_D)(R_x - R_X R_D'\xi\left[\xi'(R_D R_D')\xi\right]^{-1}\xi' R_D)'\right| \\
&= -\frac{T}{2}\ln\left|\frac{1}{T}R_X(\mathbf{I} - R_D'\xi\left[\xi'(R_D R_D')\xi\right]^{-1}\xi' R_D)(\mathbf{I} - R_D'\xi\left[\xi'(R_D R_D')\xi\right]^{-1}\xi' R_D)' R_X'\right| \\
&= -\frac{T}{2}\ln\left|\frac{1}{T}R_X\left[\mathbf{I} - \left(\xi' R_D\right)'\left[\left(\xi' R_D\right)\left(\xi' R_D\right)'\right]^{-1}\left(\xi' R_D\right)\right]R_X'\right| \\
&= -\frac{T}{2}\ln\left|S_{XX} - S_{XD}\xi(\xi' S_{DD}\xi)^{-1}\xi' S_{DX}\right|.
\end{aligned}
\tag{11}
$$

Using the identity

$$
\begin{aligned}
\begin{vmatrix} S_{XX} & S_{XD}\xi \\ \xi' S_{DX} & \xi' S_{DD}\xi \end{vmatrix} &= \left|\xi' S_{DD}\xi\right|\left|S_{XX} - S_{XD}\xi(\xi' S_{DD}\xi)^{-1}\xi' S_{DX}\right| \\
&= \left|S_{XX}\right|\left|\xi' S_{DD}\xi - \xi' S_{DX} S_{XX}^{-1} S_{XD}\xi\right|
\end{aligned}
$$

equation 11 can be expressed as:

$$
\begin{aligned}
\ln L\left(\tilde{\eta}(\xi), \xi, \tilde{\Sigma}\left(\tilde{\eta}(\xi)\xi'\right)\right) &= -\frac{T}{2}\ln\frac{\left|S_{XX}\right|\left|\xi' S_{DD}\xi - \xi' S_{DX} S_{XX}^{-1} S_{XD}\xi\right|}{\left|\xi' S_{DD}\xi\right|} \\
&= -\frac{T}{2}\ln\frac{\left|S_{XX}\right|\left|\xi'\left(S_{DD} - S_{DX} S_{XX}^{-1} S_{XD}\right)\xi\right|}{\left|\xi' S_{DD}\xi\right|}.
\end{aligned}
$$

Hence, the maximum of $\ln L$ is given by

$$
\min_{\xi}\frac{\left|\xi'\left(S_{DD} - S_{DX} S_{XX}^{-1} S_{XD}\right)\xi\right|}{\left|\xi' S_{DD}\xi\right|}
$$

and following a basic theorem of matrix analysis (see, for example,. Johansen, 1995, Lemma A.8), this factor is minimized among all $n \times r$ matrices ξ by solving the eigenvalue problem

$$
\left|\rho S_{DD} - \left(S_{DD} - S_{DX} S_{XX}^{-1} S_{XD}\right)\right| = 0
$$

or, for $\lambda = 1 - \rho$, the eigenvalue problem

$$
\left|\lambda S_{DD} - S_{DX} S_{XX}^{-1} S_{XD}\right| = 0
$$

for eigenvalues λ_i and eigenvectors v_i, such that

$$
\lambda_i S_{DD} v_i = S_{DX} S_{XX}^{-1} S_{XD} v_i.
$$

If we normalize ξ such that $\xi' S_{DD}\xi = \mathbf{I}_r$ then the vectors of $\tilde{\xi}$ are given by the eigenvectors corresponding to the r smallest eigenvalues of $S_{DD} - S_{DX} S_{XX}^{-1} S_{XD}$. The maximum log-likelihood under the rank$(\mathcal{M}) = r$ restriction is given by:

$$
\max \ln L = -\frac{nT}{2}\ln 2\pi - \frac{T}{2}\left[\ln\left|S_{XX}\right| + \sum_{i=1}^{r}\ln\left(1 - \widehat{\lambda}_i\right)\right] - \frac{nT}{2} \tag{12}
$$

since by choice of $\tilde{\xi}$ we have that $\xi' S_{DD}\xi = \mathbf{I}_r$, as well as $\xi' S_{DX} S_{XX}^{-1} S_{XD}\xi = \operatorname{diag}(\widehat{\lambda}_1, \ldots, \widehat{\lambda}_r)$

3.5 Testing for the CDS rank

Since CDS($n-r$) implies CDS($n-r-1$), it seems natural to seek the maximum degree of CDS. In general two cases have to be distinguished: (i) the number of breaks s is less than the dimension of the system $n,\ m = \min(s,n) = s < n$. In the second case, the number of breaks s is not less than the dimension of the system n, i.e. $m = \min(s,n) = n \leq s$.

Suppose in the following that $n \leq s$. Then the following hypotheses might be of interest:

I. CDS(n) : $rank(\mathcal{M}) = r$, $r = 0$:. No breaks.
II. CDS($n-r$): $rank(\mathcal{M}) = r$, $0 < r < n$. There are breaks common to both process.
III. CDS(0): $rank(\mathcal{M}) = r$, $r = n$. There are breaks independent to each process.

Following Anderson (1951), the likelihood ratio test statistic for testing the CDS(r) against the CDS(n) is given such that the likelihood ratio statistic is given by:

$$-2\ln Q(\ H(r) \mid H(n)) = T\sum_{i=r+1}^{m}\ln\left(1+\widehat{\lambda}_i^2\right)$$

which has a χ^2-distribution with degrees freedom equal to $(n-r)(s-r)$.

3.6 Representation theorem

One of the advantages of reduced rank regression for analyzing common deterministic shifts is that, once η and ξ have been found, we can get rid of the shifts by appropriate rotation and conditioning. Common deterministic shifts imply that $\mathcal{M}$ is of reduced rank. Thus $\mathcal{M}$ can be decomposed in the product of two matrices of full rank, η and ξ. Once we have estimated η and ξ, we can transform the model into new variables in the space of the common break and in its orthogonal complement.

Consider the CDS-VAR(1) where we drop the intercept for simplicity,

$$x_t = Ax_{t-1} + \mathcal{M}d_t + \varepsilon_t. \tag{13}$$

Let us introduce the matrices $\eta_\perp$ and $\overline{\eta}$, where $\eta_\perp$ is a $p\times(p-r)$ matrix orthogonal to η, such that η' $\eta_\perp = 0_{rx(p-r)}$, and $\overline{\eta} = \eta(\eta'\eta)^{-1}$. We can multiply through by $\eta_\perp'$ and $\overline{\eta}'$, in order to obtain:

$$\begin{aligned}\eta_\perp' x_t &= \eta_\perp' Ax_{t-1} + \eta_\perp'\varepsilon_t\\ \overline{\eta}'x_t &= \overline{\eta}'Ax_{t-1} + \xi' d_t + \overline{\eta}'\varepsilon_t\end{aligned}$$

If we define the new variables $\widetilde{y_t} = \eta_\perp' X_t$ and $\widetilde{z}_t = \overline{\eta}'X_t$, then the conditional model $\widetilde{y}_t|\widetilde{z}_t$ can be expressed as:

$$\eta_\perp' x_t|\overline{\eta}'x_t = \eta_\perp' Ax_{t-1} + \varpi\overline{\eta}'x_t - \varpi\overline{\eta}'Ax_{t-1} + \widetilde{\varepsilon}_t \tag{14}$$

where

$$\begin{aligned}\widetilde{\varepsilon}_t &= \eta_\perp'\varepsilon_t - \varpi\overline{\eta}'\varepsilon_t\\ \varpi &= \Pi_{12}\Pi_{22}^{-1},\end{aligned}$$

with $\Pi_{12} = \eta_\perp'\Sigma\overline{\eta}$ and $\Pi_{22} = \overline{\eta}'\Sigma\overline{\eta}$.

Invoking the decomposition:

$$\eta\overline{\eta}' + \overline{\eta}_\perp\eta_\perp' = \mathbf{I},$$

and inserting it in 14,we get:

$$\eta_\perp' x_t | \overline{\eta}' x_t = \eta_\perp' A \left[\eta\overline{\eta}' + \overline{\eta}_\perp \eta_\perp' \right] x_{t-1} + \varpi\overline{\eta}' x_t - \varpi\overline{\eta}' A \left[\eta\overline{\eta}' + \overline{\eta}_\perp \eta_\perp' \right] x_{t-1} + \widetilde{\varepsilon}_t$$

And the resulting conditional model is free of shifts. That is:

$$\widetilde{y}_t = \left[\eta_\perp' - \varpi\overline{\eta}' \right] A\eta\widetilde{z_{t-1}} + \varpi\widetilde{z}_t + \left[\eta_\perp' - \varpi\overline{\eta}' \right] A\overline{\eta}_\perp\widetilde{y_{t-1}} + \widetilde{\varepsilon}_t. \tag{15}$$

4 A Monte Carlo analysis of the reduced rank regression technique to estimate common deterministic shifts:

In this section we analyze the size and power of the rank test for common deterministic shifts. The data generation process(DGP) will be given by the two dimensional process with two breaks in the intercept term at time t_1 and t_2:

$$x_t = Ax_{t-1} + v + \Phi D_t + \varepsilon_t$$

where $x_t = \begin{pmatrix} y_t \\ z_t \end{pmatrix}$, with $D_t = \begin{pmatrix} d_{t_1} \\ d_{t_2} \end{pmatrix}$, $\varepsilon_t \sim NID\left(0, \Omega\right)$ and $\Omega = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ without loss of generality.

As experimental design variables we have: the matrix A, the matrix Φ, the sample size T and the points at which the breaks occur t_1 and t_2. For simplicity A will have the following structure:

$$A = \begin{pmatrix} 0.75 & 0.5 \\ 0 & \alpha \end{pmatrix}$$

That is z_t is strongly exogenous. The benchmark case will have $\alpha = 0.8$. We would also be interested in analyzing the size of the test when the process for z_t, becomes close to the unit root. That is, $\alpha = 0.95$, $\alpha = 0.975$ and $\alpha = 0.99$. The matrix Φ embeds information about the size of the shift, the relationship of the shift across variables and the relationship of the shifts across time. We let the relationship between the breaks change. That is,

$$\Phi = k \begin{pmatrix} 1 \\ \frac{\eta_1}{\eta_2} \end{pmatrix} \begin{pmatrix} 1 & \frac{\xi_1}{\xi_2} \end{pmatrix}$$

where $\frac{\eta_1}{\eta_2}(\frac{\eta_1}{\eta_2} = 0.25, 0.5 \text{ and } 1)$, define the relationship of the breaks across equations, $\frac{\xi_1}{\xi_2}(\frac{\xi_1}{\xi_2} = 0.25, 0.5$ and $1)$ define the relationship of breaks across time and $k(k = 1, 1.5, 2)$ define the magnitude in terms of the standard deviations.

The sample size, T, is 50,100 and 150. $t_1 = \tau_1 T$ with $\tau_1 = (0.30, 0.31, ..., 0.6)$ and $t_2 = \tau_2 T$ with $\tau_2 = 0.70$. How the distance between breaks affect the size. The number of replications is $N = 10000$.

If we rely on a full factorial design the number of experiments would be 18225.

We think that the different hypothesis could be interpreted as follows:

- $rank(\Phi) = 0$. No break
- $rank(\Phi) = 1$. There are breaks common to both process
- $rank(\Phi) = 2$(full rank). There are breaks independent to each process.

As the benchmark case we will take the following values:

- $A = \begin{pmatrix} 0.75 & 0.5 \\ 0 & 0.8 \end{pmatrix}$
- $k = 2$ and $\frac{\xi_1}{\xi_2} = \frac{\eta_1}{\eta_2} = 0.25$
- $\alpha = 0.8$
- $\tau_1 = 0.3$ and $\tau_1 = 0.7$
- $T = 50, 100$ and 150

Figure 1 plots the histogram of the test statistic for $T = 50, 100$ and 150, which should resemble a Chi square with 1 degree of freedom. Table 1 present the size and power of the test for the three different sample sizes at the 5 % level of significance. We next analyze how the position of the breaks may affect the size of the test. Thus for the benchmark case we let τ_1 vary from 0.3 to 0.6 with the second break fixed at $\tau_2 T$ with $\tau_2 = 0.7$. The results for the size of the test for the three different sample sizes($T = 50, 100$ and 150) are shown in figure 2. The results for the power of the test for the three different sample sizes($T = 50, 100$ and 150) are shown in figure 3. A similar analysis is done with $\alpha = 0.95$, $\alpha = 0.975$ and $\alpha = 0.99$ and are presented in figure 2. Finally we depart again from the benchmark case and allow k, $\frac{\xi_1}{\xi_2}$ and $\frac{\eta_1}{\eta_2}$ to change. The results are presented in table 2.

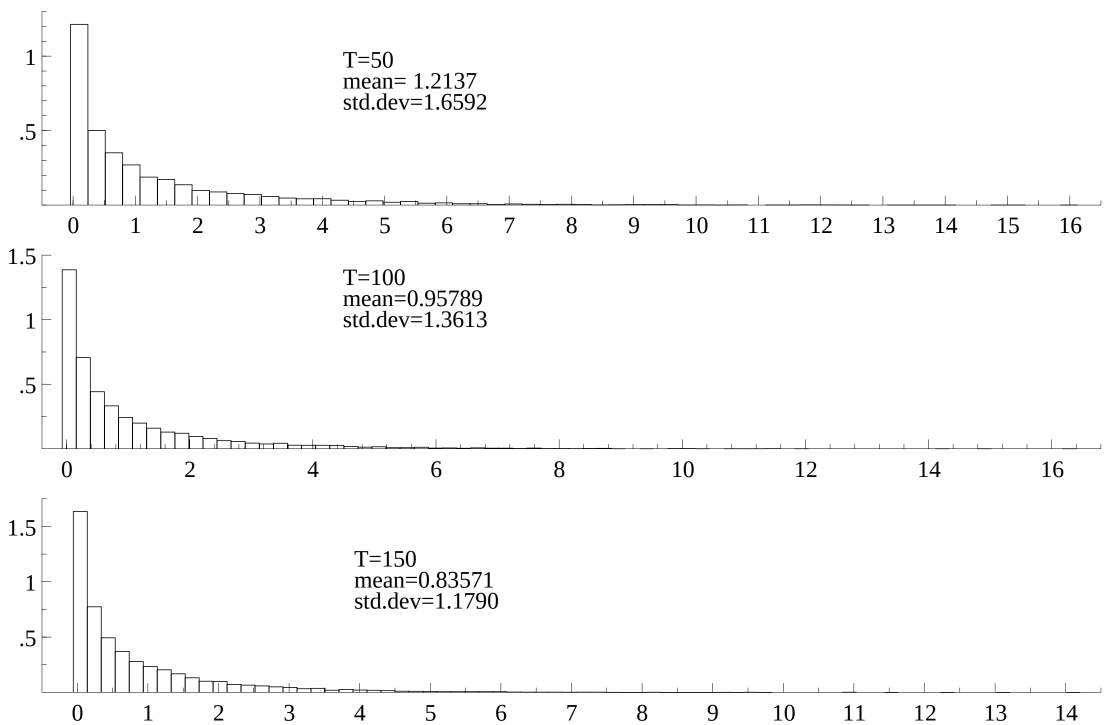

Figure 1 Histogram of the test for the reduced rank of Φ.

5 Testing for super exogeneity

5.1 The unrestricted system

Before arguing how common deterministic shifts in the conditional mean are related to the concept of super-exogenity, let us introduce the definitions of weak and super-exogeneity. Let us consider a

Table 1   Size and power of the test statistic for the benchmark DGP.

	T	
Size		
	50	0.077
	100	0.047
	150	0.031
	T	
Power		
	50	0.80
	100	0.94
	150	0.98

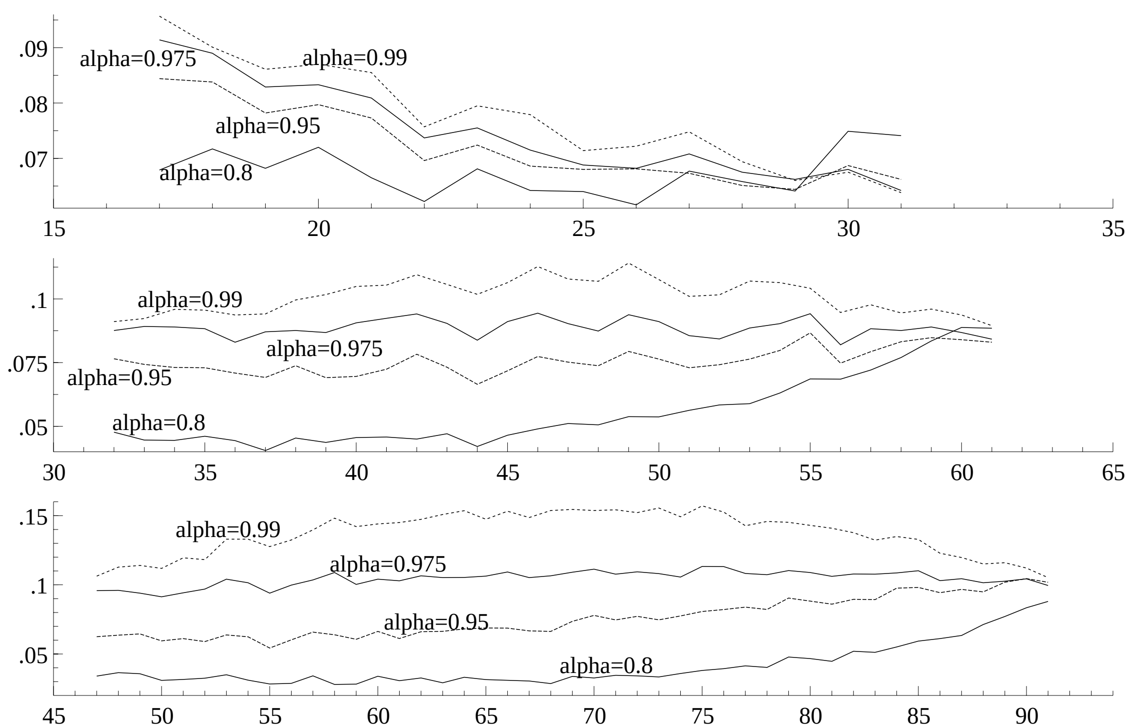

Figure 2   Size of the test statistic for different values of α and τ_1 varying from 0.3 to 0.6.

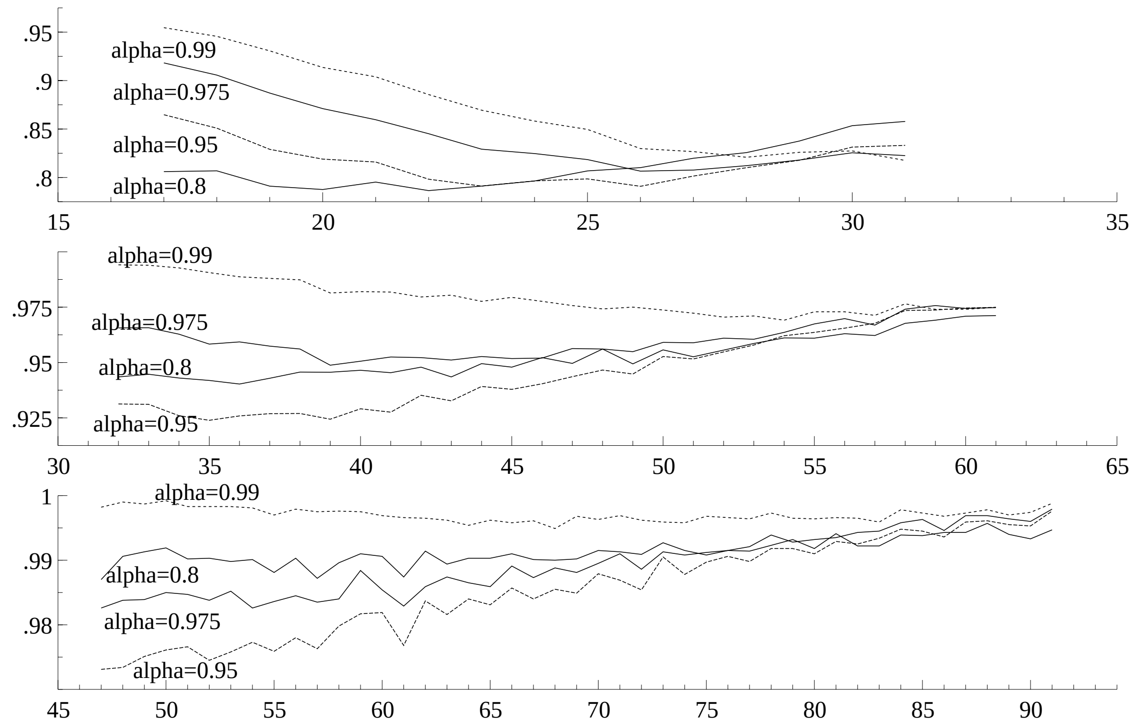

Figure 3 Power of the test statistic for different values of α and τ_1 varying from 0.3 to 0.6.

Table 2 Size and power of the test statistic for the benchmark case and allowing k, $\frac{\xi_1}{\xi_2}$ and $\frac{\eta_1}{\eta_2}$ to vary.

T=50	Power Range	Size Range
k=1	(0.45,0.59)	(0.049,0.060)
k=1.5	(0.64,0.86)	(0.065,0.083)
k=2	(0.88,0.96)	(0.087,0.091)
T=100	Power Range	Size Range
k=1	(0.65,0.77)	(0.031,0.045)
k=1.5	(0.84,0.93)	(0.040,0.062)
k=2	(0.94,0.99)	(0.047,0.073)
T=150	Power Range	Size Range
k=1	(0.69,0.81)	(0.022,0.044)
k=1.5	(0.98,0.97)	(0.046,0.055)
k=2	(0.97 0.99)	(0.046,0.060)

statistical model for the bivariate process $x_t = (y_t, z_t)'$ with parameters $\theta \in \Theta$. We are interested in a function of θ, called the parameters of interest, i.e. $\tau = \tau(\theta)$.

Definition 2. *Weak exogeneity: The process z_t is called weakly exogenous for the parameter of interest τ if there exist a parametrization of the model such that:*

$$f(x_1,..,x_T,\alpha,\beta) = \prod_{t=1}^{T} h_t(z_t|x_1,..,x_{t-1},\alpha) \prod_{t=1}^{T} g_t(y_t|z_{t,}x_1,..,x_{t-1},\beta)$$

$(\alpha,\beta) \in AxB$

$\tau = \tau(\beta)$ *is identified.*

Definition 3. *Super exogeneity: The process z_t is called super exogenous for the parameter of interest $\tau(\beta)$ and the class $\digamma$ of interventions if there exist a parametrization of the model such that:*

$$f(x_1,..,x_T,\alpha_1,...,\alpha_T,\beta) = \prod_{t=1}^{T} h_t(z_t|x_1,..,x_{t-1},\alpha_t) \prod_{t=1}^{T} g_t(y_t|z_{t,}x_1,..,x_{t-1},\beta)$$

$(\alpha_1,...,\alpha_T,\beta) \in \digamma xB$

$\tau = \tau(\beta)$ *is identified.*

The previous section used reduced rank regressions as a modelling devise to find common deterministic shift features. In this section we present an alternative analysis of common deterministic shifts. Though the concept of common deterministic shifts was presented in terms of the unrestricted model, its most insightful applications had been shown within the conditional model. Hendry and Mizon (1998) have advanced two different situations in which common deterministic shifts could play an essential role in modelling. They refer to this situations as the *contemporaneous correlation* case and the *behavioral relation* case.

Reconsider the VAR(p) in equation (1). If we apply the partition $x_t = (y_t' : z_t')'$ and consider just one lag, $p = 1$, we have:

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} \nu_{yt} \\ \nu_{zt} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{z,t} \end{pmatrix} \tag{16}$$

If we consider all information as of time $t-1$ and denote it by $\mathbf{I}_{t-1}$, the unrestricted model can be written as:

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} \mid \mathbf{I}_{t-1} \sim \mathsf{NID}\left(\begin{pmatrix} \mu_{y,t} \\ \mu_{z,t} \end{pmatrix}, \begin{pmatrix} \Sigma_{yy} & \Sigma_{yz} \\ \Sigma_{zy} & \Sigma_{zz} \end{pmatrix} \right) \tag{17}$$

where $\mu_{y,t} := E\left(y_t \mid \mathbf{I}_{t-1}\right)$,$\mu_{z,t} = E\left(z_t \mid \mathbf{I}_{t-1}\right)$ and the intercept term ν is subject to regime shifts.

In the *contemporaneous correlation case* y_t can be seen as a policy variable whereas z_t is a instrument that policy makers can use in order to reach their goal in terms of y_t.The *behavioral relation* case refers to the situation in which agents form rational expectations (about z_t) and there is an interest in analyzing how changes in the expectations may affect the plan of the agents (y_t). In both cases common deterministic shifts is introduced to justify invariance of the conditional model due to changes in the marginal model. The existence of a specific linear relationship relating breaks ($\Sigma_{yz}\Sigma_{zz}^{-1}$) together with weak exogeneity (see Engle, Hendry and Richard, 1983) would introduce the necessary conditions in order to investigate policy analysis.

5.2 The conditional system

Valid conditioning requires weak exogeneity of the marginal process with respect to the parameters of interest in the conditional model. Using the normality of ε_t, the model in equation (16) can be expressed in terms of the conditional and marginal model as:

$$
\begin{aligned}
y_t \quad & | \quad z_t, \mathbf{I}_{t-1} \sim \mathsf{NID}\left(\mu_{y|z,t}, \Omega\right) && (18)\\
z_t \quad & | \quad \mathbf{I}_{t-1} \sim \mathsf{NID}\left(\mu_{z,t}, \Sigma_{zz}\right), && (19)
\end{aligned}
$$

where the density of y_t conditional on $z_t, \mathbf{I}_{t-1}$ is determined by:

$$
\begin{aligned}
\mu_{y|z,t} \quad &: \quad = \mathsf{E}\left(y_t \mid z_t, \mathbf{I}_{t-1}\right) = \mu_{y,t} + \Sigma_{yz}\Sigma_{zz}^{-1}\left(z_t - \mu_{z,t}\right)\\
\Omega \quad &: \quad = \mathsf{Var}\left(y_t \mid z_t, \mathbf{I}_{t-1}\right) = \Sigma_{yy} - \Sigma_{yz}\Sigma_{zz}^{-1}\Sigma_{zy}.
\end{aligned}
$$

Rewriting equation in terms on the corresponding mean, we get:

$$
\begin{aligned}
\begin{pmatrix} \mathbf{I} & -\Sigma_{yz}\Sigma_{zz}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} y_t \\ z_t \end{pmatrix} \quad = & \quad \begin{pmatrix} \mathbf{I} & -\Sigma_{yz}\Sigma_{zz}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix}\\
& + \begin{pmatrix} \mathbf{I} & -\Sigma_{yz}\Sigma_{zz}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \nu_{yt} \\ \nu_{zt} \end{pmatrix} + \begin{pmatrix} u_{y,t} \\ u_{z,t} \end{pmatrix} \qquad (20)
\end{aligned}
$$

where the variance matrix of the transformed residuals is block diagonal:

$$
u_t \sim \mathsf{NID}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Omega & 0 \\ 0 & \Sigma_{zz} \end{pmatrix}\right)
$$

This type of models are prone to suffer from the Lucas critique. That is, changes in the marginal model leads to non-constancy of the conditional model. Shifts of the marginal model can induce shifts in the conditional model, but a convenient linear combination can induce constancy in the conditional process:

$$
\begin{pmatrix} \mathbf{I} & -\Sigma_{yz}\Sigma_{zz}^{-1} \end{pmatrix} \begin{pmatrix} \nu_{yt} \\ \nu_{zt} \end{pmatrix} = \begin{pmatrix} \bar{\nu}_y \\ \bar{\nu}_z \end{pmatrix}. \qquad (21)
$$

If we model $\begin{pmatrix} \nu_{yt} \\ \nu_{zt} \end{pmatrix}$ with the corresponding intervention variables, equation 16 can be rewritten as:

$$
x_t = \bar{\mu} + \sum_{i=1}^{p} A_i x_{t-i} + \Phi D_t + \varepsilon_t, \qquad (22)
$$

Note that the condition in equation 21 requires that shifts in the process for z_t implies a shift in the process for y_t. Hence super exogeneity would imply a reduced rank condition on the coefficients of the intervention variables used to model the unrestricted system(Φ), such we can rewrite the previous model as

$$
x_t = \bar{\mu} + \sum_{i=1}^{p} A_i x_{t-i} + \eta\xi' D_t + \varepsilon_t,
$$

Furthermore in other for super exogeneity to hold the conditional model should be invariant to the set of interventions in the marginal process which would imply

$$
\eta_{\perp}' = \begin{pmatrix} \mathbf{I} & -\Sigma_{yz}\Sigma_{zz}^{-1} \end{pmatrix}
$$

So the reduced rank of the coefficient of the intervention dummies with specific restrictions on the null space of η and weak exogeneity imply super-exogeneity which postulates the invariance of the conditional model under a set of interventions in the marginal model.[1]

5.3 Testing procedure for super-exogeneity:

The previous subsection showed how super-exogeneity of the y_t process with respect to a set of interventions(shifts in the conditional mean of the marginal process) required a reduced rank condition of the coefficients of the intervention variables. In order to implement a likelihood ratio test for super exogeneity with respect to these class of interventions we need to estimate the model under the null(reduced rank of $\Phi = \eta\xi'$ and specific restrictions on η, $\eta = \begin{pmatrix} \Sigma_{yz}\Sigma_{zz}^{-1} \\ -I \end{pmatrix}$) and under the alternative(the unrestricted model with the reduced rank of $\Phi = \eta\xi'$ imposed). We propose three alternative procedures to implement the super exogeneity test. They differ in the way in which the model is estimated under the null:

5.3.1 First procedure:

Estimation of the model under the restrictions $\eta = \begin{pmatrix} \Sigma_{yz}\Sigma_{zz}^{-1} \\ I \end{pmatrix}$.

Let us depart from the model:

$$R_X = \Phi R_D + E$$

with the reduced rank restriction imposed such that $\Phi = \eta\xi'$,

$$R_X = \eta\xi' R_D + E \tag{23}$$

The maximum likelihood of this model under the super-exogeneity restriction can be calculated as follows:

(1) We depart from initial estimates of $\eta(\eta_0)$ and $\Sigma(\Sigma_0)$. In model given by equation 23 we can impose the restriction that $\begin{pmatrix} \Sigma_{yz,0}\Sigma_{zz,0}^{-1} \\ I \end{pmatrix} \subset \eta_0$ leaving the rest of the space of η_0 unrestricted. Under the null, if we multiply though by $\overline{\eta}_0'$,we get:

$$\overline{\eta}_0' R_X = \xi' R_D + \overline{\eta}_0' E$$

(2) We can apply OLS in the previous equation an obtain estimates of $\xi'(\xi_1')$.
(3) Given $\xi'(\xi_1')$ and $\Sigma(\Sigma_0)$ we can then obtain new estimates of $\eta(\eta_1)$ and $\Sigma(\Sigma_1)$.

We can loop in this algorithm till convergence where in each iteration the restriction $\begin{pmatrix} \Sigma_{yz}\Sigma_{zz}^{-1} \\ I \end{pmatrix} \subset \eta$ is always updated.

[1] Many models in economics are expressed in terms of rational expectations. They can be expressed as the behavioral relation:

$$E\left(y_t \mid \mathbf{I}_{t-1}\right) = \mu_y^* + \Psi_t E\left(z_t \mid \mathbf{I}_{t-1}\right)$$

or in short form: $\mu_{y,t} = \mu_y^* + \Psi_t \mu_{z,t}$. However, when Ψ_t is constant, the plan and the expectation have a common deterministic shift. In order to achieve constant relationship between plans and expectations it was needed that $(I : -\Psi)$ was the relationship that link shifts in the mean of the *marginal* process and the conditional process.

5.3.2 Second procedure:

An alternative estimation procedure of the model under the restrictions $\eta = \begin{pmatrix} \Sigma_{yz}\Sigma_{zz}^{-1} \\ I \end{pmatrix}$ can be based on the first order conditions of the likelihood function. Let us depart from the concentrated likelihood function,

$$\ln L = -\frac{nT}{2}\ln 2\pi - \frac{T}{2}\ln|\Sigma| - \frac{1}{2}\mathrm{tr}\Sigma^{-1}\left[(S_{XX} - S_{XD}\xi\eta' - \eta\xi' S_{DX} + \eta\xi' S_{DD}\xi\eta')\right]$$

We can write the linear restriction for ξ and η, as:

$$vec(\xi) = H\varphi + h$$

$$vec(\eta') = G\chi + g$$

The derivatives for the likelihood with respect to ξ and η are given by:

$$\frac{\partial lnL}{\partial\varphi} = H'vec(S_{DX}\Sigma^{-1}\alpha) - H'vec(S_{DD}\xi\eta'\Sigma^{-1}\eta) \tag{24}$$

$$\frac{\partial lnL}{\partial\chi} = G'vec(\xi' S_{DX}\Sigma^{-1}) - H'vec(\xi' S_{DD}\xi\eta'\Sigma^{-1}) \tag{25}$$

On substituting the restrictions of ξ in 24 we get:

$$\varphi(\chi,\Sigma) = \left[H'(\eta'\Sigma^{-1}\eta\otimes S_{DD})H\right]^{-1}\left[H'(\eta'\Sigma^{-1}\otimes I)vec(S_{DX}) - H'\eta\prime\Sigma^{-1}\eta\otimes S_{DD}h\right] \tag{26}$$

Similarly for χ we can substitute the restrictions for η in 25 and we get:

$$\chi(\varphi,\Sigma) = \left[G'(\Sigma^{-1}\otimes\xi' S_{DD}\xi)G\right]^{-1}\left[G'vec(\xi' S_{XD}\Sigma^{-1}) - G'\left(\Sigma^{-1}\otimes\xi' S_{DD}\xi\right)g\right]. \tag{27}$$

It can easily be seen that the first order condition for Σ for given χ and φ is given by:

$$\Sigma(\varphi,\chi) = S_{XX} - S_{XD}\xi\eta' - \eta\xi' S_{DX} + \eta\xi' S_{DD}\xi\eta' \tag{28}$$

Hence for initial φ and Σ we can imposed the restrictions on η and obtain estimates of ξ from 26. For given ξ and Σ new estimates of φ can be obtained from 27. For given ξ and η, equation 28 delivers new estimates of Σ. We can then iterate in this algorithm with the restriction $\eta = \begin{pmatrix} \Sigma_{yz}\Sigma_{zz}^{-1} \\ -I \end{pmatrix}$ always updated.

The likelihood ratio test can be shown to have a Chi squared distribution with $p - r$ degrees of freedom, where p is the dimension of the system an r is the rank of Φ. The degrees of freedom result just from comparing the tangent space of $\eta\xi'$ with and without restrictions.

5.3.3 Third procedure:

An alternatively testing procedure can be implemented just with linear regressions[2]. In order to show this alternative procedure let us depart from model 16

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \Phi \begin{pmatrix} d_{t1} \\ d_{t2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{z,t} \end{pmatrix} \tag{29}$$

where Φ is a matrix of coefficients of the intervention variables, $\Phi = \begin{pmatrix} \mu_{t1}^y & \mu_{t2}^y \\ \mu_{t1}^z & \mu_{t2}^z \end{pmatrix}$.

The conditional model is given by:

$$y_t = A_{11}y_{t-1} + A_{12}z_{t-1} + \mu_{t1}^y d_{t1} + \mu_{t2}^y d_{t2} + E(\varepsilon_{y,t}|\varepsilon_{z,t}),$$

or,

$$y_t = A_{11}y_{t-1} + A_{12}z_{t-1} + \mu_{t1}^y d_{t1} + \mu_{t2}^y d_{t2} + \omega(z_t - A_{21}y_{t-1} - A_{22}z_{t-1} - \mu_{t1}^z d_{t1} - \mu_{t2}^z d_{t2}) + \widetilde{\varepsilon}_{y,t},$$

with $\widetilde{\varepsilon}_{y,t} = \varepsilon_{y,t} - \omega\varepsilon_{z,t}$ which can be rewritten as:

$$y_t = \omega z_t + (A_{11} - \omega A_{21})y_{t-1} + (A_{12} - \omega A_{22})z_{t-1} + (\mu_{t1}^y - \omega\mu_{t1}^z)d_{t1} + (\mu_{t2}^y - \omega\mu_{t2}^z)d_{t2} + \widetilde{\varepsilon}_{y,t}$$

and the marginal model is given by:

$$z_t = A_{21}y_{t-1} + A_{22}z_{t-1} + \mu_{t1}^z d_{t1} + \mu_{t2}^z d_{t2} + \varepsilon_{z,t} \tag{30}$$

Under the super exogeneity condition we have the restrictions $(\mu_{t1}^y - \omega\mu_{t1}^z) = 0$ and $(\mu_{t2}^y - \omega\mu_{t2}^z) = 0$. Which are the reduced rank conditions of equation 21. These restrictions can also be written as $\frac{\mu_{t1}^y}{\mu_{t1}^z} = \frac{\mu_{t2}^y}{\mu_{t2}^z} = \omega$, where the specific restrictions on the null space of η becomes clearer. Under the null of super-exogeneity we have that the conditional process reduces to

$$y_t = \omega z_t + (A_{11} - \omega A_{21})y_{t-1} + (A_{12} - \omega A_{22})z_{t-1} + \widetilde{\varepsilon}_{y,t} \tag{31}$$

The parameters in the conditional model are $\theta_c = \{\omega, A_{11} - \omega A_{21}, A_{12} - \omega A_{22}, \Omega\} = \{\theta_c^1, \theta_c^2 \theta_c^3, \theta_c^4\}$ and the parameters in the marginal model are $\theta_m = \{A_{21}, A_{22}, \mu_{t1}^z, \mu_{t2}^z, \Sigma_{zz}\} = \{\theta_m^1, \theta_m^2, \theta_m^3, \theta_m^4, \theta_m^5\}$. The properties of the Gaussian distribution imply that the parameters in the marginal process are variation independent of the parameters in the conditional process. The two equations(equation 31 and 30) can be estimated separately and the full maximum likelihood estimate is made up of two factors corresponding to the marginal and conditional distribution. The maximum likelihood estimator of the unrestricted model is just the likelihood of model under the reduced rank restriction, which can be obtained by a reduced rank regression. The maximum likelihood of the unrestricted model is jus the estimated model with the reduced rank condition imposed.

6 Conclusions

This paper puts together two topics of research, common deterministic shifts and super-exogeneity issues. We have shown how common deterministic shifts can be analyzed with simple and widely known techniques, reduced rank regressions. Deterministic shifts in the conditional mean of economic variables is a recurrent feature in empirical economics. These shifts happen to affects not just one single

[2]This altenative procedure was suggested to us by Soren Johansen.

economic variable but affect contemporaneously other related variables. Furthermore these shifts that repeat themselves in time, might be related linearly and this linear relationship might prevail throughout time. We have here proposed a technique that can be used to analyze such phenomena, and can help to gather important information about how breaks are related thought economic variables and across time. Frequently, deterministic shifts are induced by policy changes. Policy makers move the level of some variables in order to affect some target variables and reach specific goals. When deterministic shifts are induced by policy makers, the relationship between common deterministic shifts and super-exogeneity become apparent.

One important advantage of this approach is that departing from the unrestricted model we depart from a more general model that does not impose a priori any relationship among shifts in the mean of the individual process. This could be useful in case in which there did not existed an exact knowledge about a classification between target and instrument nor about their relationship. Think of a monetary model where some short interest rates are included (say the discount rate and the interbank rate) together with some variables of interest that the policy maker wants to influence. The identification of the linear relationship linking shifts in the mean of the different processes would give valuable information both about the exact relationship between them and about the transmission (the weights of the linear relationship that govern shifts) .

The methodology developed in this paper is restricted by two major assumptions. First that, conditional on the breaks, the system is stationary which excludes integrated-cointegrated systems. Secondly that the breaks points are known a priori. But it should be not too difficult to overcome these recent limitations.

In case of cointegration the following system is of interest:

$$\Delta x_t = \alpha\beta' x_{t-1} + \eta\xi' D_t + \varepsilon_t$$

which maps the I(1) system into its vector equilibrium representation. The system can then be estimated by a switching algorithm which does reduced-rank estimations of (i) α, β conditional on η, ξ and (ii) η, ξ conditional on α, β. Thus one would combine the techniques developed in this paper with the well established cointegration analysis of Johansen (1995)(see Toro (1999)).

The main drawback of the reduced rank regression techniques proposed in this paper comes from the fact that the shift points are assumed to be known in advance. If the break points are unknown, then the Markov-switching approach provides a powerful tool to model the system.

$$x_t = Ax_{t-1} + \Phi D_t + \varepsilon_t,$$

where D_t contains the indicator functions. Krolzig (1997) considers the statistical analysis of this system when the (potentially reduced) rank of $\Phi = \eta\xi'$ is imposed to the system. In an LIML approach, each equation could be estimated separately. Assume that the number of regime is $M = 2$. Then $M - 1$ smoothed regime probabilities associated with each equation can be collected to the matrix D_t and be used in the reduced rank regression approach discussed above. Again one might consider the potential cointegration of the variables which suggests the combination of including the switching algorithm on each M-step of the EM algorithm for MS-VAR models.

References

Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Annals of Mathematical Statistics*, **22**, 327–5.

Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* New York: John Wiley & Sons.

Engle, R. F., and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica*, **55**, 251–276.

Engle, R. F., Hendry, D. F., and Richard, J.-F. (1983). Exogeneity. *Econometrica*, **51**, 277–304. Reprinted in Hendry D. F. (1993), *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Engle, R., and Kozicki, S. (1993). Testing for common features. *Journal of Business Economic and Statistics*, **11**, 369–395.

EngleR.F., .D.F., H., and Richard, J. (1983). Exogeneity. *Econometrica*, **51**, 277–304.

Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics*, **48**, 213–228.

Hendry, D. (1997). A theory of co-breaking. Manuscript, Institute of Economics and Statistics, Oxford.

Hendry, D., and Mizon, G. (1998). Exogeneity,causality and co-breaking in economic policy analysis of a small econometric money of money in the uk. *Empirical Economics*, **23**, 267–295.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Krolzig, H.-M. (1997). *Markov Switching Vector Autoregressions. Modelling, Statistical Inference and Application to Business Cycle Analysis*. Berlin: Springer.

Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*: Berlin: Springer.

Toro, J. (1999). Modelling common deterministic shifts in $I(1)$ variables. Mimeo, Department of Economics, European University Institute.

Tso, K. K.-S. (1981). Reduced-rank regression and canonical analysis. *Journal of the Royal Statistical Society B*, **43**, 183–189.