# Testing for Super-Exogeneity in the Presence of Common Deterministic Shifts

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

October 30, 2001

**Abstract**

This paper introduces the concept of common deterministic shifts (CDS). This concept is simple, intuitive and relates to the common structure of shifts or policy interventions. We propose a Reduced Rank technique to investigate the presence of CDS. The proposed testing procedure has standard asymptotics and good small-sample properties. We further link the concept of CDS to that of super-exogeneity. It is shown that CDS tests can be constructed which allow to test for super-exogeneity. The Monte Carlo evidence indicates that the CDS test for super-exogeneity dominates testing procedures proposed in the literature.

*Keywords:* Co-breaking, Super-Exogeneity, Reduced Rank Regression, Regime Shifts, Markov Switching.

*JEL classification:* E32, E37, C32, E24

## 1 Introduction

Deterministic shifts in the conditional mean of economic variables are a recurrent feature in empirical economics. These shifts happen to affect not just one single economic variable but, contemporaneously, also other related economic variables. Furthermore, some economic time series are affected by several shifts. These shifts might be related linearly and this linear relationship might prevail throughout time. Here we propose a technique that can be used to analyze such phenomena, and can help to gather important information about how breaks are related through economic variables and across time.

Frequently, deterministic shifts are induced by policy changes. Policymakers move the level of some variables in order to affect some target variables and reach specific goals. When deterministic shifts are induced by policymakers, the relationship between common deterministic shifts and super-exogeneity become apparent. Super-exogeneity (see Engle, Hendry and Richard, 1983) establishes conditions under which the parameters of the partial model are invariant to changes in the parameters of the marginal model. In an economic context, the marginal model can be thought of as the instrument of the policymaker (say, the interest rate). The partial model could be thought of as the process for the goal variable (say, the rate of inflation). Super-exogeneity sets conditions under which the partial model has invariant parameters, which allows its use for policy analysis despite changes in the marginal model. The

*We are grateful to the guest editor, Stephane Gregoir, and two anonymous referees for valuable comments and suggestions on a previous version of this paper. We also thank 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 L138251009 is gratefully acknowledged by the first author.

concepts of common deterministic shifts and super-exogeneity are closely related if we limit the set of policymakers interventions to changes in the conditional mean of the marginal process (say, the level of interest rates or the rate of growth of money).

The structure of the paper is as follows: In the next section, we introduce the concept of common deterministic shifts. In §3 we define the model and introduce a reduced rank technique to estimate and test for common deterministic shifts. The size and power of the proposed technique are investigated in §4 with a Monte Carlo simulation experiment. §6 discusses the concept of super-exogeneity in the framework of common deterministic shifts. In §5, testing procedures for super-exogeneity are developed which are based on the existence of common deterministic shifts. Their small-sample properties are analyzed in a small Monte Carlo experiment discussed in §7. The empirical applicability of the introduced tests is illustrated in §8, which also comments on their limitations and possible extensions. §9 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 does not have the feature though each of the series individually has it. Some particular examples of this concept are the idea of common cycles, introduced by Engle and Kozicki (1993), and the idea of 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 an $n$ dimensional vector process, which is modeled as a vector autoregressive (VAR) process of order $p$, $A(L)x_t = \mu_t + \varepsilon_t$. 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 $s \times 1$ with $D_t = (d_{t,1} : d_{t,2} : \cdots : d_{t,s})'$. The condition for common deterministic shifts can be written as $\Phi' \mathcal{M} = \mathbf{0}$. Thus we have that $r = \operatorname{rank}[\mathcal{M}] < \mathrm{n}$ is necessary and sufficient for $\Phi' \mu_t = 0$ for all $t \in \mathcal{T}$, where $\Phi \neq \mathbf{0}$ is $n \times (n-r)$. 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 be related across variables and through time, which can be expressed as a convenient reduced rank condition in the coefficients of the intervention variables. This concept is different from the co-breaking concept of Hendry (1997) which requires that linear combinations of variables cancel the shifts in the processes. In the following, we consider the $n$-dimensional linear Gaussian VAR($p$):

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

where $\varepsilon_t \sim \mathsf{NID}(\mathbf{0}, \Sigma)$ 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 the possible non-stationarity is due to the deterministic breaks. This implies that the process possesses the infinite-order vector moving-average representation

$$x_t = \sum_{i=0}^{\infty} \Psi_i (\varepsilon_{t-i} + \mu_{t-i}),$$

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

The methodology proposed in this paper relies on using appropriate shift dummies for known dates. It is assumed that the points at which these shifts occur are known, which avoids the problem of some nuisance parameters. In order to illustrate the reduced rank approach consider a VAR(1) with shifts in the intercept:

$$x_t = A_1 x_{t-1} + \sum_{i=1}^{s} \boldsymbol{\mu}_i d_{t,i} + \varepsilon_t. \tag{2}$$

Furthermore, suppose that the shifts are permanent. Then we can use the corresponding shift dummies to model them: $d_{t,i} = I\left(t > t_i\right)$ , where $I(\bullet)$ is the indicator function and $1 < t_i < T$.

CDS is at least of order $(n-r)$ if there exist $(n-r)$ linearly independent vectors $\phi_i$ satisfying

$$\phi_i' \mu_t = \phi_i' \mathcal{M} D_t = 0$$

such that the $n \times (n-r)$ matrix $\Phi = (\phi_1 : \cdots : \phi_{(n-r)})$ has rank $(n-r)$. Then $\Phi' \mathcal{M} = \mathbf{0}$ so rank $(\mathcal{M}) < n$ and the nullity of $\mathcal{M}$ determines the order of CDS. In other words, CDS implies that $\mathcal{M}$ is of reduced rank. Thus, $\mathcal{M}$ can be decomposed into the product of two matrices, $\eta$ and $\xi$ of full rank $r$, such that we can rewrite (2) as:

$$x_t = A_1 x_{t-1} + \eta \xi' D_t + \varepsilon_t. \tag{3}$$

Note that the matrices $\eta$ and $\xi$ are not unique without suitable normalization: if $H$ is any $r \times r$ non-singular matrix, then $\mathcal{M} = \eta\xi'$ implies that $\eta H H^{-1} \xi' = \eta H (\xi H'^{-1})' = \eta^* \xi^{*\prime} = \mathcal{M}$ as well. If common deterministic shifts are a property of the data, the coefficient matrix of the dummy regressors will have a reduced rank.

# 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 (1951) and Tso (1981). The analogy with the cointegration model is straightforward if one bears in mind that the regime-dummies $d_t$ behave like non-stationary processes if there are structural breaks. In this case the matrix $\mathcal{M}$ determines how the non-stationarity feeds into the variables of the system: the rank $r$ of the matrix $\mathcal{M}$ gives the number of *common deterministic shifts*, and the CDS rank ($n-r$ or $s-r$) gives the dimension of the space whose one-step predictions are free from these *deterministic breaks*. In contrast to the cointegration problem, 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 $pn \times T$ with $z_t' := (x_{t-1}' : \cdots x_{t-p}')$, $\mathcal{B} : =(A_1 : \cdots : A_p)$ is $n \times pn$, $E = (\varepsilon_1 : \varepsilon_2 : \cdots : \varepsilon_T)$ is $n \times T$ and $D$ is $s \times T$.[1]

[1] Note that, for reasons of simplicity, we assume here that the variables have a zero mean before the first break. If the mean of the variables is non-zero before the first break, an intercept $\nu$ can be added to the model, such that $z_t' := (1 : x_{t-1}' : \cdots x_{t-p}')$ and $\mathcal{B} : =(\nu : A_1 : \cdots : A_p)$ is $n \times (1 + pn)$.

$\mathcal{M}$ is $n \times s$ and of rank $r$ and can be decomposed in two full rank matrices, $\mathcal{M} = \eta\xi'$, where $\eta$ are the loadings and $\xi$ the linear relationships across breaks.

The log-likelihood function for a sample size $T$ is easily seen to be

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

### 3.2 Estimation of $\mathcal{B}$ and $\Sigma$ conditional on $\mathcal{M}$

Note that for any fixed $\eta$ and $\xi$, the maximum of $\ln L$ is obtained by

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

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

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

where $H := (\mathbf{I}_T - Z'(ZZ')^{-1}Z)$ is a $T \times T$ matrix. Hence, we just have to maximize this expression with respect to $\eta$ and $\Sigma$. For given $\eta$ and $\xi$, the maximum is obtained for

$$\tilde{\Sigma}(\eta\xi') = T^{-1}(X - \eta\xi' D)H(X - \eta\xi' D)'.$$

Consequently we must maximize:

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

or, equivalently, minimize the determinant with respect to $\eta$ and $\xi$ (see Lütkepohl, 1991).

### 3.3 Estimation of $\eta$ conditional on $\xi$

Note that in (7), $X$ and $D$ are corrected for $Z$. Define the corresponding residuals as:

$$\begin{aligned} (n \times T)R_X &: = XH, \\ (s \times T)R_D &: = DH, \end{aligned}$$

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 $\xi$, (8) is maximized with respect to matrix $\eta$ 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 $\xi$

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

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

By 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 $\xi$ 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 $\lambda_i$ and eigenvectors $v_i$, such that

$$\lambda_i S_{DD}v_i = S_{DX}S_{XX}^{-1}S_{XD}v_i. \qquad (12)$$

If we normalize $\xi$ such that $\xi' S_{DD}\xi = \mathbf{I}_r$ then the vectors of $\tilde{\xi}$ are given by the eigenvectors corresponding to the $r$ largest 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:

$$\max\ln L = -\frac{nT}{2}\ln 2\pi - \frac{T}{2}\left[\ln\left|S_{XX}\right| + \sum_{i=1}^{r}\ln\left(1 - \widehat{\lambda}_i\right)\right] - \frac{nT}{2}, \qquad (13)$$

since, by choice of $\tilde{\xi}$, we have that $\xi' S_{DD}\xi = \mathbf{I}_r$, as well as $\xi' S_{DX}S_{XX}^{-1}S_{XD}\xi = \mathrm{diag}(\widehat{\lambda}_1,\ldots,\widehat{\lambda}_r)$,where $\widehat{\lambda}_1,\ldots,\widehat{\lambda}_r$ are the $r$ largest eigenvalues solutions of (12).

### 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. In the first case , the number of potential breaks $s$ is less than the dimension of the system $n,\ m=\min(s,n)=s<n$. In the second case, the number of potential 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($n$) : $\operatorname{rank}(\mathcal{M})=0$. No breaks.
(ii) CDS($n-r$): $\operatorname{rank}(\mathcal{M})=r$, $0<r<n$. There are breaks which are common to the processes.
(iii) CDS(0): $\operatorname{rank}(\mathcal{M})=n$. There are as many linearly independent breaks as variables in the system.

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

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

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

## 4 Some Monte Carlo results

### 4.1 Design of the Monte Carlo

In this section we analyze the size and power of the rank test for common deterministic shifts. The data generation process (DGP) is a two-dimensional process $x_t=(y_t:z_t)'$ where $z_t$ does not feedback into the first equation and there are two breaks in the intercept term at time $t_1$ and $t_2$,

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0.75 & 0.5 \\ 0 & \alpha \end{pmatrix}\begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix}+\mathcal{M}\begin{pmatrix} d_{t_1} \\ d_{t_2} \end{pmatrix}+\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{14}$$
$$\text{where } \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}\sim \mathsf{NID}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix},\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

The matrix of intervention coefficients $\mathcal{M}$ embeds information about the size of the shifts, and the relationship of the shifts across variables and time. The interpretation of the model will critically depend on the assumed rank of the matrix $\mathcal{M}$:

(i) $\operatorname{rank}(\mathcal{M})=0$. No breaks.
(ii) $\operatorname{rank}(\mathcal{M})=1$. There are breaks which are common to both processes.
(iii) $\operatorname{rank}(\mathcal{M})=2$ (full rank). There are breaks independent to each process.

We consider two specifications for $\mathcal{M}$. In the first specification, $\mathcal{M}$ is of full rank:

$$\mathcal{M}_1=k\begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix}.$$

The second specification for $\mathcal{M}$, denoted $\mathcal{M}_2$, has reduced rank:

$$\mathcal{M}_2=k\begin{pmatrix} 1 \\ \varphi \end{pmatrix}\begin{pmatrix} 1 & \pi \end{pmatrix}.$$

This corresponds to the case of CDS, where $\varphi$ defines the relationship of the breaks across equations, $\pi$ defines the relationship of breaks across time and $k$ defines the magnitude of the breaks measured in terms of the standard deviation. In the Monte Carlo experiment, we analyze the properties of the CDS test when we allow $\varphi$, $\pi$ and $k$ to vary. We will consider the following values: $\varphi \in \{0.25, 0.5, 1\}$ , $\pi \in \{0.25, 0.5, 1\}$ and $k \in \{1, 1.5, 2\}$.

The benchmark case will have $\alpha = 0.8$, but one might be interested in analyzing the size of the test when the process for $z_t$ becomes close to the unit root. That is, $\alpha \in \{0.95, 0.975, 0.99\}$. Furthermore, in order to study how the distance between breaks affects the CDS test, we consider different distances between the breaks. As the break points are defined as fractions of the sample size, $t_1 = \tau_1 T$ and $t_2 = \tau_2 T$, we will parametrically vary the first break point, $\tau_1 \in \{0.30, 0.31, \ldots, 0.6\}$, while keeping the second break point fixed, $\tau_2 = 0.7$.

We consider three different sample sizes: $T = 50, 100$ and $150$. The number of replications is $N = 10000$. If we had relied on a full factorial design, the number of experiments would have been 18225. We therefore focus on two benchmark DGPs in our Monte Carlo study: $\mathsf{DGP}_1$ is characterized by $(\mathcal{M}_1; k = 2; \alpha = 0.8; \tau_1 = 0.3, \tau_2 = 0.7)$ and will be used to study the power properties of the CDS test. $\mathsf{DGP}_2$ is defined by $(\mathcal{M}_2; \varphi = 1, \pi = 1, k = 2; \alpha = 0.8; \tau_1 = 0.3, \tau_2 = 0.7)$ and allows to investigate the size properties of the CDS test given the reduced rank of its matrix of intervention coefficients.

### 4.2 Evaluation of the Monte Carlo

Table 1 presents the rejection frequencies of the CDS test for our simulated $\mathsf{DGP}_2$ for nominal sizes from $1\%$ to $20\%$. It can be seen that the real size differs from the nominal size and this difference cannot be explained by Monte Carlo sampling variation. Table 2 presents the rejection frequencies for the test statistic $H_0 : \mathrm{rank}(\mathcal{M}) = 1$ against $H_1 : \mathrm{rank}(\mathcal{M}) = 2$ when the model simulated is $\mathsf{DGP}_1$. This delivers the power of the CDS test for our benchmark case.

Table 1 Size of the CDS rank test.

| T | 20% | 10% | 5% | 1% |
|---|---|---|---|---|
| 50 | 0.243 | 0.133 | 0.068 | 0.015 |
| 100 | 0.231 | 0.123 | 0.064 | 0.016 |
| 150 | 0.222 | 0.117 | 0.062 | 0.012 |

*Note:* Rejection frequencies for the test statistic $H_0 : \mathrm{rank}(\mathcal{M}) = 1$ against $H_1 : \mathrm{rank}(\mathcal{M}) = 2$ for the benchmark $\mathsf{DGP}_2$.

Table 2 Power of the CDS rank test.

| T | 20% | 10% | 5% | 1% |
|---|---|---|---|---|
| 50 | 0.954 | 0.909 | 0.810 | 0.673 |
| 100 | 0.996 | 0.989 | 0.949 | 0.924 |
| 150 | 0.999 | 0.998 | 0.985 | 0.978 |

*Note:* Rejection frequencies for the test statistic $H_0 : \mathrm{rank}(\mathcal{M}) = 1$ against $H_1 : \mathrm{rank}(\mathcal{M}) = 2$ for the benchmark $\mathsf{DGP}_1$.

Figure 1 compares the theoretical and empirical quantiles of the CDS test statistic for the reduced rank of $\mathcal{M}$: $H_0 : \mathrm{rank}(\mathcal{M}) = 1$ against $H_1 : \mathrm{rank}(\mathcal{M}) = 2$. The simulated model is the benchmark $\mathsf{DGP}_2$ and the sample sizes are $T = 50, 100$ and $150$. The distribution of the test follows asymptotically a $\chi^2$ with one degree of freedom. The results indicate that the empirical size is slightly higher than the nominal size, but the simulated quantiles approach the theoretical quantiles of the asymptotic distribution with increasing sample size relatively fast for $T \geq 50$.

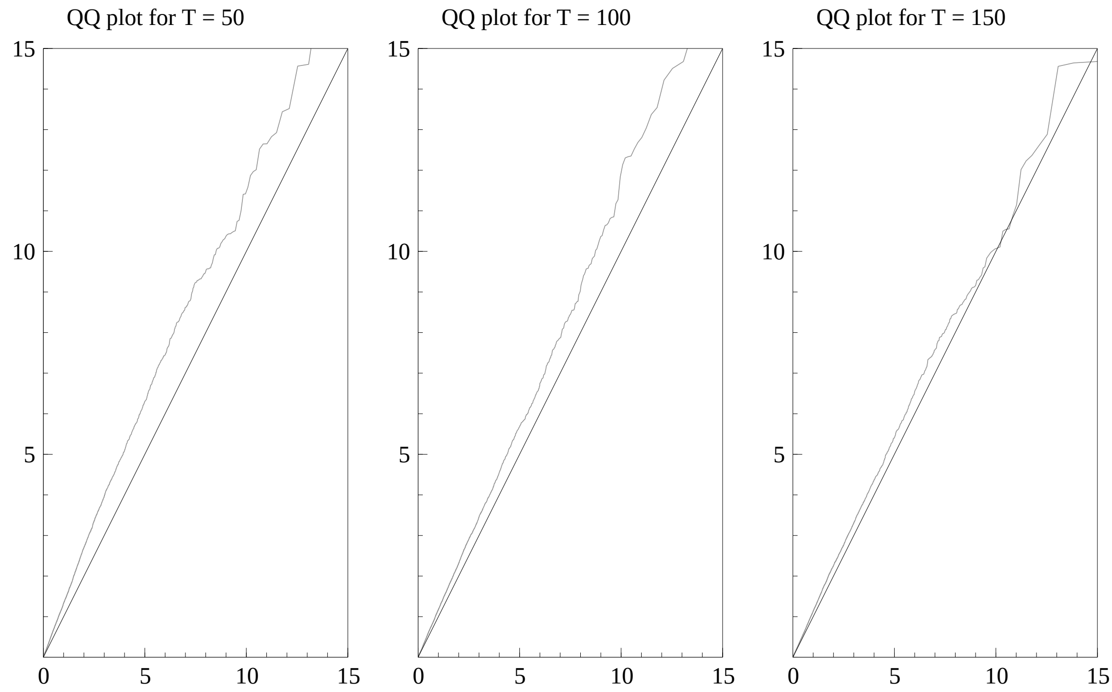


Figure 1 QQ plot of the CDS test for the reduced rank of $\mathcal{M}$.

In order to compare the power of the CDS test against its true size we will make use of size-power curves. We thus construct two different experiments: in the first experiment the DGP satisfies the null of common deterministic shifts, and in the second experiment the matrix of interventions coefficients has full rank.

Let us denote the test statistic in replication $j$ by $\vartheta_j$. With $\vartheta_j$ there is an associated $p$ value ($p_j = p(\vartheta_j)$), obtained from our asymptotic distribution, which is $\chi^2(1)$. For any point $x_i$ in the $(0, 1)$ interval we can define the empirical distribution function (EDF) of $p_j$ as:

$$F(x_i) = \frac{1}{N} \sum_{j=1}^{N} I(p_j < x_i).$$

The EDF $F(x_i)$ reports the probability of getting a $p$ value less than $x_i$ under the null. Following the suggestion of Davidson and MacKinnon (1997) the EDF is evaluated at $m = 215$ points $x_i$, $i = 1, \ldots, m$, where $m = 215$ and $x_i \in \{0, 0.001, 0.002, \ldots, 0.999\}$. We can then obtain a $p$-value plot plotting $F(x_i)$ against $x_i$. If the test is well-behaved (i.e. the test is rejected just about the right proportion of times) the plot should be close to the 45 degree line.

If we denote by $F^*(x_i)$ the EDF under the alternative, and the same random numbers are used when simulating the errors under the null and the alternative, then the size-power curve is obtained plotting $F^*(x_i)$ against $F(x_i)$.

The $p$-value plots in Figure 2 help to summarize and complement the information contained in Tables 1 and 2 and Figure 1. In the left panel of Figure 2, a $p$-value plot of the test statistic of $H_0 : \mathrm{rank}(\mathcal{M}) = 1$

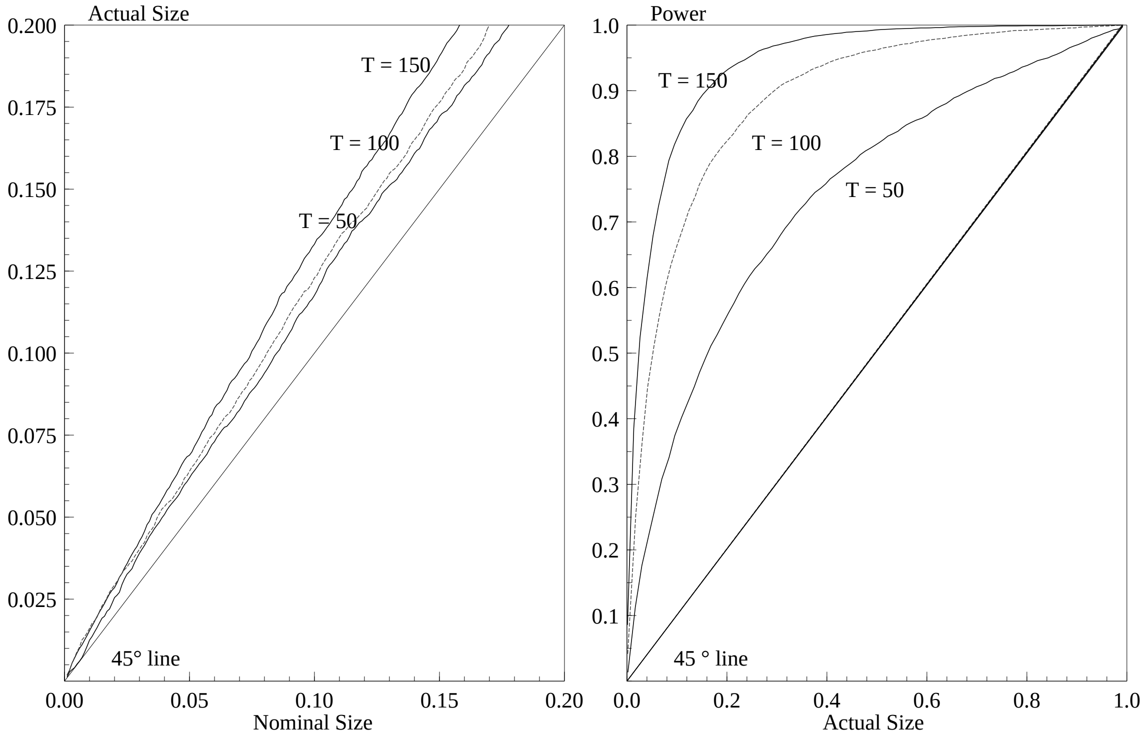


Figure 2 Size-power curve analysis.

against $H_1 : \mathrm{rank}(\mathcal{M}) = 2$ is plotted which is based on the empirical distribution of the $p$-values of our test statistic when the simulated DGP is the benchmark $\mathsf{DGP}_2$. We can see that the CDS test tends to overreject slightly. This overrejection tends to decrease quickly as the sample size increases with the $p$-value plot approaching the 45 degree line. The plot is truncated for values of $x_i > 0.2$ since relative small test sizes are of interest. The right panel of Figure 2 depicts the corresponding power-size curve. The power always exceeds hugely the actual size and is excellent for $T = 100$ and $150$.

The robustness of the previous results is analyzed in Figure 3, which presents evidence of the size-power trade-off under parametric variations of the benchmark $\mathsf{DGP}_2$. Each plot contains the size-power curve (bold line) for the CDS test statistic for the presence of one common deterministic shift, $H_0 : \mathrm{rank}(\mathcal{M}) = 1$ against $H_1 : \mathrm{rank}(\mathcal{M}) = 2$ for three sample sizes $T = 50, 100$ and $150$. For each parameterization of the DGP, the size-power curve is compared with the size-power curve in the benchmark $\mathsf{DGP}_2$ (thin line).

Plots [A] and [B] compare the size-power curves in the benchmark $\mathsf{DGP}_2$ with $k = 1.5$ and $k = 1$ respectively. Plot [C] compares the size-power curves in the benchmark $\mathsf{DGP}_2$ with that obtained when $\varphi = 0.5$. Plot [D] visualizes the effects of $\pi = 0.5$. Finally, Plot [E] graphs the size-power curve obtained when $k = 1$ and $\varphi = \pi = 0.5$. From the plots, we can see that the power for a given size remains largely unchanged with different parameterization of the benchmark DGP. There is, however, a substantial reduction in power relative to our benchmark $\mathsf{DGP}_2$ when the size of the breaks is lowered as in the case of $k = 1$ and $\varphi = \pi = 0.5$.

We next analyze how the position of the breaks may affect the size and the power of the CDS test by changing the point of occurrence of the first break, allowing $\tau_1$ to vary from $0.3$ to $0.6$ of the sample ($\tau_1 \in \{0.30, 0.31, \ldots, 0.6\}$). The second breaks is always fixed at $t_2 = \tau_2 T$ for $\tau_2 = 0.7$. The resulting size and power is analyzed for four different values of $\alpha$ in the benchmark $\mathsf{DGP}s$,

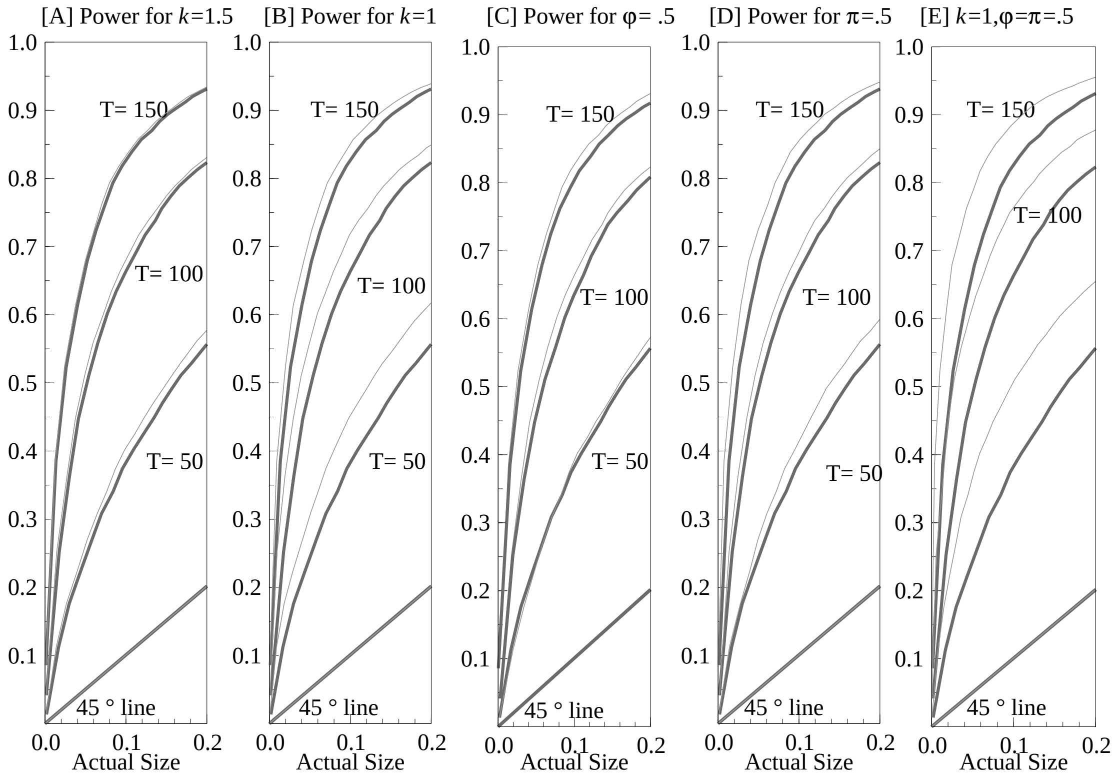


Figure 3 Size-power curve analysis for the CDS rank test.

$\alpha \in \{0.8, 0.95, 0.975, 0.99\}$, which allows the investigation of the behavior of the test as the process for $z_t$ gets closer to a unit root process. As in the previous experiments, three different sample sizes ($T = 50, 100$ and $150$) are considered.

The Monte Carlo results shown in Figures 4 and 5 depict the size and power of the CDS test for a given nominal size of $5\%$. In these plots the vertical axis measures the size and the power, respectively, and the horizontal axis measures the location of the first break as a fraction of the sample size. When interpreting the plots, the reader should be aware that Figure 5 depicts the power for the given *nominal* size. From the graphs one can see that the actual size clearly deviates from the nominal size when $\alpha$ gets closer to unity. Furthermore, for $\alpha = .8$ the actual size increases rapidly as the break points get closer to the second break point. For $T = 50$ power increases with greater distance from the second break. The power is uniformly strong for $T \geq 100$. Overall, the Monte Carlo results are very promising. The asymptotic $\chi^2$ distribution of the CDS test is found to provide a good approximation of the small-sample properties of the test.

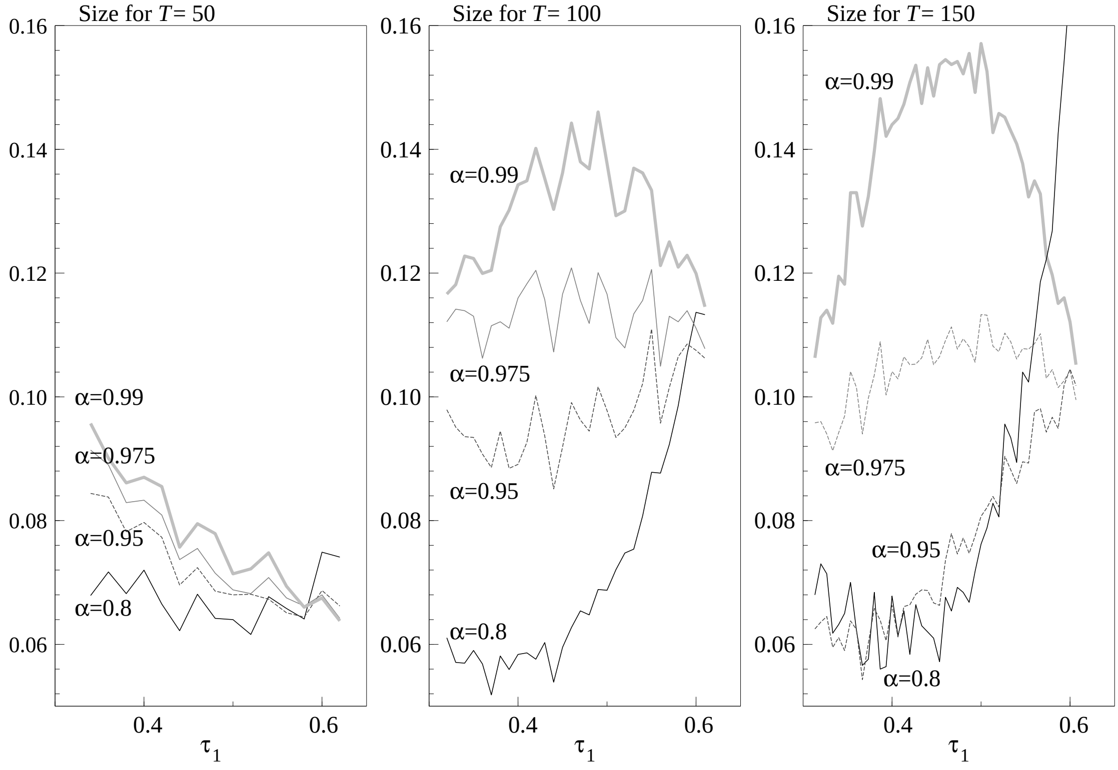


Figure 4 Size of the test statistic for different values of $\alpha$ and $\tau_1$ varying from $0.3$ to $0.6$.

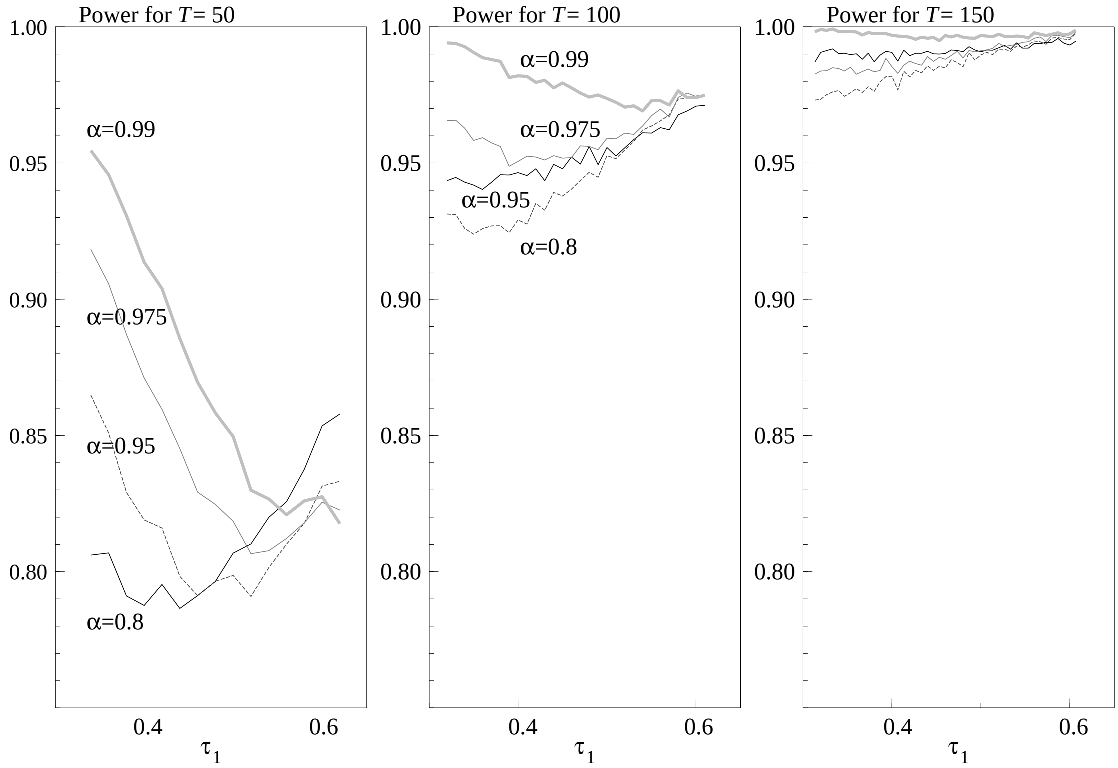


Figure 5 Power of the test statistic for different values of $\alpha$ and $\tau_1$ varying from $0.3$ to $0.6$.

# 5 Super-exogeneity in the Presence of Common Deterministic Shifts

The previous section used reduced rank regressions as a modelling device for finding common deterministic shift features. In this section we present an alternative analysis. Though the concept of common deterministic shifts has been introduced in terms of an unrestricted model, its most insightful application involves conditional models.

## 5.1 The concept of super-exogeneity

Before we discuss how common deterministic shifts in the conditional mean are related to the concept of super-exogeneity, let us introduce the definitions of weak and super-exogeneity. We will consider a statistical model for the vector process $x_t = (y_t' : z_t')'$ with parameters $\theta \in \Theta$. We are interested in a function of $\theta$, called the parameters of interest, i.e. $\vartheta = \vartheta(\theta)$.

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

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

*where $(\alpha, \beta) \in A \times B$ and $\vartheta = \vartheta(\beta)$ is identified.*

In order to define super-exogeneity we need to define the class of interventions affecting the DGP. We define the class of interventions under consideration, as any action $\lambda_t \in \Lambda_t$ by an agent from his available action set $\Lambda_t$, which alters $\theta$ from its current value to a different value $\theta_t = g(\lambda_t, \theta)$. The class of interventions[2] $\Lambda$ can then be formally defined as:

$$\Lambda = \{\lambda_t : \theta_t = g(\lambda_t, \theta),\ \lambda_t \in \Lambda_t\}.$$

**Definition 3.** *Super-exogeneity: The process $z_t$ is called super exogenous for the parameter of interest $\vartheta(\beta)$ and the class $\Lambda$ of interventions if there exists a parameterization of the model such that:*

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

*where $(\alpha_1, ..., \alpha_T, \beta) \in \Lambda \times B$ and $\vartheta = \vartheta(\beta)$ is identified.*

Super-exogeneity establishes conditions under which the Lucas critique can be refuted. Consider $y_t$ as the variable about which agents form plans given all information $\mathcal{I}_{t-1}$ available at time $t-1$. We would be interested in analyzing the invariance of $\beta$ when the marginal process changes. Favero and Hendry (1992) distinguishes two levels of the critique applicable to $\beta$.

*Level A*: At this level of the critique, $\beta$ may change due to changes in economic policy control rules.

*Level B* : At this level of the critique, $\beta$ might vary because of changes in economic environment that alter expectations.

At *level A*, changes in the distribution of variables that are under the control of a policymaker lead to variations in the parameters of empirical models. In this case future expectations are not necessarily involved. A case of this critique is the Lucas (1975) supply function. The *level B* of the critique corresponds to the use of backward-looking econometric specifications when agents use forward-looking specifications.

[2] Among the class of interventions that we will consider are those interventions induced by policymakers through changes of the level of some variables under their control.

### 5.2 The model

Consider 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:

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} \mu_{yt} \\ \mu_{zt} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{z,t} \end{pmatrix}. \tag{15}$$

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

$$\left( \begin{matrix} y_t \\ z_t \end{matrix} \middle| \mathcal{I}_{t-1} \right) \sim \mathsf{NID} \left( \left( \begin{matrix} \mu_{y,t}^c \\ \mu_{z,t}^c \end{matrix} \right), \left( \begin{matrix} \Sigma_{yy} & \Sigma_{yz} \\ \Sigma_{zy} & \Sigma_{zz} \end{matrix} \right) \right), \tag{16}$$

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

Hendry and Mizon (1998) have advanced two different situations in which common deterministic shifts could play an essential role in modelling. They refer to these situations as the *contemporaneous correlation* case and the *behavioral relation* case. The contemporaneous and behavioral cases can be identified with the *level A* and *level B* of the Lucas critique presented above. In the *contemporaneous correlation case* $y_t$ can be seen as a policy variable whereas $z_t$ is an instrument that policymakers 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 ($y_t$).[3] In both cases, common deterministic shifts are introduced to justify invariance of the conditional model due to changes in the marginal model. The existence of a specific linear relationship relating breaks ($\Sigma_{yz}\Sigma_{zz}^{-1}$) under the presence of weak exogeneity (see Engle *et al.*, 1983) define necessary conditions for a valid policy analysis.

In order to illustrate the empirical importance of the Lucas critique and the empirical applicability of the super-exogeneity test that we will propose later, we will briefly discuss an economic example of the *level A* of the critique. We mentioned above that a case of *level A* of the critique is the Lucas supply function. In the model presented in Lucas (1975) and discussed in Lucas (1976), the economy is considered as formed by suppliers of goods distributed over $N$ distinct markets $i = 1, \ldots, N$. The quantity supplied in each market is composed of two components:

$$y_{it} = y_{it}^p + y_{it}^c,$$

where $y_{it}^p$ and $y_{it}^c$ are respectively the permanent and transitory components of output supplied by individual $i$ to the market. Transitory supply varies with the relative prices of good in market $i$:

$$y_{it}^c = \beta(p_{it} - p_{it}^e), \tag{17}$$

where $p_{it}$ and $p_{it}^e$ are the local and general level of prices in the economy. The local level of prices in market $i$ $(p_{it})$ consist of two components:

$$p_{it} = p_t + z_{it}.$$

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

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

or in short form: $\mu_{y,t} = \mu_y^* + \Psi \mu_{z,t}$, where $\mu_{y,t}$ are the agent's plan about variables they control and $\mu_{z,t}$ are the policymakers plans about which the agents hold rational expectations. Changes in the marginal process will induce changes in the conditional model but it can be shown that the presence of common deterministic shifts could lead to a conditional model with invariant parameters.

Suppliers do not observe these components separately and just observe $p_{it}$. Conditioned on all information prior to time $t$ ($\mathcal{I}_{t-1}$), $p_t$ is assumed to be normally distributed with mean $\overline{p}_t$ and variance $\sigma^2$:

$$p_t \sim \mathsf{NID}(\overline{p}_t, \sigma^2).$$

The component $z_{it}$ is independent across time and across markets and is distributed normally with mean 0 and variance $\tau^2$:

$$z_{it} \sim \mathsf{NID}(0, \tau^2).$$

The general price is an average over all markets and considering a law of large number for $z_{it}$, the general price level is given by $p_t$. The estimate of the conditional mean of $p_t$ is $p_{it}^e$ and can be obtained as:

$$p_{it}^e = \mathsf{E}[p_t|\mathcal{I}_{t-1}, p_{it}] = (1-\theta)p_{it} + \theta\overline{p}_t \text{ with } \theta = \frac{\tau^2}{\tau^2+\sigma^2}. \tag{18}$$

Substituting expression for $p_{it}^e$ in (17), averaging over all markets and adding up the permanent component we have the Lucas supply function:

$$y_t = \theta\beta(p_t - \overline{p}_t) + y_{pt}.$$

In the Lucas supply function, output is viewed as made up of a permanent component and a transitory component that depends on deviations of the current level of nominal prices from the expected level given $\mathcal{I}_{t-1}$. The expected price level conditional on past information $\overline{p}_t$ will vary with the average inflation rate.

Lucas (1976) emphasized that though the econometrician might infer a stable trade-off between (transitory) output and the level of inflation, whenever this trade-off was exploited, the relationship broke down. Assume the price level follows a random walk with inflation a stationary process:

$$p_t = p_{t-1} + \epsilon_t,$$

with $\epsilon_t \sim \mathsf{NID}(\pi, \sigma^2)$. Conditional on all information at time $t-1$, $\overline{p}_t = p_{t-1} + \pi$, and the relation between output and inflation is given by:

$$y_t = \theta\beta(p_t - p_{t-1}) - \theta\beta\pi + y_{pt}. \tag{19}$$

If policymakers want to exploit an empirical relationship such as that implied by (19) and alter the rate of inflation $\pi$, the relationship that initially seemed characterized by stable parameters will be subject to shifts.

Policy-induced changes in the distribution of the marginal process lead to changes in the conditional model but the invariance of the parameters of interest in guaranteed in the presence of super-exogeneity. In an empirical model of output and inflation shifts in the marginal model for inflation induced by policymakers, will induce changes in the conditional parameters of output unless super-exogeneity holds. Shifts in the conditional and marginal process occur simultaneously and thus interventions could be modeled within the framework we will discuss in the following. A CDS test for super-exogeneity could be implemented to analyze the invariance of the conditional parameters and will be presented in the next section .

### 5.3 The conditional system

Valid inference from a conditional system requires weak exogeneity of the marginal process with respect to the parameters of interest in the conditional model. Using the normality of $\varepsilon_t$, the reduced-form model in (15) can be expressed in terms of the conditional model and the marginal model as:

$$\begin{aligned} y_t \mid z_t, \mathcal{I}_{t-1} &\sim \mathsf{NID}\left(\mu^c_{y|z,t}, \Omega\right), \quad (20) \\ z_t \mid \mathcal{I}_{t-1} &\sim \mathsf{NID}\left(\mu^c_{z,t}, \Sigma_{zz}\right), \quad (21) \end{aligned}$$

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

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

Rewriting (15) in terms of the conditional and marginal model results in:

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

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

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

In general, this type of model is prone to suffer from the Lucas critique. That is, changes in the marginal model lead to non-constancy of the conditional model. Shifts of the marginal model may induce shifts in the conditional model, but a convenient linear combination can induce constancy in the conditional process, such that:

$$\begin{pmatrix} \mathbf{I} & -\Sigma_{yz}\Sigma_{zz}^{-1} \end{pmatrix} \begin{pmatrix} \mu_{yt} \\ \mu_{zt} \end{pmatrix} = \text{constant}. \quad (23)$$

If we model $\left(\mu'_{yt} : \mu'_{zt}\right)$ with the corresponding intervention variables, Equation (15) can be rewritten as:

$$x_t = A_1 x_{t-1} + \mathcal{M} D_t + \varepsilon_t, \quad (24)$$

Note that the condition in (23) requires a linear relationship of breaks across equations. Hence, in the presence of super exogeneity, we would have a reduced rank condition on the coefficients of the intervention variables used to model the unrestricted system ($\mathcal{M}$), such that we can rewrite the previous model as

$$x_t = A_1 x_{t-1} + \eta\xi' D_t + \varepsilon_t. \quad (25)$$

Furthermore, in order for super-exogeneity to hold, the conditional model should be invariant to the set of interventions in the marginal process, which would require

$$\left(\mathbf{I} : -\Sigma_{yz}\Sigma_{zz}^{-1}\right)\eta = 0. \quad (26)$$

This implies that $\eta_\perp = (\mathbf{I} : -\Sigma_{yz}\Sigma_{zz}^{-1})$, such that $\eta_\perp \supset (\mathbf{I} - \Sigma_{yz}\Sigma_{zz}^{-1})$. So the reduced rank of the coefficient of the intervention dummies with specific restrictions on the null space of $\eta$ and weak exogeneity imply super-exogeneity. This postulates the invariance of the conditional model under a set of interventions in the marginal model.

# 6 Testing 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 this class of interventions, we firstly need to estimate the model under the null characterized by a reduced rank of $\mathcal{M} = \eta\xi'$ and specific restrictions on $\eta$, $\eta' = \left(\Sigma_{zz}^{-1}\Sigma_{zy} : \mathbf{I}\right)$; secondly under the alternative, i.e. the unrestricted model with the reduced rank of $\mathcal{M} = \eta\xi'$ imposed. In what follows, we present the original procedure to test for super-exogeneity proposed by Engle and Hendry (1993) and two alternative new procedures. These three procedures differ in the way in which the model is estimated under the null and how the test is constructed. The first procedure reparameterizes the original approach of Engle and Hendry (1993) as a reduced rank restriction where the null corresponds to Engle and Hendry (1993), but the alternative is defined differently. The second method involves reduced-rank-regression estimations of the reduced-form model under the conditions of super-exogeneity. The third method introduces a very general approach to the restricted estimation of $\eta$ and $\xi$, where additional restrictions can be imposed.

### 6.1 The Engle and Hendry procedure

A simple testing procedure can be implemented with just linear regressions. In order to show this procedure let us start from the model in (15)

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \mathcal{M} \begin{pmatrix} d_{t1} \\ d_{t2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{z,t} \end{pmatrix}. \tag{27}$$

where $\mathcal{M}$ is a matrix of coefficients of the intervention variables,

$$\mathcal{M} = \begin{pmatrix} \mu_1^y & \mu_2^y \\ \mu_1^z & \mu_2^z \end{pmatrix}.$$

The conditional model is given by:

$$\begin{aligned} y_t &= A_{11}y_{t-1} + A_{12}z_{t-1} + \mu_1^y d_{t1} + \mu_2^y d_{t2} + E(\varepsilon_{y,t}|\varepsilon_{z,t}), \\ &= A_{11}y_{t-1} + A_{12}z_{t-1} + \mu_1^y d_{t1} + \mu_2^y d_{t2} \\ &\quad + \omega(z_t - A_{21}y_{t-1} - A_{22}z_{t-1} - \mu_1^z d_{t1} - \mu_2^z d_{t2}) + \widetilde{\varepsilon}_{y,t}, \end{aligned}$$

with $\omega = \Sigma_{yz}\Sigma_{zz}^{-1}$ and $\widetilde{\varepsilon}_{y,t} = \varepsilon_{y,t} - \omega\varepsilon_{z,t}$, which can be rewritten as:

$$\begin{aligned} y_t &= \omega z_t + (A_{11} - \omega A_{21})y_{t-1} + (A_{12} - \omega A_{22})z_{t-1} + (\mu_1^y - \omega\mu_1^z)d_{t1} \\ &\quad + (\mu_2^y - \omega\mu_2^z)d_{t2} + \widetilde{\varepsilon}_{y,t}, \end{aligned}$$

and the marginal model is given by:

$$z_t = A_{21}y_{t-1} + A_{22}z_{t-1} + \mu_1^z d_{t1} + \mu_2^z d_{t2} + \varepsilon_{z,t}. \tag{28}$$

Under the super-exogeneity condition, we have the restrictions $(\mu_1^y - \omega\mu_1^z) = \mathbf{0}$ and $(\mu_2^y - \omega\mu_2^z) = \mathbf{0}$, which are the reduced rank conditions of Equation (23). For bivariate models, these restrictions can also be written as $\mu_1^y/\mu_1^z = \mu_2^y/\mu_2^z = \omega$, where the specific restrictions on the null space of $\eta$ become clearer. Under the null of super- exogeneity, we have that the conditional process reduces to

$$y_t = \omega z_t + (A_{11} - \omega A_{21})y_{t-1} + (A_{12} - \omega A_{22})z_{t-1} + \widetilde{\varepsilon}_{y,t}. \tag{29}$$

The parameters in the conditional model are $\theta_c = \{\omega, A_{11} - \omega A_{21}, A_{12} - \omega A_{22}, \Omega\}$ and the parameters in the marginal model are $\theta_m = \{A_{21}, A_{22}, \mu_t^z, \Sigma_{zz}\}$. The Gaussianity of the errors implies that the parameters in the marginal process are variation free of the parameters in the conditional process. Equations (29) and (28) can be estimated separately and the full maximum likelihood estimate is made up of two factors corresponding to the marginal and conditional density. The maximum likelihood estimation under the alternative just requires the estimation of the model under the reduced rank restriction along the lines in §3. The construction of a likelihood ratio test is thus straightforward.

### 6.2 The CDS testing procedure

In order to show the estimation procedure of the model under the super-exogeneity restriction $\eta' = \left(\Sigma_{zz}^{-1}\Sigma_{zy} : \mathbf{I}\right)$, let us start from the model equation:

$$R_X = \mathcal{M} R_D + E,$$

where $R_X$ and $R_D$ are the corrected residuals from Equation (4). Imposing the reduced rank restriction $\mathcal{M} = \eta\xi'$ we get,

$$R_X = \eta\xi' R_D + E. \tag{30}$$

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

(1) Start with initial estimates $\{\eta^{(0)}, \Sigma^{(0)}\}$ of $\eta$ and $\Sigma$ satisfying the restriction:

$$\eta^{(0)\prime} = \left(\Sigma_{zz}^{(0)^{-1}}\Sigma_{zy}^{(0)} : \mathbf{I}\right).$$

(2) Multiplying (30) through by $\overline{\eta}^{(0)\prime} = \left(\eta^{(0)\prime}\eta^{(0)}\right)^{-1}\eta^{(0)\prime}$ results in an ordinary linear regression problem in $\xi$:

$$\overline{\eta}^{(0)\prime} R_X = \xi' R_D + \overline{\eta}_{(0)}{}' E.$$

Applying OLS to the previous equation delivers new estimates of $\xi$, denoted $\xi^{(1)}$.

(3) Given $\{\xi^{(1)}, \Sigma^{(0)}\}$, new estimates of $\eta$ and $\Sigma$ can be obtained from (30):

$$R_X = \eta\left(\xi^{(1)\prime} R_D\right) + E.$$

We can loop in this algorithm till we obtain convergence, where in each iteration the restriction $\eta' = \left(\Sigma_{zz}^{-1}\Sigma_{zy} : \mathbf{I}\right)$ is always updated.

### 6.3 A generalized CDS testing procedure

An alternative estimation procedure of the model under the super-exogeneity restriction can be based on the first-order conditions of the maximum likelihood problem. Let us start from the concentrated likelihood function,

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

Introducing a very general formulation, we can write the linear restriction for $\xi$ and $\eta$ as:

$$\begin{aligned} \mathrm{vec}(\xi) &= H\varphi + h, \\ \mathrm{vec}(\eta') &= G\chi. \end{aligned}$$

Where $\xi$ denotes the matrix of the linear combinations of the intervention variables with identifying restrictions imposed. $H$ is a known $nr \times k$ matrix and $\varphi \in R^k$ is an unrestricted parameter vector. The vector $h$ corresponds to normalization restrictions on the vector of $\xi$. $G$ is a known $nr \times l$ matrix of full rank and $\chi \in R^l$ is an unrestricted parameter vector. The purpose of this formulation is that we can impose jointly further restrictions to those implied by the super-exogeneity restrictions. The two previous procedures dealt with the bivariate case. Now, we exemplify this procedure using a trivariate system, where $x_t' = (y_t' : z_t' : w_t')$ . One might want to investigate the super-exogeneity of $w_t$ with respect to a class of interventions $\Lambda$ in the conditional model $y_t|w_t$. We have seen in the previous section that this implies a set of restrictions on $\eta$. With this procedure one could test the super-exogeneity restrictions jointly with further restrictions on the coefficients of the loading $(\eta)$ on $z_t$ or restrictions on $\xi$.

The derivatives of the likelihood function with respect to $\varphi$ and $\chi$ are given by:

$$\frac{\partial lnL}{\partial \varphi} = H'\text{vec}(S_{DX}\Sigma^{-1}\alpha) - H'\text{vec}(S_{DD}\xi\eta'\Sigma^{-1}\eta). \tag{31}$$

$$\frac{\partial lnL}{\partial \chi} = G'\text{vec}(\xi' S_{DX}\Sigma^{-1}) - G'\text{vec}(\xi' S_{DD}\xi\eta'\Sigma^{-1}). \tag{32}$$

By substituting the restrictions of $\xi$ in (31) we get:

$$\varphi(\chi, \Sigma) = \left[H'(\eta'\Sigma^{-1}\eta \otimes S_{DD})H\right]^{-1} \left[H'(\eta'\Sigma^{-1} \otimes \mathbf{I})\text{vec}(S_{DX}) - H'\left(\eta\prime\Sigma^{-1}\eta \otimes S_{DD}\right)h\right]. \tag{33}$$

Similarly for $\chi$, we can substitute the restrictions for $\eta$ in (32) and we get:

$$\chi(\varphi, \Sigma) = \left[G'(\Sigma^{-1} \otimes \xi' S_{DD}\xi)G\right]^{-1} \left[G'(\Sigma^{-1} \otimes \xi')\text{vec}(S_{DX})\right]. \tag{34}$$

It can easily be seen that the first-order condition for $\Sigma$, for given $\chi$ and $\varphi$, is given by:

$$\Sigma(\varphi, \chi) = S_{XX} - S_{XD}\xi\eta' - \eta\xi' S_{DX} + \eta\xi' S_{DD}\xi\eta'. \tag{35}$$

Hence for initial $\chi$ and $\Sigma$ we can impose the restrictions on $\eta$ and obtain estimates of $\xi$ from (33). For given $\xi$ and $\Sigma$ new estimates of $\chi$ can be obtained from (34). For given $\xi$ and $\eta$, equation (35) delivers new estimates of $\Sigma$. We can then iterate in this algorithm with the restriction $\eta' = \left(\Sigma_{zz}^{-1}\Sigma_{zy} : \mathbf{I}\right)$ always updated.

The likelihood ratio test can be shown to have a $\chi^2$ distribution. The degrees of freedom result from comparing the tangent space of $\eta\xi'$ with and without restrictions (see Johansen, 1995).

# 7 A small Monte Carlo experiment

The power of different tests for super-exogeneity has been analyzed by Psaradakis and Sola (1996) and Favero and Hendry (1992) for *level B* of the Lucas critique: while agents are forward-looking, the econometric model is based on backward-looking expectations, and changes in the economic environment alter expectations inducing changes in the conditional model. The power results obtained in their simulation experiment lead Psaradakis and Sola (1996) to question the usefulness of super-exogeneity and invariance tests for the assessment of the empirical relevance of forward-looking models. While Favero and Hendry (1992) also question the usefulness of the tests for the assessment of forward-looking models, they interpret the Monte Carlo results as a lack of strength of the Lucas critique: *"for changes of a magnitude that are large enough to be policy relevant, conditional model approximating expectations process do not in practice experience dramatic predictive failure"* **(**p. 290**)**. In contrast to the previous literature, we focus on the power of super-exogeneity test at *level A* of the Lucas critique.[4] The following Monte Carlo study is designed to analyze the size and power of the CDS tests for super-exogeneity proposed in $\S 6.2$. We will compare the results with the test introduced by Engle and Hendry (1993).

In the Monte Carlo simulations, we use the benchmark $\mathsf{DGP}_2$ of $\S 4$ with a reduced-rank matrix of intervention coefficients:

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0.75 & 0.5 \\ 0 & 0.8 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} \mu_1^y & \mu_2^y \\ \mu_1^z & \mu_2^z \end{pmatrix} \begin{pmatrix} d_{t_1} \\ d_{t_2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{36}$$

$$\text{where } \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim \mathsf{NID}\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{yy} & \sigma_{yz} \\ \sigma_{zy} & \sigma_{zz} \end{pmatrix} \right).$$

The break points are again at point $\tau_1 = 0.3$ and $\tau_1 = 0.7$ of the sample. Under super-exogeneity the restrictions that we will impose in the matrix $\mathcal{M}$ will imply restrictions in the matrix $\Omega$, and this resulting matrix will be our new $\Omega$ in the benchmark $\mathsf{DGP}_2$.

We will first consider the case in which the shifts in the marginal process $z_t$ also affect the $y_t$ process, but the parameters of the conditional process $y_t|z_t$ remain constant. The corresponding restrictions on the intervention coefficients are given by the reduced rank restriction, $\mathcal{M} = \eta\xi'$, and the super-exogeneity restriction $\eta' = \left(\Sigma_{zz}^{-1}\Sigma_{zy} : \mathbf{I}\right)$. For our bivariate case, these restrictions can be written as:

$$\frac{\mu_1^y}{\mu_1^z} = \frac{\mu_2^y}{\mu_2^z} = \frac{\sigma_{zy}}{\sigma_{zz}}. \tag{37}$$

Under condition (37), the parameters in the conditional process are invariant to the class of interventions in the marginal process. In other words, despite changes in the marginal distribution induced by the policymaker's alteration of his policy rule (here, intercept shift), the parameters in the conditional model are invariant to this intervention. The Lucas critique does not hold for this specific parameterization of the DGP. Let us call this model $\mathsf{DGP}_1^*$.

Alternatively, we consider the case where the parameters of the conditional model are not invariant to changes in the intercept of the policy rule. The intervention matrix $\mathcal{M} = \eta\xi'$ still has reduced rank, but the super-exogeneity condition does not hold:

$$\frac{\mu_1^y}{\mu_1^z} = \frac{\mu_2^y}{\mu_2^z} = \frac{\sigma_{zy}}{\sigma_{zz}} + h,$$

[4] It is worth mentioning that in most settings when expectations are modeled as forward-looking, weak exogeneity of the marginal process with respect to the parameters of interest is generally violated. In this case, testing for the constancy of the parameters of the conditional process or additional variable tests for the absence of the moments of the marginal process in the conditional models are invariance tests.

with $h \neq 0$. In the simulations we will use a value of 0.5 for $h$. This correspond to a situation in which shifts in the $y_t$ and $z_t$ process are linearly related, the matrix of the intervention coefficients is of reduced rank but super-exogeneity does not hold. So the conditional model has not constant parameters. We will refer to this case as $\mathsf{DGP}_2^*$.

In the previous section it has been shown that the test for super-exogeneity is a test of $H_0$ : $\operatorname{rank}(\mathcal{M}) = 1 \wedge \eta' = \left(\sigma_{zz}^{-1}\sigma_{zy} : 1\right)$ against $H_1$ : $\operatorname{rank}(\mathcal{M}) = 1 \wedge \eta' \neq \left(\sigma_{zz}^{-1}\sigma_{zy} : 1\right)$, which follows a $\chi^2(1)$. In order to analyze the size of the test, we will consider $\mathsf{DGP}_1^*$ for sample sizes of $T = 50$, $100$ and $150$.

Table 3 Size of the CDS test for super-exogeneity.

| T | 20% | 10% | 5% | 1% |
|---|---|---|---|---|
| 50 | 0.298 | 0.183 | 0.114 | 0.0377 |
| 100 | 0.266 | 0.153 | 0.086 | 0.0232 |
| 150 | 0.256 | 0.147 | 0.081 | 0.0217 |

*Note:* Rejection frequencies for the CDS test statistic for $\mathsf{DGP}_1^*$.

Table 4 Power of the CDS test for super-exogeneity.

| T | 20% | 10% | 5% | 1% |
|---|---|---|---|---|
| 50 | 0.971 | 0.948 | 0.917 | 0.809 |
| 100 | 0.997 | 0.994 | 0.988 | 0.958 |
| 150 | 0.999 | 0.999 | 0.997 | 0.989 |

*Note:* Rejection frequencies for the CDS test statistic for $\mathsf{DGP}_2^*$.

Size and power of the CDS test for super-exogeneity are reported in Tables 3 and 4. Table 3 presents the rejection frequencies for the CDS test for super-exogeneity of $H_0$ : $\operatorname{rank}(\mathcal{M}) = 1 \wedge \eta' = \left(\sigma_{zz}^{-1}\sigma_{zy} : 1\right)$ against $H_1$ : $\operatorname{rank}(\mathcal{M}) = 1 \wedge \eta' \neq \left(\sigma_{zz}^{-1}\sigma_{zy} : 1\right)$ under the null ($\mathsf{DGP}_1^*$). Table 4 shows the rejection frequencies of the test when the simulated model is $\mathsf{DGP}_2^*$. The actual size of the CDS test for super-exogeneity differs only slightly from nominal size and power (non size corrected) is high.

Table 5 Size of the Engle and Hendry test for super-exogeneity.

| T | 20% | 10% | 5% | 1% |
|---|---|---|---|---|
| 50 | 0.349 | 0.216 | 0.129 | 0.0353 |
| 100 | 0.331 | 0.202 | 0.121 | 0.0323 |
| 150 | 0.309 | 0.184 | 0.106 | 0.0282 |

*Note:* Rejection frequencies for the Engle and Hendry test for $\mathsf{DGP}_1^*$

For comparison purposes, we also consider the Engle and Hendry (1993) test for super-exogeneity. This is a variable addition tests constructed as a conventional $F$-test for the joint significance of the intervention variables in the conditional model: $H_0$: $\mu_{1,c}^y = \mu_{2,c}^y = 0$ against $H_1$: $\mu_{1,c}^y \neq$ and $\mu_{2,c}^y \neq 0$.[5] The results for size and power of this test are presented in Tables 5 and 6. Table 5 reports the rejection frequencies for the test on the conditional model of $H_0$: $\mu_{1,c}^y = \mu_{2,c}^y = 0$ against $H_1$: $\mu_{1,c}^y \neq$ and $\mu_{2,c}^y \neq 0$

[5]Where $\mu_{1,c}^y = (\mu_1^y - \omega\mu_1^z)$ and $\mu_{2,c}^y = (\mu_2^y - \omega\mu_2^z)$ are the coefficients of the intervention variables in the conditional model.

Table 6 Power of the Engle and Hendry test for super-exogeneity.

| T | 20% | 10% | 5% | 1% |
|---|---|---|---|---|
| 50 | 0.971 | 0.943 | 0.901 | 0.753 |
| 100 | 0.997 | 0.992 | 0.983 | 0.941 |
| 150 | 0.999 | 0.998 | 0.996 | 0.982 |

*Note:* Rejection frequencies for the Engle and Hendry test for $\mathsf{DGP}_2^*$

when the simulated model is $\mathsf{DGP}_1^*$. Table 6 presents the rejection frequencies of the test when the simulated model is $\mathsf{DGP}_2^*$. The size and (non-size corrected) power results obtained for the Engle and Hendry test for super-exogeneity are similar to the ones for the CDS test, though there is some indication that the Engle and Hendry test does not perform as good as the CDS test.

In order to be able to compare the properties of the two tests further, size-power curves are plotted in Figure 6 for both tests for sample sizes of $T = 50$ and $100$. The left hand panel of Figure 6 shows the results for a sample size of 50. The continuous line graph corresponds to the size-power curve for the CDS and the discontinuous line corresponds to the Engle and Hendry test. These graphs show that for a given size the CDS test for super-exogeneity has a higher power than the Engle and Hendry test. For a sample size of 100 (RHS of Figure 6), the behavior of the CDS test is still better, though the differences between the behavior are greatly reduced. For larger samples, results did not differ much.

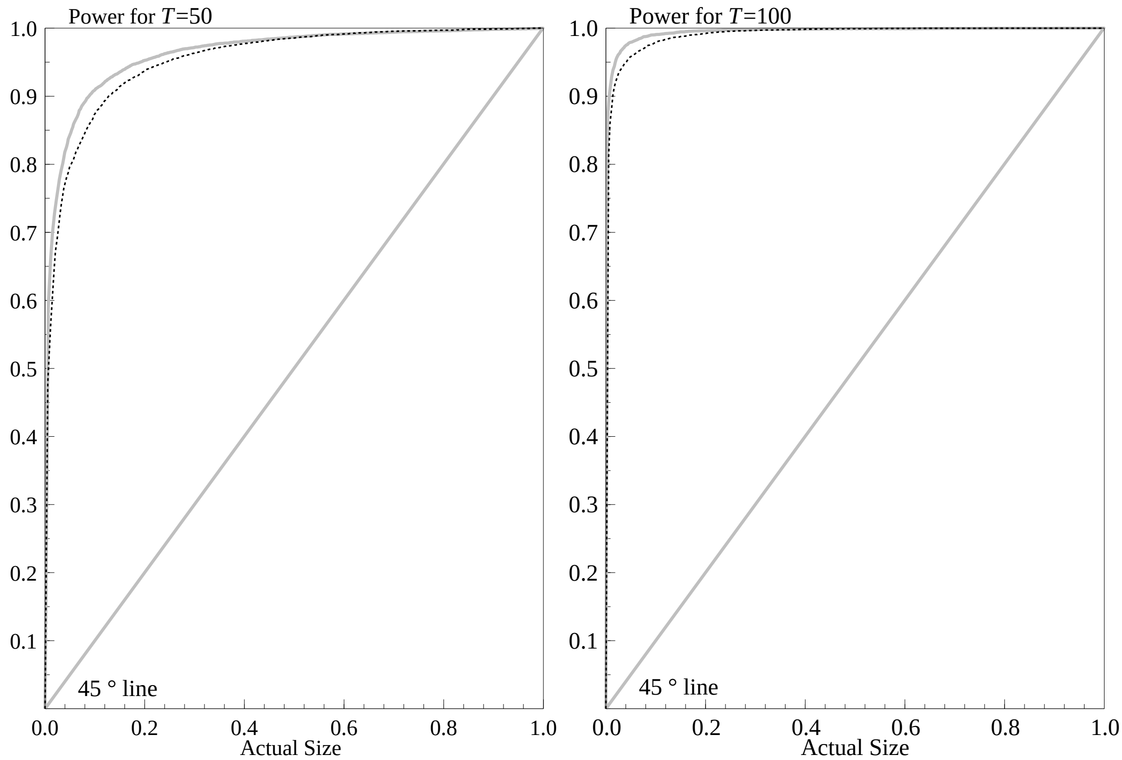


Figure 6 Size-power curve analysis for the Engle-Hendry and CDS tests for super-exogeneity.

In the light of these results, the use of the CDS tests for super-exogeneity in empirical applications can be recommended. As we have seen, the CDS test for super-exogeneity has some comparative advantages versus tests proposed in the literature when sample sizes are small, which is often the case in economics.

# 8 Some Comments

## 8.1 Empirical Illustrations

In this paper we propose a general approach to the modelling and testing of common deterministic shifts An important advantage of this approach is that it does not rely on *prior* information regarding the relationships between the shifts in the mean of the individual processes. This can be useful in cases with limited information about the classification of targets and instruments and their relationship.

Consider, for example, a monetary model where some short-term interest rates (say, the discount rate and the inter-bank rate) are included together with the target variables the policymaker intends to control. The identification of a linear relationship linking shifts in the mean of the different processes would provide valuable information about the exact relationship between instruments and targets, particularly the transmission of interventions (by the weights of the linear relationship that govern the deterministic shifts).

It is worth emphasizing that even in the absence of super-exogeneity, the presence of CDS will deliver important information about the economic system and testing for CDS could be used to ask relevant economic questions. Let us take as an example the defense of a currency by the government. Suppose that the researcher uses a vector autoregression to model the exchange rate, the yield spread of domestic government bonds with a benchmark risk-free bond, and the short-term interest rate as the policy instrument. In a period of exchange rate turbulence, he would probably find that the government has intervened repeatedly raising the interest rate in order to defend the currency. Furthermore suppose that there have been just three interventions and that these actions of the government are represented by intervention dummies. The researcher could then hypothesize that if each of the interventions was credible, *ceteris paribus*, they would have a proportional effect on the exchange rate and the yield spreads. This would imply a reduced rank of the matrix of dummy coefficients. In contrast, a full rank would indicate the non-credibility of the government interventions (the first intervention had a proportionally different effect than the second and third intervention).

Another example from industrial organization is the analysis presented in Guerrero, Peña and Poncela (1998). Guerrero *et al.* (1998) aim to measure the effects of promotional campaigns launched by a bank in order to attract deposits. The banking institution wanted to identify the possible effects on (a) new accounts, (b) stock variation, (c) cancellations, and (d) total amount. Since the increase in total amount from the previous month to the current month is given by new accounts $+$ stock variation $-$ Cancellation, a linear restriction is satisfied by the vector of time series every month and also applies to the intervention coefficients.

## 8.2 Limitations, Extensions and Generalizations

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*.

In the case of cointegration, one could be interested in the following model:

$$\Delta x_t = \Pi x_{t-1} + \mathcal{M} D_t + \varepsilon_t, \tag{38}$$

where $x_t$ is $I(1)$, $\Pi$ has reduced rank and $\mathcal{M}$ and $D_t$ had been defined in section 2. The above model equation raises two different issues.

- The first issue is the mapping of the $I(1)$ system to the $I(0)$ system.

- The second issue is the test of CDS in the vector equilibrium representation of the system. That is, when we have mapped the system from $I(1)$ to $I(0)$.

The second issue is probably the easiest of both issues and we will address it first. A general result is that when the $I(1)$ system has been mapped to $I(0)$, a test of restrictions on $\eta$ have a $\chi^2$ distribution. In most empirical applications economic priors are used to map the system to $I(0)$ (see Favero and Hendry (1992) for a situation in which an $I(1)$ system is mapped to an $I(0)$ system and then the Lucas critique is investigated).

If we map the $I(1)$ system into the vector equilibrium representation:

$$\Delta x_t = \alpha\beta' x_{t-1} + \eta\xi' D_t + \varepsilon_t, \tag{39}$$

the system can then be estimated by a switching algorithm which provides reduced-rank estimations of (i) $\alpha, \beta$ conditional on $\eta, \xi$ and (ii) $\eta, \xi$ conditional on $\alpha, \beta$. Thus one would combine the techniques developed in this paper with the well-established cointegration analysis of Johansen (1995) (see Toro, 1999). So, for known rank of $\Pi$ and $\mathcal{M}$, one could impose a CDS restriction together with restrictions in $\alpha$ or $\beta$ and these are $\chi^2$ tests.

A probably more technical and complicated issue is the mapping of the $I(1)$ to the $I(0)$ system. In order to do this we need the joint determination of the rank on $\Pi$ and $\mathcal{M}$. This issue is quite involved. We will try to summarize the main issues without introducing new notation and going into much technicality. Here one could investigate the determination of the rank of $\Pi$ in the presence of intervention variables and then investigate the rank of $\mathcal{M}$ for given $rank(\Pi) = r$. We first review the literature on the determination of the rank of $\Pi$ in the presence of intervention variables and then discuss the difficulties involved when $\mathcal{M}$ has also reduced rank suggesting potential solutions to this problem in empirical applications.

Most empirical applications that use the cointegrated VAR system with intervention variables assess the rank of $\Pi$ based on the traditional tables tabulated in Johansen (1995). However Johansen and Nielsen (1993) showed that in the presence of deterministic dummy variables in the cointegrated VAR system, different distributions apply. These are generalized Dickey-Fuller distributions. Johansen and Nielsen (1993) propose a program for the simulation of the specific distribution. Johansen, Nielsen and Mosconi (2000) analyze cointegration in Gaussian vector autoregressive model with a broken linear trend and known break points. They investigate three models which they denote by $H_l$, $H_{lc}$ and $H_c$. In model $H_l$ in each subsample the deterministic component is linear both for non-stationary and cointegrating relationships. Model $H_c$ has no linear trend but a broken constant level. In model $H_{lc}$ the non-stationary relations have a broken linear trend whereas the cointegration relations have broken constant levels. They demonstrate that in model $H_{lc}$, the third model, the asymptotic analysis is heavily burdened with nuisance parameters. The asymptotic distribution is investigated for each model and the moments of the distributions are summarized with a response surface. They suggest to approximate the critical values with a $\Gamma$-distribution using the moments obtained in a response surface analysis.

A further possible model is a cointegrated VAR model where the set of intervention coefficients could also have a reduced rank structure. In this model one would be interested in investigating the joint determination of the rank of $\Pi$ and $\mathcal{M}$. It has been shown (see Toro, 1999) that the asymptotic distribution of the test for the rank of $\Pi$ when the regressors of $\mathcal{M}$ are deterministic depends on the rank of $\mathcal{M}$ or the number of deterministic trends that load into the system and whether $\alpha_\perp\mathcal{M} = 0$ or $\alpha_\perp\mathcal{M} \neq 0$. Also tests on the rank of $\mathcal{M}$ depend on the rank of $\Pi$. This seems to be a complicated setting. However it turns out that a much simpler path exist if the interpretation of the deterministic component is not questioned with the rank but is tested in a later stage (see Nielsen and Rahbek (2000)). We could consider the situation in which the intervention variables induce a broken constant level in the cointegrating relationships but they do not induce any broken trend. This is probably the most relevant situation in macroeconomics and it would be a relevant situation in the case of a test of the Lucas critique. We would be considering interventions that alter the constant term in an equilibrium relationship but that does not induce a changing trend. This can be formulated as $\alpha_{\perp}\mathcal{M} = 0$, such that Model Equation 38 can be reparametized as:

$$\Delta Y_t = \alpha(\beta' Y_{t-1} + \mathcal{M}_r D_{t-1}) + \mathcal{M}\Delta D_t + \varepsilon_t, \tag{40}$$

with $\mathcal{M}_r = (\alpha'\alpha)^{-1}\alpha'\eta\xi'$. Under the further restriction $\alpha_{\perp}\mathcal{M} = 0$ there is no nuisance parameter in the distribution of the test for the cointegrating rank (see Johansen *et al.* (2000)). Once the rank of $\Pi$ has been investigated, one could then investigate hypotheses about the rank of $\alpha\mathcal{M}_r = \Upsilon$ or $\mathcal{M}$ in the form of CDS tests or CDS-super-exogeneity test, and they follow a $\chi^2$ distribution.

A further 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.

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

where $D_t$ is now a Markov chain. Krolzig (1997) considers the statistical analysis of this system when the (potentially reduced) rank of $\mathcal{M} = \eta\xi'$ 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.

# 9 Conclusions

This paper brings together two topics of econometric research: common deterministic shifts and super-exogeneity. We have shown how common deterministic shifts can be analyzed with simple and widely-known technique: reduced rank regressions. The proposed CDS tests were found of being only slightly oversized and having excellent power properties in small samples.

Deterministic shifts in the conditional mean of economic variables are a recurrent feature in empirical economics. These shifts happen to affect not just one single economic variable but also contemporaneously, other related variables. Furthermore, these shifts might be related linearly and this linear relationship might prevail throughout time. We have proposed a technique that can be used to analyze such phenomena, and can help to gather important information about how breaks are related through economic variables and across time. Frequently, deterministic shifts are induced by policy changes. Policymakers move the level of some variables in order to affect some target variables and reach specific goals. When deterministic shifts are induced by policymakers, the relationship between common deterministic shifts and super-exogeneity becomes apparent.

## References


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

Davidson, R., and MacKinnon, J. (1997). Graphical methods for investigating the size and power of hypothesis tests. Discussion paper.

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 Hendry, D. (1993). Testing super exogeneity and invariance in regression models. *Journal of Econometrics*, **56**, 119–139.

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

Favero, C., and Hendry, D. (1992). Testing the Lucas critique: A review. *Econometric Reviews*, **11**, 265–306.

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

Guerrero, V., Peña, D., and Poncela, P. (1998). Measuring intervention effects on multiple time series subjected to linear restrictions:a banking example. *Journal of Business Economics and Statistics*, **16**, 489–497.

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: OUP.

Johansen, S., and Nielsen, B. (1993). Asympotics for cointegration rank tests in the presence of intervention dummies, manual for the simulation program disco. Preprint, University of Copenhagen, http:www.math.ku.dk.

Johansen, S., Nielsen, B., and Mosconi, R. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. Manuscript, European University Institute.

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

Lucas, R. (1975). Some international evidence on output-inflation tradeoffs. *American Economic Review*, **63**, 326–334.

Lucas, R. (1976). Econometric policy evaluation: a critique. *Carnegie-Rochester conference series on public policy*, **1**, 19–46.

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

Nielsen, B., and Rahbek, A. (2000). Similiraity issues in cointegration models. *Oxford Bulletin of Economics and Statistics*, **6**, 5–22.

Psaradakis, Z., and Sola, M. (1996). On the power of tests for superexogeneity and structural invariance. *Journal of Econometrics*, **72**, 151–175.

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.