Cointegration between trends and their estimators in state space models and CVAR models

Søren Johansen and Morten Nyboe Tabor

CREATES Research Paper 2017-11
Cointegration between trends and their estimators in state space models and CVAR models

Søren Johansen† Morten Nyboe Tabor‡
University of Copenhagen University of Copenhagen
and CREATESand CREATEs
February 28, 2017

Abstract
In a linear state space model, $y_{t+1} = BT_t + \varepsilon_{t+1}$, we investigate if the unobserved trend, $T_t$, cointegrates with the extracted trend $E_tT_t$, and with the estimated trend $\hat{E}_tT_t$, in the sense that the spreads $T_t - E_tT_t$ and $E_tT_t - \hat{E}_tT_t$ are stationary. We find that this result holds for $BT_t - BE_tT_t$ and $BE_tT_t - \hat{B}\hat{E}_tT_t$. For the trends $T_t$ and $\hat{E}_tT_t$, however, this type cointegration depends on the identification of $B$ and $T_t$. The same results are found, if the observations, $y_t$, from the state space model are analysed using a cointegrated vector autoregressive model, where the trend is defined as the common trend. Finally we investigate cointegration between trends and their estimators based on the two models, and find the same results. We illustrate with two examples and confirm the results by a small simulation study.

Keywords: Cointegration of trends, State space models, CVAR models

JEL Classification: C32.

1 Introduction and Summary
In connection with a project on long-run causal order for nonstationary processes, see Hoover, Johansen, Juselius and Tabor (2014), we used the state space model (SSM)

$$T_{t+1} = T_t + \eta_{t+1},$$
$$y_{t+1} = BT_t + \varepsilon_{t+1},$$

t = 1, \ldots, n, to estimate the effect, $B$, of unobserved independent random walks, $T_t$, on the observed series, $y_t$. When analysing some simulation studies, we noticed that the estimator

†The first author is grateful to CREATEs - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. The second author is grateful to the Carlsberg Foundation (grant reference 2013_01_0972). We have benefitted from discussions with Siem Jan Koopman and Eric Hillebrand on state space models.
‡Address for correspondence: Department of Economics University of Copenhagen, Øster Farimagsgade 5 Building 26, 1353 Copenhagen K, Denmark. E-mail: soren.johansen@econ.ku.dk.
§Department of Economics, University of Copenhagen. E-mail: morten.nyboe.tabor@econ.ku.dk
of the extracted trend, \( \hat{E}_t T_t \), sometimes did not cointegrate with the simulated trend, \( T_t \), in the sense that \( T_t - \hat{E}_t T_t \) was not stationary. In order to understand this phenomenon we have analysed the state space model and found the solution, namely that we always have cointegration \( (1, -1) \) between \( B T_t \) and its estimator \( B \hat{E}_t T_t \), but not in general between \( T_t \) and \( \hat{E}_t T_t \). The latter result depends on how the parameter \( B \) and the trend \( T_t \) are identified.

We then investigated the same questions using a cointegrated vector autoregressive model (CVAR) to analyse \( y_{1,t}, \ldots, y_{n,t} \), and found the same results. Finally we compared the estimated trends extracted from the SSM and the CVAR models and found the same results.

The results are given in Theorems 1, 2 and 3. For the trends from the CVAR, the trends from the SSM, and finally for a comparison for the estimated trends from the two models.

## 2 The DGP and the statistical models

The data generating process (DGP) is formulated as

\[
\begin{align*}
T_{t+1}^0 &= T_t^0 + \eta_{t+1}, \\
y_{t+1} &= B^0 T_t^0 + \varepsilon_{t+1},
\end{align*}
\]  

(2)

where \( T_t^0 \in \mathbb{R}^m \) and \( y_t \in \mathbb{R}^p \), \( t = 1, \ldots, n \). We assume \( m < p \), and that \( B^0 \) has full column rank. Moreover \( \eta_t \) and \( \varepsilon_t \) are mutually independent and i.i.d. Gaussian with mean zero and variances \( \Omega^0_\eta = \text{diag}(\sigma^{0}_{11}, \ldots, \sigma^{0}_{pp}) > 0 \) and \( \Omega^0_\varepsilon > 0 \) respectively. Note that the trend enters with a lag in the observation equation. This is of course just a question of notation, as we can define the trend \( S_{t+1}^0 = T_t^0 \), and the observation equation becomes \( y_{t+1} = B^0 S_{t+1}^0 + \varepsilon_{t+1} \).

### 2.1 The statistical models

We analyse the observations \( y_1, \ldots, y_n \) from the DGP given in (2) using two different statistical models each containing the DGP.

#### 2.1.1 The first statistical model

We consider the state space model (SSM) defined by (1) for the observed process, \( y_t \), and unobserved process, \( T_t \), with \( p \times m \) parameter \( B \) of full rank, and where \( \Omega_\eta > 0 \) and \( \Omega_\varepsilon > 0 \) are freely varying positive covariance matrices of dimensions \( m \times m \) and \( p \times p \) respectively. This statistical model can be analysed using the Kalman filter to calculate the likelihood function, and an optimizing algorithm can be used to find the maximum likelihood estimator.

#### 2.1.2 The second model

Next we reformulate the DGP in error correction form for the observation \((y_t, T^0_t)\)

\[
\begin{pmatrix}
\Delta y_{t+1} \\
\Delta T^0_{t+1}
\end{pmatrix} = \begin{pmatrix}
-I_p & B^0 \\
0 & 0
\end{pmatrix} \begin{pmatrix}
y_t \\
T^0_t
\end{pmatrix} + \begin{pmatrix}
\varepsilon_{t+1} \\
\eta_{t+1}
\end{pmatrix}.
\]

It is seen that \( y_1, \ldots, y_n \) are partial observations from a CVAR for \((y'_t, T^0'_t)'\), and we use the infinite order CVAR as statistical model

\[
\Delta^2 y_t = \alpha \beta' y_{t-1} - \Gamma \Delta y_{t-1} + \sum_{i=0}^{\infty} \Gamma_i \Delta^2 y_{t-i} + \delta_t.
\]

(3)

Using results of Saikkonen (1992) and Saikkonen and Lütkepohl (1996), parameters \( \alpha, \beta, \Gamma \), and residuals \( \delta_t \) can be estimated consistently by choosing a finite lag length \( k_n \rightarrow \infty \), such that \( k_n^2/n \rightarrow 0 \).
Note that the parameters of the SSM model consist of $B, \Omega_\eta, \Omega_\varepsilon$, whereas the parameters of the CVAR are $\alpha, \beta, \Gamma, \Gamma_\varepsilon, i = 0, \ldots, \infty$, and $\Omega_\varepsilon$. Let $B_\perp^0$ be $p \times (p - m)$ of full rank satisfying $B_\perp^0B^0 = 0$. We see from (2), that $B_\perp^0y_t = B_\perp^0\varepsilon_t$ is stationary, such that the cointegrating vectors in the DGP are $\beta_\perp^0 = B_\perp^0$. Hence it is natural to assume that parameters in the two models are related by $\beta_\perp = B_\perp$.

Moreover, the long-run variance in the SSM, $\lim_{n \to \infty} n^{-1}Var(y_n) = B\Omega_\eta B'$, is the same as the long-run variance in the CVAR, $\lim_{n \to \infty} n^{-1}Var(y_n) = C\Omega C'$, where $C = B(\alpha_\perp^0 \Gamma B)^{-1} \alpha_\perp^0$, which gives another relation between the parameters, see Johansen and Juselius (2014).

### 2.2 The trends

Given the two models, there are two ways of defining trends. In the SSM generated by (2), the trend $T_t$ is part of the model formulation, and in the CVAR (3) we define the trend as the common trend

$$T_t^* = (\alpha_\perp^0 \Gamma B)^{-1} \alpha_\perp^0 \sum_{s=1}^{t} \delta_s,$$

where $\delta_t$ is the prediction error for $y_t$ given the infinite past $\{y_s, s < t\}$, see (3).

Thus we have two representations of $y_t$,

$$SSM : y_t = BT_t + \varepsilon_t - B\eta_t, \quad (4)$$

$$CVAR : y_t = C \sum_{s=1}^{t} \delta_s + u_t = B(\alpha_\perp^0 \Gamma B)^{-1} \alpha_\perp^0 \sum_{s=1}^{t} \delta_s + u_t = BT_t^* + u_t, \quad (5)$$

using the Granger Representation Theorem, see Johansen (1996, Theorem 4.2). Finally $u_t$ is an asymptotically stationary process.

### 2.3 The Kalman filter

Conditional on $y_1, \ldots, y_t$, $T_t$ is Gaussian with mean $E_tT_t = E(T_t|y_1, \ldots, y_t)$ and variance $V_t = Var(T_t|y_1, \ldots, y_t)$. It is well known, see Durbin and Koopman (2012) or Harvey (1989), that $E_tT_t$ and $V_t$ can be calculated recursively by the Kalman filter starting with $E_1T_1 = 0$ and $V_1 = \Omega_\eta$, using the equations for $t = 1, \ldots, T - 1$,

$$E_{t+1}T_{t+1} = E_tT_t + K_t(y_{t+1} - E_t y_{t+1}) \quad (6)$$

$$V_{t+1} = \Omega_\eta + V_t - V_tB'(BV_tB' + \Omega_\varepsilon)^{-1}BV_t,$$

where $K_t = V_tB'(BV_tB' + \Omega_\varepsilon)^{-1}$ is the Kalman gain. Let $\bar{B} = B(B'B)^{-1}$ and define $\Omega_B = \Omega_B = \Omega_B = Var(\bar{B}\varepsilon_t|B_\perp^0\varepsilon_t)$.

**Lemma 1** The recursion for $V_t$ can be expressed as

$$V_{t+1} = \Omega_\eta + V_t - V_t(\Omega_B + \Omega_\varepsilon)^{-1}V_t. \quad (7)$$

Solving the eigenvalue problem

$$|\lambda\Omega_B - \Omega_\eta| = 0,$$
we find eigenvectors $W$ and eigenvalues $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$, such that $W'\Omega B W = I_m$ and $W'\Omega_n W = \Lambda$. Then, if $V_1 = \Omega_n$, we find $W'V_1 W \rightarrow W'VW = \Lambda_\infty$, where

$$
\lambda_{i,\infty} = \frac{1}{2}\{\lambda_i + (\lambda_i^2 + 4\lambda_i)^{1/2}\}.
$$

It follows that

$$
K_t B = V_t B'(BV_t B' + \Omega_\varepsilon)^{-1} B \rightarrow KB = V(V' + \Omega_B)^{-1},
$$

such that $KB$ has positive eigenvalues less than one.

### 2.4 Estimation

With the above definitions of trends, there is an identification issue between $B$ and $T_t$ (or $B$ and $\Omega_n$) and between $B$ and $T_t^*$, because for any $m \times m$ matrix $M$ of full rank, we can use $BM^{-1}$ as parameter and $MT_t$ as trend and $M\Omega_T M'$ as variance, and similarly for $T_t^*$. In order to estimate $B$, $T$, and $\Omega_n$, we need to impose restrictions, and we shall give two examples of identification.

#### 2.4.1 Identification 1.

Because $B$ has rank $m$, we can permute the rows and assume that $B_0 = (B_0^1, B_0^2)$, where $B_1$ is $m \times m$ and has full rank. Then we redefine the parameters and trend as

$$
\hat{B} = \begin{pmatrix}
I_m \\
B_2 B_1^{-1}
\end{pmatrix}, \hat{\Omega}_n = B_1\Omega_n B_1', \hat{T}_t = B_1 T_t.
$$

This parametrization is useful because it separates parameters that are $n$-consistently estimated, $\gamma$, from those that are $n^{1/2}$-consistently estimated, $\Omega_n, \Omega_\varepsilon$, see Lemma 2. Note that we simply identify the (correlated) trends by defining $T_{1t}$ as the trend in $y_{1t}$, $T_{2t}$ as the trend in $y_{2t}$, and so on.

#### 2.4.2 Identification 2.

We also want to consider the normalization where $\hat{\Omega}_n$ is diagonal, and define a Cholesky decomposition $\hat{\Omega}_n = C_n diag(\sigma_{11}, \ldots, \sigma_{mm}) C_n'$, and the new parameters and the trend

$$
\tilde{B} = \begin{pmatrix}
C_n \\
\gamma' C_n
\end{pmatrix}, \tilde{\Omega}_n = diag(\sigma_{11}, \ldots, \sigma_{mm}), \tilde{T}_t = C_n^{-1}\hat{T}_t,
$$

such that $\tilde{B}\tilde{T}_t = \tilde{B}\hat{T}_t = BT_t$ and $\tilde{B}\tilde{\Omega}_n \tilde{B}' = \tilde{B}\hat{\Omega}_n \hat{B}' = B\Omega_B B'$.

This parametrization is also useful because it defines independent trends and how they load into the observations. An example is to define $T_{1t}$ as the trend in $y_{1t}$, and $T_{2t}$ as the trend in $y_{2t}$, but orthogonalized on $T_{1t}$, such that the trend in $y_{2t}$ is a combination of $T_{1t}$ and $T_{2t}$, etc.

#### 2.4.3 A simple estimator

We can estimate the state space model using the Kalman filter to calculate the likelihood function for given parameters and then maximize the likelihood function using a general optimizing algorithm. But we can also find very simple (but not efficient) estimators, which are easier to analyse.
Irrespective of the identification, we find the relations

\[ \text{Var}(\Delta y_t) = B \Omega_y B' + 2 \Omega_\varepsilon, \]  
\[ \text{Cov}(\Delta y_t, \Delta y_{t+1}) = -\Omega_\varepsilon. \]  
(12)  
(13)

In the identified parametrization (10), where \( B = (I_m, \gamma)' \), we take \( B_\perp = (\gamma', -I_{p-m})' \), define \( z_{1t} = (y_{1t}, \ldots, y_{mt})' \) and \( z_{2t} = (y_{m+1,t}, \ldots, y_{pt}) \) with a similar decomposition of \( \varepsilon_t = (\varepsilon'_{1t}, \varepsilon'_{2t})' \). Then the equation for \( y_{t+1} = (z_{1,t+1}, z_{2,t+1})' \) becomes

\[
\begin{align*}
    z_{1,t+1} &= T_t + \xi_{1,t+1}, \\
    z_{2,t+1} &= \gamma' T_t + \xi_{2,t+1} = \gamma' z_{1,t+1} - \gamma' \xi_{1,t+1} + \xi_{2,t+1} = \gamma' z_{1,t+1} - \varepsilon_t B_\perp
\end{align*}
\]  
(14)  
(15)

**Lemma 2** Assume \( B \) is identified as in (10), that is \( B' = (I_m, \gamma), B_\perp = (\gamma, -I_{p-m}) \), and \( \Omega_\eta \) is adjusted accordingly. Using (12) and (13), we can find \( n^{1/2} \)-consistent estimators, which are asymptotically Gaussian for the variance matrices \( \Omega_\eta \) and \( \Omega_\varepsilon \).

From equation (15), we find by regressing \( z_{2,t+1} \) on \( z_{1,t+1} \), that the estimator \( \hat{\gamma}_{\text{reg}} \) is \( n \)-consistent with asymptotic Mixed Gaussian distribution

\[
    n(\hat{\gamma}_{\text{reg}} - \gamma) = nB'(\hat{B}_\perp - B_\perp) \overset{D}{\to} \left( \int_0^1 W_\eta W_\eta'(du) \right)^{-1} \int_0^1 W_\eta (dW_\varepsilon)' B_\perp.
\]  
(16)

If \( B \) is identified as in (11), that is \( B = (C_\eta', C_\eta \gamma)' \), and \( \Omega_\eta = \text{diag}(\sigma_{11}, \ldots, \sigma_{pp}) \), then \( \hat{B} - B = O_P(n^{-1/2}) \) but (16) still holds for \( nB'(\hat{B}_\perp - B_\perp^0) \).

Using the Kalman filter, we can calculate the extracted trend \( \hat{E}_i(T_t) \) based on observations and known parameters, and the estimator of the extracted trend \( \hat{T}_t = \hat{E}_t T_t \), based on observations and estimated parameters.

Using the CVAR we find maximum likelihood estimators of the common trend \( \hat{T}_t^* = (\hat{\gamma}'_\perp \hat{\Gamma} \hat{B})^{-1} \hat{\gamma}'_\perp \sum_{s=1}^t \hat{\delta}_s \). Note that we use \( \hat{B} \) to indicate an estimator in SSM, and \( \hat{T} \) to indicate an estimator in CVAR.

Based on these trends, we can ask whether the trends cointegrate (1, -1) with the process \( y_t \), whether they cointegrate with the estimated trends, and whether the estimated trends from the two models cointegrate similarly with each other.

## 3 Cointegration between trends and their estimators

This section gives the main results in three theorems with proofs in the Appendix. In Theorem 1 we show that in the CVAR model, the estimated trend \( \hat{B} \hat{T}_t^* \) cointegrates with the trend \( BT_t^* \), and hence with \( y_t \), by showing that \( y_t - BT_t^* \) and \( BT_t^* - \hat{B} \hat{T}_t^* \) are asymptotically stationary. However, whether \( T_t^* - \hat{T}_t^* \) is asymptotically stationary depends on the identification of \( B, T_t \), and \( \Omega_\eta \). Theorem 2 gives the same results for the state space model, and finally in Theorem 3, we compare the estimated trends in the two models and show that \( \hat{B}T_t - \hat{B} \hat{T}_t^* \) is asymptotically stationary, but the same does not in general hold for \( \hat{T}_t - \hat{T}_t^* \).

The conclusion is that in terms of cointegration of the trends, it does not matter which model we use, as long as we focus on identified trends \( BT_t \) and \( BT_t^* \).

The missing cointegration between \( T_t \) and \( \hat{T}_t \), say, can be explained in terms of the identity

\[
    \hat{B}(T_t - \hat{T}_t) = (\hat{B} - B) T_t + (BT_t - \hat{B} \hat{T}_t),
\]
where \( BT_t - \hat{B} \hat{T}_t \) is asymptotically stationary by Theorem 2 (b), but \((\hat{B} - B)T_t\) is not necessarily asymptotically stationary, because in general \( \hat{B} - B = O_P(n^{-1/2}) \) and \( T_t = O_P(n^{1/2}) \). If, however, \( B \) is identified as in (10), then we have asymptotic stationarity for \( T_t - \hat{T}_t = (I_m, 0_{m \times (p-m)})(BT_t - \hat{B} \hat{T}_t) \). Thus, identification and estimation of parameters can explain the lack of cointegration between estimated and true trends, see Examples 1 and 2 for an illustration.

**Theorem 1** Let \( y_t \) and \( T_t \) be generated by the DGP given in (2). If we use the CVAR (3) for inference, and define the trend \( T_t^* = (\alpha_1' \Gamma B)^{-1} \alpha_1' \sum_{s=1}^t \delta_s \), with estimator \( \hat{T}_t^* \), then

(a) \( y_t - BT_t^* \) is asymptotically stationary,
(b) \( BT_t^* - \hat{B} \hat{T}_t^* \) is asymptotically stationary,
(c) \( T_t^* - \hat{T}_t^* \) is not necessarily asymptotically stationary.

If we choose the parametrization (10), then \( (I_m, 0_{m \times (p-m)})(BT_t^* - \hat{B} \hat{T}_t^*) = T_t^* - \hat{T}_t^* \), such that also \( T_t^* - \hat{T}_t^* \) is asymptotically stationary, but for the parametrization (11), the result does not hold.

In the state space model we can prove similar results

**Theorem 2** Let \( y_t \) and \( T_t \) be generated by the DGP given in (2). If we use the state space model defined by (2) for inference, see Lemma 1, then

(a) \( T_t - E_t T_t \) is asymptotically stationary,
(b) \( E_t BT_t - \hat{E}_t \hat{B} \hat{T}_t \) is asymptotically stationary,
(c) \( E_t T_t - \hat{E}_t \hat{T}_t \) is not in general asymptotically stationary.

Finally we compare trends estimated in the two models and prove the main result that trends derived from the SSM and CVAR models cointegrate \((1, -1)\), as long as we consider \( B \hat{T}_t \) and \( \hat{B} \hat{T}_t^* \). For \( \hat{T}_t \) and \( \hat{T}_t^* \) we get the result again, that asymptotic stationarity depends on how \( B \) and the trends are identified.

**Theorem 3** Let \( y_t \) and \( T_t \) be generated by the DGP given in (2). We estimate the trend \( T_t^* = (\alpha_1' \Gamma B)^{-1} \alpha_1' \sum_{s=1}^t \delta_s \) by analysing CVAR (3), and estimate \( E_t T_t \) using SSM (2). If \( \hat{B} \) be the estimator for \( B \) derived from SSM, whereas \( \hat{B} \) is derived from CVAR, then

(a) \( T_t - T_t^* \) is asymptotically stationary,
(b) \( \hat{B} \hat{T}_t - \hat{B} \hat{T}_t^* \) is asymptotically stationary,
(c) \( \hat{T}_t - \hat{T}_t^* \) is not necessarily asymptotically stationary.

### 4 Two examples

We give two examples where \( p = 3 \) and \( m = 2 \). The parameters \( B \) and \( \Omega_\eta \) contain \( 6 + 3 \) parameters, but the \( 3 \times 3 \) matrix \( B \Omega_\eta B' \) is of rank 2 and has only 5 estimable parameters. Thus, we need to impose 4 restrictions to identify the parameters.

**Example 1.** We first consider the model with the identification given in (11),

\[
B = \begin{pmatrix}
1 & 0 \\
a_{21} & 1 \\
a_{31} & a_{32}
\end{pmatrix}, \quad \Omega_\eta = \begin{pmatrix}
\sigma_{11} & 0 \\
0 & \sigma_{22}
\end{pmatrix},
\]

(17)
such that
\[ E_t B T_t - \hat{E}_t B T_t = \begin{pmatrix} a_{21} E_t T_{1t} + E_t T_{2t} - \hat{a}_{21} \hat{E}_t T_{1t} - \hat{E}_t T_{2t} \\ a_{31} E_t T_{1t} + a_{32} E_t T_{2t} - \hat{a}_{31} \hat{E}_t T_{1t} - \hat{a}_{32} \hat{E}_t T_{2t} \end{pmatrix}. \] (18)

By the results in Theorem 2 (b), we see that \( E_t T_{1t} - \hat{E}_t T_{1t} \) is asymptotically stationary, because it can be recovered as the first row of \( E_t B T_t - \hat{E}_t B T_t \) in (18). Moreover the second row of (18) \( (a_{21} E_t T_{1t} - \hat{a}_{21} \hat{E}_t T_{1t}) + (E_t T_{2t} - \hat{E}_t T_{2t}) \) is asymptotically stationary. Thus to prove that \( E_t T_{2t} - \hat{E}_t T_{2t} \) is asymptotically stationary, it is enough to show stationarity of
\[ a_{21} E_t T_{1t} - \hat{a}_{21} \hat{E}_t T_{1t} = (a_{21} - \hat{a}_{21}) E_t T_{1t} - \hat{a}_{21} (E_t T_{1t} - \hat{E}_t T_{1t}). \] (19)

Here \( \hat{a}_{21}(E_t T_{1t} - \hat{E}_t T_{1t}) \) is asymptotically stationary because \( E_t T_{1t} - \hat{E}_t T_{1t} \) is, but the first term is not, because \( \hat{a}_{21} \) is \( n^{1/2} \)-consistent, and in this case \( n^{1/2}(a_{21} - \hat{a}_{21}) \) is \( Z \) Gaussian, such that
\[ (a_{21} - \hat{a}_{21}) E_t T_{1t}\big|_{t=[n]} = n^{1/2}(a_{21} - \hat{a}_{21}) n^{-1/2} E_t T_{1t}\big|_{t=[n]} \xrightarrow{D} Z W_{n_t}(u), \]
which is nonstationary. This argument is a special case of the proof (22).

To illustrate theresults, we simulate data from the model with \( n = 100 \) observations starting with \( T_1 = 0 \), and parameter values \( a_{21} = 0.0 \), \( a_{31} = a_{32} = 0.5 \), \( \sigma_{11} = \sigma_{22} = 1 \). We estimate the parameters by Gaussian maximum likelihood using SsfPack 3.0, see Koopman, Shepard, and Doornik (2008), where we fix \( y_1 = 0 \) and assume a diffuse prior for the initial value \( T_1 \). The results are summarized in Figure 1. The panels a and b show plots of \( (E_t T_{1t}, \hat{E}_t T_{1t}) \) and \( (E_t T_{2t}, \hat{E}_t T_{2t}) \) respectively, and we see that they co-move in each panel. In panels c and d we show the differences \( E_t T_{1t} - \hat{E}_t T_{1t} \) and \( E_t T_{2t} - \hat{E}_t T_{2t} \). We note that the first looks stationary, whereas the second is clearly nonstationary. When comparing with the plot of \( E_t T_{1t} \) in panel a, it appears that the process \( \hat{E}_t T_{1t} \) can explain the nonstationarity of \( E_t T_{2t} - \hat{E}_t T_{2t} \), consistently with equation (19) and Table 1. Finally we analysed the four variables \( (E_t T_{1t}, \hat{E}_t T_{1t}, E_t T_{2t}, \hat{E}_t T_{2t}) \) using a CVAR and we found two cointegrating relations given in Table 1.

<table>
<thead>
<tr>
<th>( E_t T_{1t} )</th>
<th>( \hat{E}<em>t T</em>{1t} )</th>
<th>( E_t T_{2t} )</th>
<th>( \hat{E}<em>t T</em>{2t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_1^1 )</td>
<td>-1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>( \beta_2^1 )</td>
<td>0</td>
<td>-0.26</td>
<td>-1</td>
</tr>
<tr>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The result of a cointegration analysis of \( E_t T_{1t}, \hat{E}_t T_{1t}, E_t T_{2t}, \) and \( \hat{E}_t T_{2t} \) using model (16). We found two cointegrating relations given in the Table. Note that \( \hat{E}_t T_{1t} \) has a significant coefficient in the second cointegrating relation. See also Figure 1 panel d, where \( E_t T_{2t} - \hat{E}_t T_{2t} \) is plotted, and it is seen that \( \hat{E}_t T_{2t} \) is need to get a stationary relation.

**Example 2.** If instead we choose the parametrization (10)
\[ B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ a_{31} & a_{32} \end{pmatrix}, \quad \Omega_\eta = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}, \] (20)
we still have just identified parameters, but now

\[
E_t BT_{1t} - \hat{E}_t BT_{1t} = \begin{pmatrix}
E_t T_{1t} - \hat{E}_t T_{1t} \\
E_t T_{2t} - \hat{E}_t T_{2t} \\
a_{31} E_t T_{1t} + a_{32} E_t T_{2t} - \hat{a}_{31} \hat{E}_t T_{1t} - \hat{a}_{32} \hat{E}_t T_{2t}
\end{pmatrix}.
\] (21)

Thus, both \( E_t T_{1t} - \hat{E}_t T_{1t} \) and \( E_t T_{2t} - \hat{E}_t T_{2t} \) are asymptotically stationary, as the first two rows of \( E_t BT_{1t} - \hat{E}_t BT_{1t} \), in (21), by Theorem 2.

In the simulation of this example, we use the values \( a_{21} = 0.0, a_{31} = a_{32} = 0.5 \) and \( \sigma_{11} = \sigma_{22} = 1, \sigma_{12} = 0 \), such that in fact the DGP for this example is the same as the DGP for Example 1, namely

\[
B = \begin{pmatrix}
1 & 0 \\
0 & 1 \\
0.5 & 0.5
\end{pmatrix}, \quad \Omega_{\eta} = \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}.
\]

In fact we analysed the same data as in Example 1. We plot the data in Figure 2. The panels \( a \) and \( b \) show plots of \((E_t T^1_{1t}; \hat{E}_t T^1_{1t})\) and \((E_t T^2_{1t}; \hat{E}_t T^2_{1t})\) respectively, and we see that they co-move. In panels \( c \) and \( d \) we show the differences \( \hat{E}_t T_{1t} - E_t T_{1t} \) and \( \hat{E}_t T_{2t} - E_t T_{2t} \), which appear to be stationary in this parametrization of the model. A cointegration analysis of \((E_t T^1_{1t}; \hat{E}_t T^1_{1t}, E_t T^2_{2t}; \hat{E}_t T^2_{2t})\) shows that there two cointegrating relations, see Table 2.

<table>
<thead>
<tr>
<th>( \beta'_1 )</th>
<th>( \beta'_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-1)</td>
<td>(1)</td>
</tr>
<tr>
<td>(1)</td>
<td>(0)</td>
</tr>
<tr>
<td>(-1)</td>
<td>(1)</td>
</tr>
</tbody>
</table>

Table 2: The result of a cointegration analysis of \( E_t T^1_{1t}, \hat{E}_t T^1_{1t}, E_t T^2_{2t}, \hat{E}_t T^2_{2t} \) using model (19). We found the two cointegrating relations given here.

5 Conclusion

We have analysed data generated by a multivariate homogenous linear state space model

\[
\begin{align*}
y_{t+1} &= B^0 T_t + \varepsilon_{t+1}, \\
T^0_{t+1} &= T^0_t + \eta_{t+1},
\end{align*}
\]

where the state variable, \( T^0_t \), is an unobserved multivariate random walk. We have used a state space model and a CVAR model to estimate the trend in the data. In both cases we find that the estimated trend cointegrates \((1, -1)\) with the trend in the model, provided we only consider \( BT_t \) and its estimator. The same result does not hold in general for the trend \( T_t \) and its estimator. The reason is that if the parameters are identified using only restrictions on \( \Omega_{\eta} \), the estimator of \( B \) is not \( n \)-consistent, and recovering the trend \( \hat{T}_t \) from the \( \hat{B} T_t \) causes some trends not to cointegrate with estimated trends.
6 Appendix

Proof of Lemma 1. Proof of (7): We define $M=(\hat{B}, B_\perp)$, for $\hat{B} = B(B' B)^{-1}$, and find
\[
K_t B = V_t B' M [(M' B V_t B' M + M' \Omega_v M)]^{-1} M' B \\
= V_t \left( I_m \begin{array}{c} 0 \end{array} \right)' \left( V_t B' \Omega_v \hat{B} B' \Omega_v B_\perp B' \Omega_v B_\perp \right)^{-1} \left( I_m \begin{array}{c} 0 \end{array} \right) = V_t (V_t + \Omega_B)^{-1},
\]
where
\[
\Omega_B = \hat{B}' [\Omega_v - \Omega_v B_\perp (B_\perp \Omega_v B_\perp)^{-1} B_\perp \Omega_v] \hat{B} = \text{Var}(\hat{B}' \varepsilon_t | B'_\perp \varepsilon_t).
\]

Proof of (8): If the recursion starts with $V_1 = \Omega_\eta$, then all $V_t$ can be diagonalized by $W$, such that the recursion for $\lambda_{i,t}$, the eigenvalue of $V_t$, becomes
\[
\lambda_{i,t+1} = \lambda_i + \lambda_{i,t} - \frac{\lambda_{i,t}^2}{1 + \lambda_{i,t}}.
\]
It is seen that $\lambda_{i,t}$ is increasing in $t$ and the limit is $\lambda_{i,\infty} = \{\lambda_i + (\lambda_i^2 + 4\lambda_i)^{1/2}\}/2$. See for instance Chan, Goodwin and Sin (1984) for more general results.

Proof of Lemma 2. Consider first the product moments and their limits (12) and (13),
\[
S_{1n} = n^{-1} \sum_{i=1}^n \Delta y_i \Delta y_i' \overset{P}{\to} B \Omega_\eta B' + 2 \Omega_\varepsilon,
\]
\[
S_{2n} = n^{-1} \sum_{i=1}^n \Delta y_i \Delta y_{i+1}' \overset{P}{\to} -\Omega_\varepsilon.
\]
Irrespective of the identification, $B \Omega_\eta B'$ and $\Omega_\varepsilon$ can be estimated $n^{1/2}$-consistently with Gaussian limit distribution using the Central Limit Theorem.

To prove (16), we find from (14) and (15), that the least squares estimator $\hat{\gamma}_{\text{reg}}$ satisfies
\[
n(\hat{\gamma}_{\text{reg}} - \gamma) = -(n^{-2} \sum_{i=1}^{n-1} z_{1,t+1} z_{1,t+1}')^{-1} n^{-1} \sum_{i=1}^{n-1} z_{1,t+1} \varepsilon_{t+1}' B_\perp \\
\overset{D}{\to} - (\int_0^1 W_\eta W_\eta' (du)^{-1} \int_0^1 W_\eta (dW_\varepsilon)' B_\perp).
\]
Here $W_\eta$ is Brownian motion generated from $\eta_t$, adjusted to the identification of $B = (I_m, \gamma)'$, and $W_\varepsilon$ is Brownian motion generated from $\varepsilon_t$. We choose $B_\perp = (\gamma_0', -I_{p-m})'$ and find the relations
\[-(\hat{B} - B)' B_\perp = \hat{\gamma}_{\text{reg}} - \gamma = B'(\hat{B}_\perp - B_\perp).
\]
Note that for the other parametrization (11), where $B = (C_\eta', C_\eta' \gamma)'$ and we can choose the same $B_\perp = (\gamma_0', -I_{p-m})'$, such that for both parametrizations we have (16). The estimator of $B$, however, changes in the parametrization (11), and we find
\[
\hat{B} = \begin{pmatrix} \hat{C}_\eta' \\ \hat{\gamma}' C_\eta \end{pmatrix},
\]
where $\hat{C}_n$ is derived from the $n^{1/2}$-consistent estimator of $\Omega_\eta$, such that for this parametrization, we do not get $n$-consistent estimation of $B$, but only that $\hat{B} - B = O_P(n^{-1/2})$.

**Proof of Theorem 1.** Proof of (a): This follows from (5).

Proof of (b): The Granger representation (5) holds for parameters and residuals, but because $y_t$ is also a solution to the equations with estimated parameters and estimated residuals, we have the same representation in terms of these. This implies that

$$BT_t^* - \hat{B}T_t^* = C \sum_{s=1}^{t} \delta_s - \hat{C} \sum_{s=1}^{t} \hat{\delta}_s = y_t - u_t - (y_t - \hat{u}_t) = \hat{u}_t - u_t,$$

which is an asymptotically stationary process.

Proof of (c): We next find from

$$\hat{B}(T_t^* - \hat{T}_t^*) = (BT_t^* - \hat{B}T_t^*) + (\hat{B} - B)T_t^*,$$

that the first term is stationary by (b), and the last term $(\hat{B} - B)T_t^* = O_P(1)$ is not stationary because for the parametrization (11) we have only $n^{1/2}$-consistent estimators, such that

$$n^{1/2}(\hat{B} - B) = O_P(1) \text{ and } n^{-1/2}T_n^* \xrightarrow{D} W_\delta(u), \quad (22)$$

where $W_\delta$ is Brownian motion generated by $\delta_t$. Thus, in general $\hat{B}(T_t^* - \hat{T}_t^*)$ and therefore $T_t^* - \hat{T}_t^*$ is not asymptotically stationary.

**Proof of Theorem 2.** Proof of (a): We define the deviation between $T_t$ and the extracted trend, $v_t = T_t - \hat{E}_tT_t$, and find

$$y_{t+1} - \hat{E}_t y_{t+1} = BT_t + \varepsilon_{t+1} - \hat{E}_t BT_t = Bv_t + \varepsilon_{t+1}, \quad (23)$$

$$T_{t+1} - \hat{E}_t T_t = T_t - \hat{E}_t T_t + \eta_{t+1} = v_t + \eta_{t+1}. \quad (24)$$

From the Kalman filter equations (6), we find

$$v_{t+1} = T_{t+1} - \hat{E}_{t+1}T_{t+1} = T_{t+1} - \hat{E}_t T_t - K_t(y_{t+1} - \hat{E}_t y_{t+1}), \quad (25)$$

such that, using (23) and (24),

$$v_{t+1} = v_t + \eta_{t+1} - K_t(Bv_t + \varepsilon_{t+1}) = (I_m - K_tB)v_t - K_t\varepsilon_{t+1} + \eta_{t+1}. \quad (26)$$

From (9) we have that $I_m - K_tB \rightarrow \Omega_B(V + \Omega_B)^{-1}$, which has positive eigenvalues less than 1, and is hence stable. Moreover $\varepsilon_{t+1}$ is i.i.d. and the AR coefficient $I_p - K_tB$ is stable for large $t$, so we find that (25) determines an asymptotically stationary process for $T_t - \hat{E}_tT_t$.

Proof of (b): We define $u_t = BT_t - \hat{E}_tBT_t$, and find

$$\hat{E}_tBT_t - \hat{E}_tBT_t = (\hat{E}_tBT_t - BT_t) + (BT_t - \hat{E}_tBT_t) = B(E_tT_t - T_t) + u_t.$$

The first term is stationary by (a), such that it is enough to show that $u_t$ is asymptotically stationary. From (23), (24), and (25) we find, using estimated values of parameters,

$$y_{t+1} - \hat{E}_t y_{t+1} = BT_t + \varepsilon_{t+1} = \hat{E}_tBT_t = u_t + \varepsilon_{t+1}, \quad (27)$$

where $\hat{E}_tBT_t$ is derived from the $n^{1/2}$-consistent estimation of $B$, and is hence stable. Moreover $\varepsilon_{t+1}$ is i.i.d. and the AR coefficient $I_p - \hat{E}_tBT_t$ is stable for large $t$, so we find that (27) determines an asymptotically stationary process for $\hat{E}_tBT_t$.
and the recursion
\[ u_{t+1} = (I_p - \hat{B} \hat{K}_t)u_t - \hat{B} \hat{K}_t \varepsilon_{t+1} + B \eta_{t+1}. \]  
(26)

From the definition of \( u_t \), we find from (16), that
\[ \hat{B}' u_t = \hat{B}' BT_t = \hat{B}'(B - \hat{B})T_t = -\hat{B}'(B_\perp - \hat{B}_\perp)T_t = O_P(n^{-1})O_P(n^{1/2}) = O_P(n^{-1/2}). \]

Next we multiply (26) by \( \hat{B}' \) and use \( \hat{B}' B = \hat{B}' (B - \hat{B}) + I_m = I_m + O_P(n^{-1/2}) \),
to find
\[
\hat{B}' u_{t+1} = (\hat{B}' - \hat{K}_t)u_t - \hat{K}_t \varepsilon_{t+1} + \hat{B}' B \eta_{t+1}
= (\hat{B}' - \hat{K}_t)(\hat{B} \hat{B}' + \hat{B}_\perp \hat{B}_\perp)u_t - \hat{K}_t \varepsilon_{t+1} + \hat{B}' (B - \hat{B}) \eta_{t+1}
= (I_m - \hat{K}_t \hat{B}) \hat{B}' u_t - \hat{K}_t \varepsilon_{t+1} + \eta_{t+1} + O_P(n^{-1/2}),
\]
because both \( \hat{B}' (B - \hat{B}) \eta_{t+1} \) and \( (I_m - \hat{K}_t) \hat{B}_\perp \hat{B}_\perp u_t \) are \( O_P(n^{-1/2}) \). From (9) we see that
\[ I_m - \hat{K}_t \hat{B} \xrightarrow{P} \Omega_B(V + \Omega_B)^{-1} \]
is stable for large \( n \) and \( t \), which shows that \( \hat{B}' u_{t+1} \) and hence \( u_t \) is asymptotically stationary.

**Proof of (c):** We apply previous results and find from the identity
\[ \hat{B}(T_t - \hat{E}_t T_t) = (\hat{B} - B)T_t + (B T_t - \hat{B} \hat{E}_t T_t), \]
that \( B T_t - \hat{B} \hat{E}_t T_t \) is asymptotically stationary by (b). For the second term we find \( n^{1/2}(\hat{B} - B) = O_P(1) \) and \( n^{-1/2}T_{[m]} \xrightarrow{D} W_n \), see (22), such that \( (\hat{B} - B)T_t \) is not necessarily asymptotically stationary. Hence \( T_t - \hat{E}_t T_t \) is not in general a stationary process. □

**Proof of Theorem 3.** **Proof of (a):** This follows from (4) and (5).

**Proof of (b):** We compare each estimator with the corresponding trend and find
\[ \hat{B} \hat{T}_t - \hat{B} \hat{T}_t^* = (\hat{B} \hat{T}_t - BT_t) + (BT_t - BT_t^*) + (BT_t^* - \hat{B} \hat{T}_t^*). \]
Here the first term is asymptotically stationary using Theorem 2(b), and the last is asymptotically stationary by Theorem 1(b, c), and the middle term is asymptotically stationary by (a).

**Proof of (c):** We find similarly
\[ \hat{T}_t - \hat{T}_t^* = (\hat{T}_t - T_t) + (T_t - T_t^*) + (T_t^* - \hat{T}_t^*), \]
(27)
where \( T_t - T_t^* \) is asymptotically stationary by (a). For the first term, we decompose as
\[ \hat{B}(T_t - \hat{T}_t) = (\hat{B} - B)T_t + (B T_t - \hat{B} \hat{T}_t), \]
where \( (\hat{B} - B)T_t \) is asymptotically nonstationary, because \( \hat{B} \) is in general only \( n^{1/2} \)-consistent. The same argument applies to the last term in (27), which shows that in general we do not get asymptotic stationarity. □
7 References


Figure 1: The result of a simulation of model (17) are plotted. Panels a and b show plots of $\hat{E}_t T_1$ and $E_t T_1$, and $\hat{E}_t T_2$ and $E_t T_2$, respectively. We note that in both cases, the processes seem to co-move. In panel c we have plotted $\hat{E}_t T_1 - E_t T_1$ which appears stationary, but in panel d we note that the spread $\hat{E}_t T_2 - E_t T_2$ is nonstationary. This is accordance with the finding in Table 1, where it is seen that $\hat{E}_t T_1$ is needed to find cointegration, see (18).

Figure 2: The results of a simulation of model (20) are plotted. Panels a and b show plots of $\hat{E}_t T_1$ and $E_t T_1$, and $\hat{E}_t T_2$ and $E_t T_2$, respectively. We note that in both cases, the processes seem to co-move. In panels c and d, we have plotted $\hat{E}_t T_1 - E_t T_1$ and $\hat{E}_t T_2 - E_t T_2$, which now appear stationary, because they are both recovered from $\hat{E}_t AT_t - E_t AT_t$ as the first two coordinates see (19).
2016-27: Kim Christensen, Ulrich Hounyo and Mark Podolskij: Testing for heteroscedasticity in jumpy and noisy high-frequency data: A resampling approach

2016-28: Kim Christensen, Roel Oomen and Roberto Renò: The Drift Burst Hypothesis


2016-30: Morten Ørregaard Nielsen and Sergei S. Shibaev: Forecasting daily political opinion polls using the fractionally cointegrated VAR model

2016-31: Carlos Vladimir Rodríguez-Caballero: Panel Data with Cross-Sectional Dependence Characterized by a Multi-Level Factor Structure

2016-32: Lasse Bork, Stig V. Møller and Thomas Q. Pedersen: A New Index of Housing Sentiment

2016-33: Joachim Lebovits and Mark Podolskij: Estimation of the global regularity of a multifractional Brownian motion

2017-01: Nektarios Aslanidis, Charlotte Christiansen and Andrea Cipollini: Predicting Bond Betas using Macro-Finance Variables

2017-02: Giuseppe Cavaliere, Morten Ørregaard Nielsen and Robert Taylor: Quasi-Maximum Likelihood Estimation and Bootstrap Inference in Fractional Time Series Models with Heteroskedasticity of Unknown Form

2017-03: Peter Exterkate and Oskar Knapik: A regime-switching stochastic volatility model for forecasting electricity prices

2017-04: Timo Teräsvirta: Sir Clive Granger’s contributions to nonlinear time series and econometrics

2017-05: Matthew T. Holt and Timo Teräsvirta: Global Hemispheric Temperatures and Co-Shifting: A Vector Shifting-Mean Autoregressive Analysis

2017-06: Tobias Basse, Robinson Kruse and Christoph Wegener: The Walking Debt Crisis

2017-07: Oskar Knapik: Modeling and forecasting electricity price jumps in the Nord Pool power market

2017-08: Malene Kallestrup-Lamb and Carsten P.T. Rosenskjold: Insight into the Female Longevity Puzzle: Using Register Data to Analyse Mortality and Cause of Death Behaviour Across Socio-economic Groups


2017-10: Jeroen V.K. Rombouts, Lars Stentoft and Francesco Violante: Dynamics of Variance Risk Premia, Investors’ Sentiment and Return Predictability

2017-11: Søren Johansen and Morten Nyboe Tabor: Cointegration between trends and their estimators in state space models and CVAR models