Consistent cotrending rank selection when both stochastic and nonlinear deterministic trends are present

Zheng-Feng Guo* and Mototsugu Shintani †

This version: April 2010

Abstract

This paper proposes a model-free cotrending rank selection procedure based on the eigenstructure of a multivariate version of the von Neumann ratio, in the presence of both stochastic and nonlinear deterministic trends. Our selection criteria are easily implemented and the consistency of the rank estimator is established under very general conditions. Simulation results suggest good finite sample properties of the new rank selection criteria. The proposed method is then illustrated through an application of Japanese money demand function allowing for the cotrending relationship among money, income and interest rates.

*Department of Economics, Vanderbilt University, VU Station B #351819, Nashville, TN 37235-1819. Email: zhengfeng.guo@vanderbilt.edu
†Department of Economics, Vanderbilt University, VU Station B #351819, Nashville, TN 37235-1819. Email: mototsugu.shintani@vanderbilt.edu
1 Introduction

For decades, one of the most important issues in the analysis of macroeconomic time series has been how to incorporate a trend. Two popular approaches that have been often employed in the literature are (i) to consider a stochastic trend (with or without a linear deterministic trend), such as the one suggested in Nelson and Plosser (1982) and (ii) to consider a nonlinear deterministic trend such as the one with trend breaks considered in Perron (1989, 1997). Cointegration, introduced by Engle and Granger (1987), is a useful concept in understanding the nature of comovement among variables based on the first approach. In cointegration analysis, the cointegrating rank, defined as the number of linearly independent cointegrating vectors, provides valuable information regarding the trending structure of a multivariate system with stochastic trends. Several model-free consistent cointegrating rank selection procedures have been developed in the literature. Analogous to cointegration analysis, we can also investigate the nature of comovement based on the second approach, namely the nonlinear deterministic trend. The cotrending analyses of Bierens (2000), Hatanaka (2000) and Hatanaka and Yamada (2003) are along this line of research. However, the consistent selection procedure of the cotrending rank, defined similarly as the cointegrating rank with a stochastic trend replaced by a nonlinear deterministic trend, has not yet been developed.

This paper proposes a model-free consistent cotrending rank selection procedure when both stochastic and nonlinear deterministic trends are present in a multivariate system. Consistency here refers to the property that the probability of selecting the wrong cotrending rank approaches zero as sample size tends to infinity. Our procedure selects the cotrending rank by minimizing the von Neumann criterion, similar to the one used by Shintani (2001) and Harris and Poskitt (2004) in their analyses of cointegration. This approach exploits the fact that identification of cotrending rank can be interpreted as identification among three groups of eigenvalues of the generalized von Neumann ratio. Using this property of the von Neumann criterion, we propose two types of cotrending rank selection procedures that are (1) invariant
to linear transformation of data; (2) robust to misspecification of the model; and (3) valid not only with a break in the trend function but also with a certain class of a smooth nonlinear trend function. The simulation results also suggest that our cotrending rank procedures perform well in the finite sample.

Our analysis is closely related to Harris and Poskitt (2004) and Cheng and Phillips (2009), who propose consistent cointegrating rank procedures that do not require a parametric vector autoregressive model of cointegration such as the one in Johansen (1991). While we mainly discuss a class of smooth transition trend models as our preferred specification of the nonlinear trend function, our cotrending rank procedure does not rely on the parametric specification of the trend function, nor the parametric specification of serial dependence structure. Thus, our approach is more general than that of Harris and Poskitt (2004) and Cheng and Phillips (2009) in the sense it allows both common stochastic trends and common deterministic trends. Consequently, we can use the procedure to determine the cointegrating rank in the absence of nonlinear deterministic trends. To illustrate this flexibility, we also provide the simulation results to evaluate the performance of our procedure when it is applied to cointegrating rank selection.

As emphasized in Stock and Watson (1988), the cointegrated system can be interpreted as a factor model with a stochastic trend being a common factor. Thus, determining the cointegrating rank is identical to determining the number of common stochastic trends because the latter is the difference between the dimension of the system (number of variables) and the cointegrating rank.\footnote{PANIC method proposed by Bai and Ng (2004) utilizes the consistent selection of the number of common stochastic trends in a very large dynamic factor system based on information criteria. See also Bai and Ng (2002) for the case of consistent selection of the number of stationary common factors.} In the presence of both stochastic and nonlinear deterministic trends, however, the number of common nonlinear deterministic trends does not correspond to the difference between the dimension and the cotrending rank. Because the number of common deterministic trends also contains valuable information on the trending structure, we introduce the notion of weak cotrending rank so that the difference between the dimension and the weak cotrending rank becomes the number of common deterministic trends. Our procedure can
select both a cotrending rank and a weak cotrending rank.

Our cotrending notion can be interpreted as a type of a common feature discussed in Engle and Kozicki (1993). They define the common feature as “a feature that is present in each of a group of series but there exists a non-zero linear combination of the series that does not have the feature”. Our definition of cotrending also requires a non-zero linear combination to eliminate nonlinear deterministic trends. Our cotrending analysis is also related to cobreaking analysis of Hendry and Mizon (1998) and Clements and Hendry (1999). In the presence of a trend break, cobreaking is a necessary condition of cotrending but not a sufficient condition.

The remainder of this paper is organized as follows. Section 2 introduces several techniques to model deterministic trends and some key concepts in the presence of stochastic and deterministic trends. The main theoretical results are given in section 3. Section 4 reports Monte-Carlo simulation results to show the finite sample performance of our procedures. In section 5, we apply our procedures to the Japanese money demand function. Section 6 concludes and the technical proofs are presented in the Appendix.

2 Motivation

2.1 A smooth transition trend model

The traditional approach to introduce a deterministic trend in a scalar time series $\{y_t\}_{t=1}^T$ is to consider a simple linear trend model given by

$$y_t = d_t^{LIN} + \varepsilon_t$$

and

$$d_t^{LIN} = c + \mu t,$$
where $d_t^{LIN}$ is a linear trend function and $\varepsilon_t$ is a zero-mean stationary error component. Serial correlation is typically allowed in the error term, but we focus on the case of i.i.d. white noise to simplify the argument. Using this linear trend model to fit $y_t$, measured in logarithms, is appropriate when the (log) growth rate of the variable is stable over the sample period around the average (log) growth rate of $\mu$. However, as discussed in Mills (2003), many macroeconomic time series data, including GDP of the UK and Japan, and stock prices in the U.S., violated the assumption of stable growth over the sample.

A convenient approach to allow for multiple shifts in the average growth rate while maintaining the continuity of the trend function is to consider a kinked trend, or a piece-wise linear trend structure in each segment of the whole sample period. When there are $h$ time shifts in the average growth rate, a segmented linear trend model can be defined as

\[
y_t = d^KINK_t + \varepsilon_t,
\]

\[
d^KINK_t = c + \mu_0 t + \sum_{i=1}^{h} \mu_i (t - T_i) 1[t > T_i],
\]

where $d^KINK_t$ is a piece-wise linear trend function, $T_i$ is the break point, and $1[x]$ is an indicator which takes the value of 1 if $x$ is true and 0 otherwise. The segmented linear trend model above implies that the average growth rate during the first subperiod $t < T_1$ corresponds to $\mu_0$, and the remaining subperiods, $T_j \leq t < T_{j+1}$ for $j = 1, ..., h$ corresponds to $\mu_0 + \sum_{i=1}^{j} \mu_i$. As long as the break point is known, the model can be estimated using the least squares estimator.

Although the segmented trend function $d^KINK_t$ imposes continuity, its first derivative is not continuous, suggesting an abrupt change of the growth rate at each break point. To allow for a gradual change in the growth rate, we may replace the indicator function in $d^KINK_t$ with a smooth transition function. This substitution of the trend function leads to a smooth transition trend model. The smooth transition trend model was originally proposed by Bacon and Watt (1971) and has been discussed by Granger and
Teräsvirta (1993), Teräsvirta (1994) and Greenaway, Leybourne and Sapsford (1997). While there are many types of smooth transition trend functions, the most frequently used one is the logistic transition function given by

\[ G(\gamma_i, T_i) = \frac{1}{1 + \exp(-\gamma_i(t - T_i))}, \]

where \( \gamma_i \) is the scaling parameter which controls the speed of transition, and \( T_i \) becomes the timing of the transition midpoint instead of the break point.

A multiple-regime logistic smooth transition trend (LSTT) model takes the form of

\[ y_t = d_{LST} + \varepsilon_t, \]
\[ d_{LST} = c + \mu_0 t + \sum_{i=1}^{h} \mu_i(t - T_i)G(\gamma_i, T_i) \]

where \( d_{LST} \) is the nonlinear linear trend function which we mainly focus on in our cotrending analysis. It should be noted that as \( \gamma_i \) approaches infinity, the logistic transition function \( G(\gamma_i, T_i) \) approaches the indicator function \( 1[t > T_i] \). Thus, our deterministic trend \( d_{LST} \) nests both the kinked trend \( d_{KINK} \) and the linear trend \( d_{LIN} \) as a special case. Figure 2.1 shows the typical shape of kinked and smooth transition trends when \( h = 1 \). The former contains a one-time abrupt change in the first derivative while the latter shows continuous change in the first derivative.

**2.2 Cointegration and common deterministic trend**

(1) Stochastic trends only

Cointegration, a very important concept in trending analysis of macroeconomic time series, was first introduced by Engle and Granger (1987). Multivariate integrated processes of order one are cointegrated if some linear combination of the same series becomes stationary. For example, suppose that a pair of
variables $y_t = (y_{1t}, y_{2t})'$ are generated from two random walk processes as in

$$
y_{1t} = s_{1t}, \quad y_{2t} = s_{2t},
$$

$$
s_{1t} = s_{1t-1} + \varepsilon_{1t}, \quad \text{and}
$$

$$
s_{2t} = s_{2t-1} + \varepsilon_{2t}.
$$

Since there is no linear combination of $y_{1t}$ and $y_{2t}$ that yields stationarity, $y_t$ is not cointegrated. Instead, suppose that the pair are generated from one random walk component as in

$$
y_{1t} = s_{1t}, \quad y_{2t} = s_{2t},
$$

$$
s_{1t} = s_{1t-1} + \varepsilon_{1t} \quad \text{and}
$$

$$
s_{2t} = s_{1t} + \varepsilon_{2t}.
$$

Since $y_{1t} - y_{2t} = -\varepsilon_{2t}$ is stationary, $y_t$ is cointegrated with a cointegrating vector $(1, -1)$. This simple example shows that cointegration requires the pair to share a common stochastic trend $s_{1t}$.

(2) Deterministic trends only

Let us now turn to the case of deterministic trends. We refer cotrending among the variables with deterministic trends to the case when variables share a common deterministic trend. Suppose that each variable in the pair $y_t = (y_{1t}, y_{2t})'$ is generated from the linear trend model as in

$$
y_{1t} = d^{LIN}_{1t}, \quad y_{2t} = d^{LIN}_{2t},
$$

where $d^{LIN}_{1t}$ and $d^{LIN}_{2t}$ are linear trend functions with potentially different intercepts and trend slope coefficients. Unlike the preceding example, we can always find the linear combination that eliminates the trend function (namely, the linear combination becomes a constant). Thus, the notion of cotrending in
the case of a linear trend is trivial. However, if the linear trend functions, \(d_{1t}^{LIN}\) and \(d_{2t}^{LIN}\), are replaced with segmented trend functions, \(d_{1t}^{KINK}\) and \(d_{2t}^{KINK}\), we can eliminate the deterministic trend if and only if (i) all of the break points, \(T_i\)’s, are the same (common break) and (ii) all of the piece-wise trend slope coefficients \(\mu_i\) are proportional between the two trend functions. If either of the two conditions fails to hold, the two nonlinear deterministic trends are linearly independent and no common deterministic trend exists. A similar arguments apply when the segmented linear trends are replaced with smooth transition trends.

(3) Both stochastic and deterministic trends

When variables contain both stochastic and deterministic trends, there will be a layer of potential cotrending relationships. Suppose that a pair of variables, \(y_t = (y_{1t}, y_{2t})'\), are generated from two random walk with a drift processes as in

\[
y_{1t} = \mu_1 + y_{1t-1} + \varepsilon_{1t},
\]
\[
y_{2t} = \mu_2 + y_{2t-1} + \varepsilon_{2t},
\]

where \(\mu_1 \neq 0\) and \(\mu_2 \neq 0\). Then, the model can be rewritten as

\[
y_{1t} = d_{1t}^{LIN} + s_{1t}, \quad y_{2t} = d_{2t}^{LIN} + s_{2t},
\]

where

\[
d_{1t}^{LIN} = s_{10} + \mu_1 t, \quad d_{2t}^{LIN} = s_{20} + \mu_2 t,
\]
\[
s_{1t} = s_{1t-1} + \varepsilon_{1t}, \quad \text{and}
\]
\[
s_{2t} = s_{2t-1} + \varepsilon_{2t}.
\]
In this case, the vector \((1, -\mu_1/\mu_2)\) eliminates the linear deterministic trend but does not eliminate the stochastic trend. In contrast, if the pair is generated from

\[
y_{1t} = d_{1t}^{LIN} + s_{1t}, \quad y_{2t} = d_{2t}^{LIN} + s_{2t},
\]

where

\[
d_{1t}^{LIN} = s_{10} + \mu_1 t, \quad d_{2t}^{LIN} = s_{20} + \mu_2 t,
\]

\[
s_{1t} = s_{1t-1} + \varepsilon_{1t}, \quad \text{and}
\]

\[
s_{2t} = s_{1t} + \varepsilon_{2t},
\]

the vector \((1, -1)\) eliminates the stochastic trend, and the vector \((1, -\mu_1/\mu_2)\) eliminates the linear deterministic trend. Therefore, the vector \((1, -1)\) cannot eliminate both stochastic and deterministic trends except for the case when \(\mu_1 = \mu_2\). For the purpose of distinguishing between these two possibilities, Ogaki and Park (1997) introduce the notions of stochastic cointegration and deterministic cointegration. Stochastic cointegration refers to the case in which only the stochastic trend is eliminated while deterministic cointegration refers to the case in which both stochastic and deterministic trends are eliminated at the same time.

While Ogaki and Park (1997) do not consider more general nonlinear deterministic trends, similar concepts can also be used when the linear trends, \(d_{1t}^{LIN}\) and \(d_{2t}^{LIN}\), are replaced by either segmented trends, \(d_{1t}^{KINK}\) and \(d_{2t}^{KINK}\), or smooth transition trends, \(d_{1t}^{LST}\) and \(d_{2t}^{LST}\). In our analysis of cotrending, we are interested in the case of deterministic cointegration when the nonlinear deterministic trends are represented by a class of smooth transition trend models (which includes linear and kinked trends). To be more specific, we identify the total number of linearly independent vectors that can eliminate both stochastic and nonlinear deterministic trends. In this paper, we refer to the number of such cotrending
vectors as the cotrending rank and denote it by \( r_1 \). In a system of \( m \) variables with both stochastic and deterministic trends, the cotrending rank can be any integer value in the range of \( 0 \leq r_1 < m \). When no variable contains deterministic trends, our definition of cotrending rank corresponds to that of usual cointegrating rank.

In addition to the cotrending rank, we also consider the total number of linearly independent vectors that can eliminate the deterministic trend regardless of whether such vectors can eliminate the stochastic trend at the same time. In the above example, a vector \((1, -\mu_1/\mu_2)\) can eliminate the deterministic trend regardless of the values of \( \mu_1 \) and \( \mu_2 \). We refer to this number as the weak cotrending rank and denote it by \( r_2 \). Since all the cotrending vectors are also included in the weak cotrending vector, \( r_2 \) should satisfy \( r_1 \leq r_2 < m \). While it is not a cotrending rank, identification of \( r_2 \) is important since \( m - r_2 \) in the \( m \)-variable-system with both stochastic and deterministic trends corresponds to the total number of common deterministic trends.

In the next section, we propose a simple procedure to identify both \( r_1 \) and \( r_2 \) in a system of \( m \) variables, in the presence of both stochastic and deterministic trends.

3 Theory

We assume that an \( m \)-variate time series, \( y_t = [y_{1t}, \cdots, y_{mt}]' \), is generated by

\[
y_t = d_t + s_t, \quad t = 1, \cdots, T,
\]

where \( d_t = [d_{1t}, \cdots, d_{mt}]' \) is a nonstochastic trend component, \( s_t = [s_{1t}, \cdots, s_{mt}]' \) is a stochastic process, respectively defined below, and neither \( d_t \) nor \( s_t \) are observable. We denote a random (scalar) sequence \( x_T \) by \( O_p(T^\lambda) \) if \( T^{-\lambda}x_T \) is bounded in probability and by \( o_p(T^\lambda) \) if \( T^{-\lambda}x_T \) converges to zero in probability. For a deterministic sequence, we use \( O(T^\lambda) \) and \( o(T^\lambda) \), if \( T^{-\lambda}x_T \) is bounded and converges to zero,
respectively. The first difference of \( x_t \) is denoted by \( \Delta x_t \). Below, we employ a set of assumptions that are similar to the ones in Hatanaka and Yamada (2003).

Assumption 1. (i) \( s_t = s_{t-1} + \xi_t \) and \( \xi_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j} \), \( C_0 = I_n \), \( \sum_{j=0}^{\infty} j^2 \| C_j \| < \infty \), where \( \varepsilon_t \) is iid with zero mean and covariance matrix \( \Sigma_{\varepsilon\varepsilon} > 0 \). (ii) Each element of \( \sum_{t=1}^{T} d_t \) is \( O(T^2) \) and is not \( o(T^2) \). (iii) There exists an \( m \times m \) orthogonal full rank matrix \( B = [B_\perp B_2 B_1] \) such that each element of \( \sum_{t=1}^{T} B_1' y_t \) is \( O_p(T^{1/2}) \), each element of \( \sum_{t=1}^{T} B_2' y_t \) is \( O_p(T) \) and is not \( o_p(T) \), and each element of \( \sum_{t=1}^{T} B_\perp' y_t \) is \( O_p(T^2) \) is not \( o_p(T^2) \), where \( B_1, B_2, B_\perp \) are \( m \times r_1 \), \( m \times (r_2 - r_1) \) and \( m \times (m - r_2) \), respectively.

Under Assumption 1, \( B_1 \) represents a set of cotrending vectors which eliminates both deterministic and stochastic trends. \( B_2 \) represents a set of vectors which eliminates only deterministic trends, but not stochastic trends. \( B_\perp \) consists of vectors orthogonal to \( B_1 \) and \( B_2 \).

In scalar case, the von Neumann ratio is defined as the sample variances of the differences and the level of a time series. The multivariate generalization of the von Neumann ratio is defined as \( S_{11}^{-1} S_{00} \) where

\[
S_{11} = T^{-1} \sum_{t=1}^{T} y_t y_t', \quad \text{and} \quad S_{00} = T^{-1} \sum_{t=2}^{T} \Delta y_t \Delta y_t'.
\]

Shintani (2001) and Harris and Poskitt (2004) also use this multivariate version of the von Neumann ratio in cointegration analysis. Let \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0 \) be the eigenvalues of \( S_{11}^{-1} S_{00} \). We summarize the statistical properties of \( \lambda_i \)'s in the presence of both stochastic and deterministic trends in the following lemma.

Lemma 1 Under Assumption 1, we have: (i) a sequence of \( [\lambda_1, \cdots, \lambda_{r_1}] \) has a positive limit and is \( O_p(1) \) but is not \( o_p(1) \); (ii) a sequence of \( T[\lambda_{r_1+1}, \cdots, \lambda_{r_2}] \) has a positive limit and is \( O_p(1) \) but is not \( o_p(1) \), provided \( r_2 - r_1 > 0 \); and (iii) a sequence of \( T^2[\lambda_{r_2+1}, \cdots, \lambda_m] \) has a positive limit and is \( O_p(1) \).
but is not $o_p(1)$, provided $m - r_2 > 0$.

From Lemma 1, the eigenvalues of $S_{11}^{-1}S_{00}$ can be classified into three groups depending on their rates of convergence, namely, $O_p(1)$, $O_p(T^{-1})$, and $O_p(T^{-2})$. The number of eigenvalues in each group corresponds to the number of cotrending relationships, the difference between weak cotrending and (strict) cotrending relationships, and the number of common deterministic trends, respectively. We exploit this property to construct the following two types of consistent cotrending rank selection procedures based on the von Neumann criterion defined as a sum of the partial sum of eigenvalues and a penalty term. The first is a paired procedure which independently selects the cotrending rank $r_1$ and the weak cotrending rank $r_2$ by minimizing each of

\[
V N_1(r_1) = -\sum_{i=1}^{r_1} \hat{\lambda}_i + f(r_1) \frac{C_T}{T}, \text{ and}
\]

\[
V N_2(r_2) = -\sum_{i=1}^{r_2} \hat{\lambda}_i + f(r_2) \frac{C'_T}{T^2},
\]

or

\[
\hat{r}_1 = \arg \min_{0 \leq r_1 \leq m} VN_1(r_1), \text{ and}
\]

\[
\hat{r}_2 = \arg \min_{0 \leq r_2 \leq m} VN_2(r_2)
\]

where $f(r)$, $C_T$ and $C'_T$ are elements of penalty function defined in detail below.

Alternatively, we can simultaneously determine both $r_1$ and $r_2$ by minimizing

\[
V N(r_1, r_2) = -\sqrt{T} \sum_{i=1}^{r_1} \hat{\lambda}_i - \sum_{i=r_1+1}^{r_2} \hat{\lambda}_i + f(r_1) \frac{C_T}{T} + f(r_2) \frac{C'_T}{T^2},
\]

11
or

$$(\hat{r}_1, \hat{r}_2) = \arg \min_{0 \leq r_1, r_2 \leq m} VN(r_1, r_2).$$

The main theoretical result is provided in the following proposition.

**Proposition 1** (i) Suppose Assumption 1 holds, $f(r)$ is an increasing function of $r$, $C_T, C_T' \to \infty$,

$C_T/T, C_T'/T, \to 0$, then the paired procedure using $VN_1(r_1)$ and $VN_2(r_2)$ yields,

$$\lim_{T \to \infty} P(\hat{r}_1 = r_1, \hat{r}_2 = r_2) = 1.$$  

(ii) Suppose Assumption 1 holds, $f(r)$ is an increasing function of $r$, $C_T/T, C_T'/T \to \infty$, $C_T/T, C_T'/T, \to 0$, then the joint procedure using $VN(r_1, r_2)$ yields,

$$\lim_{T \to \infty} P(\hat{r}_1 = r_1, \hat{r}_2 = r_2) = 1.$$  

Remarks:

(a) The above proposition implies that both of the two cotrending rank selection procedures are consistent in selecting cotrending rank without specifying a parametric model.

(b) The criterion function $VN_1(r_1)$ in the paired procedure, can be used to select cointegrating rank in a system with only a stochastic trend. It nests the criterion function considered in Harris and Poskitt (2004) as a special case. In contrast, the information criteria for selecting cointegrating rank in Cheng and Phillips (2009) are based on the eigenstructure of reduced rank regression model.

(c) A popular choice of $C_T$ and $C_T'$, in the literature of information criteria includes $\ln(T), \ln(\ln(T))$, or 2, which respectively leads to BIC, HQ, and AIC. The proposition implies that our cotrending rank selection procedure yields consistency with a BIC and HQ type penalty but not with a AIC type penalty.
In the simulations, we also consider $C'_T = \ln(T) \times \ln(T)$ which yields consistency.

(d) For consistency, $f(r)$ can be any increasing function of $r$. In this paper, we consider $f_{HP}(r) = 2r(2m + r - 1)$ from Harris and Poskitt (2004) and $f_{CP}(r) = r(2m + (r + 1)/2)$ discussed in the footnote 1 of Cheng and Phillips (2009).

(e) Provided $C_T = C'_T$, the pair of selected cotrending ranks obtained from the paired procedure satisfies $\hat{r}_1 \leq \hat{r}_2$.

(f) In the absence of deterministic trend with $m = 1$, $r_1 = 0$ corresponds to a unit root case and $r_2 = 1$ corresponds to the stationary case.

4 Experimental evidence

4.1 Stochastic trend and cointegration

The selection procedures developed in the previous section are based on asymptotic theory. It is of interest to examine their finite sample properties by means of Monte Carlo analysis. For this purpose, this section reports the results under different settings of the true cotrending ranks, and of various penalty terms.

As noted in the previous section, our procedures are more general than that of Harris and Poskitt (2004) and Cheng and Phillips (2009), and the criterion function $VN_1(r_1)$ in the paired procedure can also determine the cointegrating rank in the absence of nonlinear deterministic trend. In this subsection, the finite sample performance of $VN_1(r_1)$ is evaluated and compared to the performance of cointegrating rank selection procedures proposed by Harris and Poskitt (2004) and Cheng and Phillips (2009).

We follow Cheng and Phillips (2009) and generate a bivariate time series $y_t = (y_{1t}, y_{2t})'$ using

$$\Delta y_t = \alpha \beta' y_{t-1} + u_t, \quad t = 1, \cdots, T,$$
where \( u_t \) follows a \( VAR(1) \) process with \( VAR \) coefficient \( 0.4 \times I_2 \) with mutually independent standard normal error term. By setting \( \alpha \beta' = 0 \),

\[
\alpha \beta' = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix} \begin{pmatrix} -1 & 1 \end{pmatrix},
\]

and

\[
\alpha \beta' = \begin{pmatrix} -0.5 & 0.1 \\ 0.2 & -0.15 \end{pmatrix},
\]

we generate a multivariate system with the true cointegrating rank being 0, 1 and 2, respectively. We generate the data for two sample sizes \( T = 100 \) and \( T = 400 \). To eliminate the effect of the initial values \( y_0 = 0 \) and \( u_0 = 0 \), the first 100 observations are discarded. We evaluate the performance of the procedures based on the frequencies of selecting the true cointegrating rank in 20,000 replications.

For the reduced rank regression procedure of Cheng and Phillips (2009), we employ the AIC, BIC and HQ criteria and denote them by RRR-AIC, RRR-BIC and RRR-HQ, respectively. For \( VN_1(r_1) \), we use \( f_{HP}(r) = 2r(2m - r + 1) \) as in Harris and Poskitt (2004), and \( C_T = 2, \log(T), \) and \( \log(\log(T)) \), and denote corresponding criteria by \( VN-AIC, VN-BIC, \) and \( VN-HQ, \) respectively. The consistent selection criterion of Harris and Poskitt (2004) (\( \Gamma_{C,T} \) in their notation) is identical to our \( VN-BIC \). It should also be noted that both RRR-AIC and VN-AIC are inconsistent in selecting true cointegrating rank.

Table 1 reports the performance of the cointegrating rank selection procedures based on six criteria with the frequencies of correctly selecting true rank denoted in bold type. Among the procedures based on the von Neumann criterion, when sample size is large, VN-BIC outperforms other procedures. The good performance of BIC-based criterion is consistent with the similar finding by Cheng and Phillips (2009) who compare the performance among reduced rank regression-based procedures. On the whole, the performance of the procedure based on the von Neumann criterion is comparable to that of Cheng.
and Phillips (2009) in selecting cointegrating rank.\footnote{While not reported, similar results are obtained for the case of different simulation design using other serial correlation structure for $u_t$.}

### 4.2 Stochastic and deterministic trends and cotrending

In this section, we conduct a Monte Carlo experiment to investigate the finite sample performance of our proposed cotrending rank selection procedure. We consider a three-dimensional vector series $y_t = (y_{1t}, y_{2t}, y_{3t})'$ with different combinations of cotrending and weak cotrending ranks. First, we generate using

$$
y_{1t} = \rho_1 y_{1t-1} + \varepsilon_{1t},
$$
$$
y_{2t} = \rho_2 y_{2t-1} + \varepsilon_{2t},
$$
$$
y_{3t} = \begin{cases} 
  c + \mu_0 t & \text{if } t \leq \tau T \\
  c + (\mu_0 - \mu_1)\tau T + \mu_1 t & \text{if } t > \tau T
\end{cases},
$$

with $(\varepsilon_{1t}, \varepsilon_{2t})' = iidN(0, \Sigma_\varepsilon)$ where

$$
\Sigma_\varepsilon = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}.
$$

By using different combinations of $\rho_i \in \{0.5, 1.0\}$ for $i = 1, 2$, we control cotrending and weak cotrending ranks. We generate the data with $(r_1, r_2) = (0,2), (1,1)$, and $(2,0)$. The parameters for the kinked trend function are set to $c = 0.5$, $\mu_0 = 2$, $\tau = 0.5$, and $\mu_1 = 0.5$.

Second, we generate the data using

$$
y_{1t} = \rho_1 y_{1t-1} + \varepsilon_{1t},
$$
$$
y_{2t} = \rho_2 y_{2t-1} + \varepsilon_{2t},
$$
$$
y_{3t} = \rho_3 y_{3t-1} + \varepsilon_{3t},
$$
with \((\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t})' = iidN(0, \Sigma_\varepsilon)\) where

\[
\Sigma_\varepsilon = \begin{bmatrix}
1 & 0.5 & 0.5 \\
0.5 & 1 & 0.5 \\
0.5 & 0.5 & 1
\end{bmatrix}.
\]

We use different combinations of \(\rho_i \in \{0.5, 1.0\}\) to generate the data with \((r_1, r_2) = (0, 3), (1, 2), (2, 1),\) and \((0, 3)\).

Third, we consider the cases of two and three deterministic trends using

\[
y_{1t} = c + \mu_0 t \text{ or } y_{1t} = \rho_1 y_{1t-1} + \varepsilon_{1t},
y_{2t} = \begin{cases}
    c + \mu_0 t & \text{if } t \leq \tau_1 T \\
    c + (\mu_0 - \mu_1)\tau_1 T + \mu_1 t & \text{if } t > \tau_1 T
\end{cases},
y_{3t} = \begin{cases}
    c + \mu_0 t & \text{if } t \leq \tau_2 T \\
    c + (\mu_0 - \mu_1)\tau_2 T + \mu_1 t & \text{if } t > \tau_2 T
\end{cases},
\]

with \(\varepsilon_{1t} = iidN(0, 1), \rho_1 \in \{0.5, 1.0\}, c = 0.5, \mu_0 = 2, \tau_1 = 0.5, \tau_2 = 1/3\) and \(\mu_1 = 0.5\). This generates with data with \((r_1, r_2) = (0, 0), (1, 0),\) and \((0, 1)\).

We employ two paired cotrending rank selection procedures and two joint cotrending rank selection procedures. For the paired procedures, we use \(C_T = C'_T = \log(T)\) and denote corresponding procedure by ‘paired VN-BIC’. In addition, we also consider the case with a stronger penalty for \(VN_2(r_2)\) with \(C'_T = \log(T) \times \log(T)\) along with BIC type penalty for \(VN_1(r_1)\) with \(C_T = \log(T)\), and denote the procedure by ‘paired VN-BIC2’. For the joint selection procedures, we consider \(VN(r_1, r_2)\) with the penalty \(C_T = C'_T = \sqrt{T} \log(T)\), and denote the procedure by ‘joint VN-BIC’. A joint selection procedure with \(C_T = \sqrt{T} \log(T)\) and \(C'_T = \sqrt{T} \log(\log(T))\) is denoted by ‘joint VN-BIC2’. Two types of \(f(r)\) are considered, namely, \(f_{HP}(r) = 2r(2m + r - 1)\) and \(f_{CP}(r) = r(2m + (r + 1)/2)\). Theoretically, all the procedures are consistent in selecting cotrending rank and weak cotrending rank.

Tables 2 through 5 report the frequencies of selecting cotrending and weak cotrending rank using various data generating processes. We investigate the performance of four procedures for sample sizes
$T = 100$ and $T = 400$ in 1,000 replications.\(^3\) The data generating process is described by the vector $(r_1, r_2 - r_1, m - r_2)$ of which each element denotes the number of cotrending rank, the difference between weak cotrending rank and cotrending rank, and the number of common deterministic trends, respectively. Both the true data generating process and the selection frequencies of the true model are shown in bold type. First, overall, performance of joint procedures seems better than that of paired procedure. Second, in both cases, increasing sample size leads to higher frequency of selecting true pairs $r_1$ and $r_2$. Finally, for some DGPs such as $(0,1,2)$ or $(0,0,3)$, relative performance among paired procedures highly depends on the choice of penalty terms.

### 4.3 Sensitivity to the choice of scale parameter

In this section, we study the sensitivity of the performance of our procedures to the choice of scale parameters in the logistic transition function. We generate the artificial data using

\[
y_{1t} = y_{1t-1} + \varepsilon_{1t}; \\
y_{2t} = c_0 + \mu_0 t; \\
y_{3t} = (c_0 + \mu_0 t)G(\gamma, \tau T) + (c_1 + \mu_1 t)(1 - G(\gamma, \tau T))
\]

where $G(\gamma, \tau T)$ is a logistic transition function defined in section 2 and $\varepsilon_{1t} = iidN(0, 1)$, $c = 0.5$, $\mu_0 = 2$, $\tau = 0.5$, and $\mu_1 = 0.5$. This model gives $(r_1, r_2) = (0, 1)$. As is noted in the previous sections, the scale parameter $\gamma$ controls the speed of transition. As $\gamma$ approaches infinity, the logistic function collapses to a index function $I(t > \tau T)$. On the other hand, as $\gamma$ goes to zero, the smooth transition trend model approaches to a linear trend. In this scenario, we can always find the linear combination that eliminates the trend function. Therefore, we expect that, when $\gamma$ becomes smaller, it would be difficult for our

\(^3\)We also run simulation for other cases with AIC type and HQ type penalties. When sample size is large, however, they are outperformed by the procedures with BIC type penalties.
procedure to identify the nonlinear deterministic trend from the linear trend.

Table 6 and Table 7 present the simulation results with $f_{HP}(r) = 2r(2m + r - 1)$ and $f_{CP}(r) = r(2m + (r + 1)/2)$ given different choices of the scale parameter $\gamma \in \{0.001, 0.05, 0.1\}$. The results are consistent with our prediction in direction.

5 Application

The simulation results in the previous section show that our procedures perform well in various experimental set-ups. In this section, we apply our procedures to the Japanese money demand function to investigate the cotrending relations among money demand, income and interest rate. A seasonally adjusted quarterly series of real GDP, two definitions of monetary aggregates, $M_1$ and $M_2 + CD$, the call rate are plotted in Figures 2 to 5. The figures show the possibility of kinked deterministic trends in these variables.

We follow Bae, Kakkar and Ogaki (2006) and consider following three different specifications of money demand functions,

Model 1 : $\log \left( \frac{M_t}{P_t} \right) = \beta_0 + \beta_1 \log(y_t) + \beta_1 i_t + \varepsilon_t$,

Model 2 : $\log \left( \frac{M_t}{P_t} \right) = \beta_0 + \beta_1 \log(y_t) + \beta_1 \log(i_t) + \varepsilon_t$, and

Model 3 : $\log \left( \frac{M_t}{P_t} \right) = \beta_0 + \beta_1 \log(y_t) + \beta_1 \log \left( \frac{i_t}{1 + i_t} \right) + \varepsilon_t$,

where $M_t$ is the money demand, $P_t$ is the aggregate price level, $y_t$ is real GDP and $i_t$ is the nominal interest rate.

We apply both paired and joint cotrending rank selection procedures to the vectors $(\log(M_t/P_t), \log(y_t), i_t)$, $(\log(M_t/P_t), \log(y_t), \log(i_t))$, and $(\log(M_t/P_t), \log(y_t), \log(i_t/(1 + i_t))$ using the sample spanning from 1958:Q1 to 2007:Q4 in case $M_1$ is used as a proxy for $M_t$, and from 1967:Q1 to 2007:Q4 when M2+CD
is used. Table 8 reports the empirical results for all three different specifications of the functional form for interest elasticity of money demand. The results are somewhat mixed depending on the choice of the penalty of the criteria and the model of money demand.

Overall, however, we have more cases of choosing cotrending with \( r_1 = 1 \) when Models 2 and 3 are used along with the penalty function \( f_{CP}(r) = r(2m + (r + 1)/2) \). Bae, Kakkar and Ogaki (2006) argue that the nonlinear functional forms (Models 2 and 3) are more appropriate for the Japanese long-run money demand. In that case, we have a stronger result in supporting the presence of cotrending relationship among these variables.

6 Conclusion

This paper has proposed a model-free cotrending rank selection procedure when both stochastic and nonlinear deterministic trends are present in a multivariate system. The procedure selects two types of cotrending rank by minimizing a new criterion based on the generalized von Neumann ratio. Our approach is invariant to linear transformation of data, robust to misspecification of the model, and consistent under very general conditions. Monte Carlo experiments have suggested good finite sample performance of the proposed procedure. Empirical applications to the money demand function in Japan have also suggested the usefulness of our procedure in detecting the cotrending relationships when there exists nonlinear deterministic trends in data.
Appendix

Proof of Lemma 1:

We want to show that \( \hat{\lambda}_1, \cdots, \hat{\lambda}_{r_1} \) is \( O_p(1) \) but is not \( o_p(1) \), \( \hat{\lambda}_{r_1+1}, \cdots, \hat{\lambda}_{r_2} \) is \( O_p(T^{-1}) \) but is not \( o_p(T^{-1}) \), and \( \hat{\lambda}_{r_2+1}, \cdots, \hat{\lambda}_m \) is \( O_p(T^{-2}) \) but is not \( o_p(T^{-2}) \) if all the eigenvalues of \( S_{11}^{-1}S_{00} \) are arranged in a descending order. We employ the data matrix notation, \( Y' = [y_1, \cdots, y_T], D' = [d_1, \cdots, d_T] \) and \( S' = [s_1, \cdots, s_T] \).

We have constructed an orthogonal full rank matrix \([B_\perp B_2 B_1]\) in Assumption 1 and further define

\[
M_{11} = B'S_{11}B, \quad \text{and} \quad M_{00} = B'S_{00}B
\]

Due to the orthogonality of the matrix \([B_\perp B_2 B_1]\), the eigenvalues of \( S_{11}^{-1}S_{00} \) arise as the same solutions to

\[
\det(\lambda M_{11} - M_{00}) = 0.
\]

Our proof can be established in the following two steps.

Step 1:

We assume \( G = \lim_{T \to \infty} T^{-3} \sum_{t=1}^{T} d_t d_t' \) exists and \( T^{-3} \sum_{t=1}^{T} d_t d_t' - G \) is \( O(T^{-1/2}) \). The eigenvalues of \( T^2 M_{11}^{-1} M_{00} \) are equivalent to the eigenvalues \( \lambda' s \) that solve

\[
\det(\lambda T^{-2} M_{11} - M_{00}) = 0
\]

For the matrix \( T^{-2} M_{11} \), the only block matrix that is not equal to zero is \( B_\perp' Y' Y B_\perp \), which converge to \( B_\perp' GB_\perp \) under Assumption 1. Because the eigenvalues are continuous functions of the matrix,

\[
p \lim_{T \to \infty} \lambda_i(T^2 M_{11}^{-1} M_{00}) = \lambda_i(p \lim_{T \to \infty} T^2 M_{11}^{-1} M_{00}).
\]
It can be easily shown that $M_{00}$ is $O_p(1)$ but is not $o_p(1)$. Therefore, for $i = r_2 + 1, \cdots, m$, we are led to

$$\lambda_i(T^2M_{11}^{-1}M_{00}) = O_p(1) \text{ but is not } o_p(1).$$

This leads to the result that $T^2\lambda_i$ is $O_p(1)$ but not $o_p(1)$ for $i = r_2 + 1, \cdots, m$.

Step 2:

Let $D_T = \text{diag}[I_{m-r_2}, T^{1/2}I_{r_2}]$, the roots of

$$\det(\lambda T^{-2}M_{11} - M_{00}) = 0$$

are equivalent to

$$\det(D_T [\lambda T^{-2}M_{11} - M_{00}] D_T) = 0 \quad (2)$$

The matrix $\lambda T^{-2}M_{11}$ can be rewritten as

$$\begin{pmatrix}
\lambda T^{-3}B'_2Y'YB_\perp & \lambda T^{-3}B'_2Y'Y[ B_2 \quad B_1 ] \\
\lambda T^{-3}[ B'_2 \quad B_1 ]Y'YB_\perp & \lambda T^{-3}[ B'_2 \quad B_1 ]Y'Y[ B_2 \quad B_1 ]
\end{pmatrix},$$

and we denote

$$Y_a = \lambda T^{-3}B'_2Y'YB_\perp - B'_2\Delta Y'\Delta YB'_2,$$

$$Y_b = \lambda T^{-2} \begin{pmatrix}
B'_2Y'YB_2 & B'_2Y'YB_1 \\
B'_1Y'YB_2 & B'_1Y'YB_1
\end{pmatrix} - \begin{pmatrix}
TB'_2\Delta Y'\Delta YB_2 & T^{1/2}B'_2\Delta Y'\Delta YB_1 \\
T^{1/2}B'_2\Delta Y'\Delta YB_2 & TB'_2\Delta Y'\Delta YB_1
\end{pmatrix},$$

and

$$Y_c = \lambda T^{-5/2} \begin{pmatrix}
B'_2Y'YB_\perp \\
B'_1Y'YB_\perp
\end{pmatrix} - T^{1/2} \begin{pmatrix}
B'_2\Delta Y'\Delta YB_\perp \\
B'_1\Delta Y'\Delta YB_\perp
\end{pmatrix}.$$
Then equation (2) is rewritten as

$$\det(Y_a) \det[Y_b - Y_c' Y_a^{-1} Y_c] = 0$$

(3)

The first determinant can on the LHS of (3) cannot be equal to zero, implying the second determinant must be zero. Concerning the first part of $Y_b$, only its first $r_2 \times r_2$ diagonal block is nonzero, and the second part of $Y_b$ and $Y_c' Y_a^{-1} Y_c$ is $O_p(T)$ but is not $o_p(T)$. Hence, we are led to

$$\det(\lambda_i T^{-2} B_1' Y' Y B_1 - O_p(T)) = 0$$

for $i = r_1 + 1, \cdots, r_2$. While we let $T$ goes to infinity and the solutions $\lambda_i$ solves the above equation satisfies

$$\lambda_i (T^2 M_{11}^{-1} M_{00}) = O_p(T) \text{ but is not } o_p(T) \text{ for } i = r_1 + 1, \cdots, r_2.$$

Therefore, one can conclude that $\hat{\lambda}_i$ is $O_p(T^{-1})$ but is not $o_p(T^{-1})$ for $i = r_1 + 1, \cdots, r_2$. Analogously, one can show that $\hat{\lambda}_i$ is $O_p(1)$ but is not $o_p(1)$ for $i = 1, \cdots, r_1$.

**Proof of Proposition 2.**

(i) Let $r_1$ be the true cotrending rank, which is estimated by minimization of $VN_1(r_1)$ for $0 \leq r_1 \leq m$.

To check the consistency of this estimator, we need to show $VN(r'_1) > VN(r_1)$ if $r'_1$ is not equal to the true cotrending rank $r_1$.

When $r'_1 < r_1$,

$$VN_1(r'_1) - VN_1(r_1) = \sum_{i=r_1+1}^{r_1} \hat{\lambda}_i + (f(r'_1) - f(r_1))C_T T^{-1}$$
In order to consistently select $r_1$ with probability 1 as $T \to \infty$, we need

$$\sum_{i=r_1' + 1}^{r_1} \lambda_i + (f(r_1') - f(r_1))C_T T^{-1} > 0, \text{ as } T \to \infty.$$ 

From Proposition 1, we know the first term is a positive number that is bounded away from zero and the second term is a negative number of order $O(C_T T^{-1})$. As long as $C_T T^{-1} \to 0$ as $T \to \infty$, the above inequality holds and we are led to the conclusion that $V N_1(r_1') > V N_1(r_1)$ when $r_1' < r_1$.

When $r_1' > r_1$,

$$VN_1(r_1') - VN_1(r_1) = -\sum_{i=r_1 + 1}^{r_1'} \lambda_i + (f(r_1') - f(r_1))C_T T^{-1}.$$

From Proposition 1, we know that $\lambda_i$ is $O_p(T^{-1})$ but is not $o_p(T^{-1})$ for $i = r_1 + 1, \cdots, r_1$. By multiplying both sides by $T$, we have

$$T\left(VN_1(r_1') - VN_1(r_1)\right) = -T \sum_{i=r_1 + 1}^{r_1'} \lambda_i + (f(r_1') - f(r_1))C_T.$$

As long as $C_T \to \infty$ as $T \to \infty$, the second term on the right hand side dominates, which leads to $VN_1(r_1') > VN_1(r_1)$ when $r_1' > r_1$. Thus the consistency of $VN_1(r_1)$ in selecting true cotrending rank is established. Analogously, one can establish the consistency of the estimator of the true weak cotrending rank by $VN_2(r_2)$.

(ii) To show the consistency of the joint selection procedure, consider all the possible cases as follows.

Case 1: $r_1' < r_1$,

We have

$$VN(r_1', r_2') - VN(r_1, r_2) = \sqrt{T} \sum_{i=r_1' + 1}^{r_1} \lambda_i + O_p(\frac{C_T}{T}),$$

where $\lambda_i$ for $i = r_1' + 1, \cdots, r_1$ is $O_p(1)$ but is not $o_p(1)$.

From Proposition 1 and Lemma 1, the first term dominates, which leads to $VN(r_1', r_2') > VN(r_1, r_2)$. 

23
when \( r'_1 < r_1 \).

Case 2: \( r'_1 > r_1 \).

\[
V N(r'_1, r'_2) - V N(r_1, r_2) = -\sqrt{T} \sum_{i=r_1+1}^{r'_1} \tilde{\lambda}_i + (f(r'_1) - f(r_1))\frac{C_T}{T} + O_p\left(\frac{C_T}{T^2}\right),
\]

where \( \tilde{\lambda}_i \) is \( O_p(T^{-1}) \) for \( i = r_1 + 1, \cdots, m \).

The dominant term in the above equation is \( (f(r'_1) - f(r_1))\frac{C_T}{T} \) provided that \( \frac{C_T}{\sqrt{T}} \to \infty \), the inequality \( V N(r'_1, r'_2) > V N(r_1, r_2) \) holds in this case.

Case 3: \( r'_1 = r_1 \).

When \( r'_2 > r_2 \),

\[
V N(r'_1, r'_2) - V N(r_1, r_2) = -\sqrt{T} \sum_{i=r_2+1}^{r'_2} \tilde{\lambda}_i + (f(r'_2) - f(r_1))\frac{C'_T}{T^2},
\]

where \( \tilde{\lambda}_i \) is \( O_p(T^{-2}) \) for \( r'_2 + 1, \cdots, m \).

Then, we have

\[
T^2 \left( V N(r'_1, r'_2) - V N(r_1, r_2) \right) = -\sqrt{T} \sum_{i=r_2+1}^{r'_2} T^2 \tilde{\lambda}_i + (f(r'_2) - f(r_2))C'_T.
\]

Provided that \( \frac{C'_T}{\sqrt{T}} \to \infty \), the dominant term is \( (f(r'_2) - f(r_2))C'_T \), which is greater than zero. Hence \( V N(r'_1, r'_2) > V N(r_1, r_2) \) in this case.

When \( r'_2 < r_2 \),

\[
V N(r'_1, r'_2) - V N(r_1, r_2) = \sqrt{T} \sum_{i=r'_2+1}^{r'_2} \tilde{\lambda}_i + (f(r'_2) - f(r_2))\frac{C'_T}{T^2}.
\]

The first term on the right hand side is \( O_p(T^{-3/2}) \) but is not \( o_p(T^{-3/2}) \), dominate the second term,
provided that $\frac{c'_T}{T} \to 0$. Hence $VN(r'_1, r'_2) > VN(r_1, r_2)$ in this case.

Combining the conditions on $C_T$ and $C'_T$, for all the preceding cases, it follows that the joint selection procedure will lead to consistent estimation of the cotrending and weak cotrending ranks.
References


Table 1. Two dimensional cointegrating rank selection

<table>
<thead>
<tr>
<th></th>
<th>$T = 100$</th>
<th></th>
<th>$T = 400$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r_1 = 0$</td>
<td>$r_1 = 1$</td>
<td>$r_1 = 2$</td>
</tr>
<tr>
<td>RRR-AIC</td>
<td>0.49</td>
<td>0.39</td>
<td>0.12</td>
</tr>
<tr>
<td>RRR-BIC</td>
<td>0.88</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>RRR-HQ</td>
<td>0.35</td>
<td>0.47</td>
<td>0.19</td>
</tr>
<tr>
<td>VN-AIC</td>
<td>0.98</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>VN-BIC</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>VN-HQ</td>
<td>0.93</td>
<td>0.07</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$r_1 = 0$</th>
<th>$r_1 = 1$</th>
<th>$r_1 = 2$</th>
<th>$r_1 = 0$</th>
<th>$r_1 = 1$</th>
<th>$r_1 = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRR-AIC</td>
<td>0.00</td>
<td>0.78</td>
<td>0.22</td>
<td>0.00</td>
<td>0.76</td>
<td>0.24</td>
</tr>
<tr>
<td>RRR-BIC</td>
<td>0.00</td>
<td>0.94</td>
<td>0.06</td>
<td>0.00</td>
<td>0.96</td>
<td>0.04</td>
</tr>
<tr>
<td>RRR-HQ</td>
<td>0.00</td>
<td>0.71</td>
<td>0.29</td>
<td>0.00</td>
<td>0.73</td>
<td>0.27</td>
</tr>
<tr>
<td>VN-AIC</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>VN-BIC</td>
<td>0.05</td>
<td>0.95</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>VN-HQ</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$r_1 = 0$</th>
<th>$r_1 = 1$</th>
<th>$r_1 = 2$</th>
<th>$r_1 = 0$</th>
<th>$r_1 = 1$</th>
<th>$r_1 = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRR-AIC</td>
<td>0.00</td>
<td>0.25</td>
<td>0.75</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>RRR-BIC</td>
<td>0.05</td>
<td>0.73</td>
<td>0.22</td>
<td>0.00</td>
<td>0.02</td>
<td>0.98</td>
</tr>
<tr>
<td>RRR-HQ</td>
<td>0.00</td>
<td>0.14</td>
<td>0.86</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>VN-AIC</td>
<td>0.00</td>
<td>0.39</td>
<td>0.61</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>VN-BIC</td>
<td>0.00</td>
<td>0.95</td>
<td>0.05</td>
<td>0.00</td>
<td>0.06</td>
<td>0.94</td>
</tr>
<tr>
<td>VN-HQ</td>
<td>0.00</td>
<td>0.19</td>
<td>0.81</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note: The frequencies of selecting each cointegrating rank are reported.
Table 2. Simulation results on cotrending rank selection with $f_{HP}(r) = 2r(2m - r + 1)$

<table>
<thead>
<tr>
<th>$T=100$</th>
<th>(1,1)</th>
<th>(1,2,0)</th>
<th>(2,1,0)</th>
<th>(2,0,1)</th>
<th>(0,2,1)</th>
<th>(1,0,2)</th>
<th>(0,1,2)</th>
<th>(3,0,0)</th>
<th>(0,3,0)</th>
<th>(0,0,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired VN-BIC</td>
<td>0.97</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.97</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.94</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.98</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$T=100$</td>
<td>(1,1)</td>
<td>(1,2,0)</td>
<td>(2,1,0)</td>
<td>(2,0,1)</td>
<td>(0,2,1)</td>
<td>(1,0,2)</td>
<td>(0,1,2)</td>
<td>(3,0,0)</td>
<td>(0,3,0)</td>
<td>(0,0,3)</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.93</td>
<td>0.07</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.93</td>
<td>0.00</td>
<td>0.00</td>
<td>0.07</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.46</td>
<td>0.49</td>
<td>0.02</td>
<td>0.03</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.04</td>
<td>0.92</td>
<td>0.04</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$T=100$</td>
<td>(1,1)</td>
<td>(1,2,0)</td>
<td>(2,1,0)</td>
<td>(2,0,1)</td>
<td>(0,2,1)</td>
<td>(1,0,2)</td>
<td>(0,1,2)</td>
<td>(3,0,0)</td>
<td>(0,3,0)</td>
<td>(0,0,3)</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.07</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$T=100$</td>
<td>(1,1)</td>
<td>(1,2,0)</td>
<td>(2,1,0)</td>
<td>(2,0,1)</td>
<td>(0,2,1)</td>
<td>(1,0,2)</td>
<td>(0,1,2)</td>
<td>(3,0,0)</td>
<td>(0,3,0)</td>
<td>(0,0,3)</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.74</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.92</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.47</td>
<td>0.49</td>
<td>0.02</td>
<td>0.03</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.04</td>
<td>0.92</td>
<td>0.04</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

30
<table>
<thead>
<tr>
<th>$T=400$</th>
<th>(1,1)</th>
<th>(1,2)</th>
<th>(2,1)</th>
<th>(2,0)</th>
<th>(0,2)</th>
<th>(1,2)</th>
<th>(3,0)</th>
<th>(3,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired VN-BIC</td>
<td>0.99</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.99</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.98</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.99</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T=400$</th>
<th>(1,1)</th>
<th>(1,2)</th>
<th>(2,1)</th>
<th>(2,0)</th>
<th>(0,2)</th>
<th>(1,2)</th>
<th>(3,0)</th>
<th>(3,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired VN-BIC</td>
<td>0.98</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.75</td>
<td>0.01</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.23</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.98</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 3. Simulation results on cotrending rank selection with $f_{HP}(r) = 2r(2m - r + 1)$
Table 4. Simulation results on cotrending rank selection with $f_{CP}(r) = r(2m + (r + 1)/2)$

<table>
<thead>
<tr>
<th>T=100</th>
<th>(1,1,1)</th>
<th>(1,2,0)</th>
<th>(2,1,0)</th>
<th>(2,0,1)</th>
<th>(0,2,1)</th>
<th>(1,0,2)</th>
<th>(0,1,2)</th>
<th>(3,0,0)</th>
<th>(0,3,0)</th>
<th>(0,0,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired VN-BIC</td>
<td>0.66</td>
<td>0.00</td>
<td>0.00</td>
<td>0.35</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.66</td>
<td>0.00</td>
<td>0.00</td>
<td>0.35</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.72</td>
<td>0.00</td>
<td>0.00</td>
<td>0.28</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.72</td>
<td>0.00</td>
<td>0.00</td>
<td>0.28</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>T=100</th>
<th>(1,1,1)</th>
<th>(1,2,0)</th>
<th>(2,1,0)</th>
<th>(2,0,1)</th>
<th>(0,2,1)</th>
<th>(1,0,2)</th>
<th>(0,1,2)</th>
<th>(3,0,0)</th>
<th>(0,3,0)</th>
<th>(0,0,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.50</td>
<td>0.48</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.50</td>
<td>0.48</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.08</td>
<td>0.50</td>
<td>0.32</td>
<td>0.09</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.58</td>
<td>0.40</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>T=100</th>
<th>(1,1,1)</th>
<th>(1,2,0)</th>
<th>(2,1,0)</th>
<th>(2,0,1)</th>
<th>(0,2,1)</th>
<th>(1,0,2)</th>
<th>(0,1,2)</th>
<th>(3,0,0)</th>
<th>(0,3,0)</th>
<th>(0,0,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.85</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.85</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.76</td>
<td>0.12</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.89</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>T=100</th>
<th>(1,1,1)</th>
<th>(1,2,0)</th>
<th>(2,1,0)</th>
<th>(2,0,1)</th>
<th>(0,2,1)</th>
<th>(1,0,2)</th>
<th>(0,1,2)</th>
<th>(3,0,0)</th>
<th>(0,3,0)</th>
<th>(0,0,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired VN-BIC</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

32
Table 5. Simulation results on cointrending rank selection with \( f_{CP}(r) = r(2m + (r + 1)/2) \)

<table>
<thead>
<tr>
<th>T=400</th>
<th>(1,1)</th>
<th>(1,2)</th>
<th>(2,1)</th>
<th>(2,0)</th>
<th>(2,1)</th>
<th>(0,2)</th>
<th>(1,2)</th>
<th>(0,3)</th>
<th>(0,3)</th>
<th>(0,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired VN-BIC</td>
<td>0.82</td>
<td>0.00</td>
<td>0.00</td>
<td>0.18</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.82</td>
<td>0.00</td>
<td>0.00</td>
<td>0.18</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.84</td>
<td>0.00</td>
<td>0.00</td>
<td>0.17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.84</td>
<td>0.00</td>
<td>0.00</td>
<td>0.16</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.71</td>
<td>0.29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.71</td>
<td>0.29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.04</td>
<td>0.69</td>
<td>0.24</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.73</td>
<td>0.27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.91</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.91</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.87</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.93</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

33
Table 6. Sensitivity to the choice of scale parameter with $f_{HP}(r) = 2r(2m - r + 1)$

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>(1,1,1)</th>
<th>(1,2,0)</th>
<th>(2,1,0)</th>
<th>(2,0,1)</th>
<th>(0,2,1)</th>
<th>(1,0,2)</th>
<th>(0,1,2)</th>
<th>(3,0,0)</th>
<th>(0,3,0)</th>
<th>(0,0,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma = 0.001$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.94</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.94</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.93</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>0.03</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.96</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\gamma = 0.05$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.94</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.92</td>
<td>0.00</td>
<td><strong>0.02</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td><strong>0.97</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.97</td>
<td>0.01</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\gamma = 0.1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.96</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.96</td>
<td>0.02</td>
<td><strong>0.98</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td><strong>0.98</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td><strong>0.98</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 7. Sensitivity to the choice of scale parameter with $f_{CP}(r) = r(2m + (r + 1)/2)$

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>(1,1,1)</th>
<th>(1,2,0)</th>
<th>(2,1,0)</th>
<th>(2,0,1)</th>
<th>(0,2,1)</th>
<th>(1,0,2)</th>
<th>(0,1,2)</th>
<th>(3,0,0)</th>
<th>(0,3,0)</th>
<th>(0,0,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma = 0.001$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.80</td>
<td>0.00</td>
<td>0.00</td>
<td>0.08</td>
<td>0.12</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.80</td>
<td>0.00</td>
<td>0.00</td>
<td>0.08</td>
<td>0.12</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.78</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.19</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.78</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.19</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\gamma = 0.05$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.61</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.40</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.61</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.40</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.55</td>
<td>0.45</td>
<td><strong>0.45</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.46</td>
<td>0.01</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\gamma = 0.1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.96</td>
<td>0.00</td>
<td><strong>0.00</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.96</td>
<td>0.02</td>
<td><strong>0.96</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td><strong>0.98</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.98</td>
<td>0.00</td>
<td><strong>0.98</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Table 8. Cotrending relationship among money, income, and interest rates

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2 + CD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Paired VN-BIC</td>
<td>(0,3,0)</td>
<td>(0,3,0)</td>
</tr>
<tr>
<td>Paired VN-BIC2</td>
<td>(0,2,1)</td>
<td>(0,2,1)</td>
</tr>
<tr>
<td>Joint VN-BIC</td>
<td>(0,1,2)</td>
<td>(0,1,2)</td>
</tr>
<tr>
<td>Joint VN-BIC2</td>
<td>(0,2,1)</td>
<td>(1,0,2)</td>
</tr>
</tbody>
</table>

Note: The first element in the parenthesis denotes cotrending rank, $r_1$, the second element denotes the difference between weak cotrending rank and cotrending rank, $r_2 - r_1$, the last element denotes the number of common deterministic trends, $m - r_2$. The first column for each model represents the results by using $f_{HP}(r)$ and the second column represents the results by using $f_{CP}(r)$. Both the results using M1 and M2+CD are reported.
Figure 1. Segmented linear trend and smooth transition trend
Figure 2. Real GDP

Figure 3. Monetary aggregate: M2 + CD
Figure 4. Monetary aggregate: M1

Figure 5. Call rate