DEPARTMENT OF ECONOMICS

Working Paper

Long-run forecasting in multicointegrated systems

Boriss Siliverstovs, Tom Engsted and Niels Haldrup

Working Paper No. 2002-15

ISSN 1396-2426

UNIVERSITY OF AARHUS • DENMARK

INSTITUT FOR ØKONOMI

AFDELING FOR NATIONALØKONOMI - AARHUS UNIVERSITET - BYGNING 350
8000 AARHUS C - ☎ 89 42 11 33 - TELEFAX 86 13 63 34

WORKING PAPER

Long-run forecasting in multicointegrated systems

Boriss Siliverstovs, Tom Engsted and Niels Haldrup

Working Paper No. 2002-15

DEPARTMENT OF ECONOMICS

SCHOOL OF ECONOMICS AND MANAGEMENT - UNIVERSITY OF AARHUS - BUILDING 350
8000 AARHUS C - DENMARK ☎ +45 89 42 11 33 - TELEFAX +45 86 13 63 34

Long-run forecasting in multicointegrated systems.*

Boriss Siliverstovs[§]
DIW Berlin

Tom Engsted[¶]
Department of Finance
The Aarhus School of Business

Niels Haldrup[‖]
Department of Economics
University of Aarhus

October 10, 2002

*For constructive criticism the authors would like to thank Hans Christian Kongsted, Philip Hans Franses, and Allan Würtz, as well as seminar participants at the University of Aarhus, University of Southampton, the 22nd International Symposium on Forecasting in Dublin, and "Common Features in Rio" conference in Rio de Janeiro.

Abstract

In this paper long-run forecasting of multicointegrating variables is investigated. Multicointegration typically occurs in dynamic systems involving both stock and flow variables whereby a common feature in the form of shared stochastic trends is present across different levels of multiple time series. Hence, the effect of imposing this "common feature" restriction on out-of-sample evaluation and forecasting accuracy of such variables is of interest. In particular, we compare the long-run forecasting performance of the multicointegrated variables between a model that correctly imposes the "common feature" restrictions and a (univariate) model that omits these multicointegrating restrictions completely. We employ different loss functions based on a range of mean square forecast error criteria, and the results indicate that different loss functions result in different ranking of models with respect to their infinite horizon forecasting performance. We consider loss functions using a standard trace mean square forecast error criterion (penalizing the forecast errors of flow variables only), and a loss function evaluating forecast errors of changes in both stock and flow variables. The latter loss function is based on the triangular representation of cointegrated systems and was initially suggested by Christoffersen and Diebold (1998). It penalizes deviations from long-run relations among the flow variables through cointegrating restrictions. We present a new loss function which further penalizes deviations in the long run relationship between the levels of stock and flow variables. It is derived from the triangular representation of multicointegrated systems. Using this criterion, system forecasts from a model incorporating multicointegration restrictions dominate forecasts from univariate models. The paper highlights the importance of carefully selecting loss functions in forecast evaluation of models involving stock and flow variables.

Keywords: Common Features, Multicointegration, Forecasting, VAR models.

JEL Classification Codes: C32, C53.

[§]DIW Berlin, Königin-Luise Straße 5, 14195 Berlin, Germany, `e-mail: bsiliverstovs@econ.au.dk`, (corresponding author)

[¶]The Aarhus School of Business, Department of Finance, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark, `e-mail: tom@asb.dk`

[‖]University of Aarhus, Department of Economics, Building 326, office 217, DK-8000 Aarhus C, Denmark, `e-mail: nhaldrup@econ.au.dk`

1 Introduction

Assessing the forecasting performance of econometric models is an important ingredient in evaluating such models. In multivariate models containing non-stationary variables, cointegration may be thought to play a key role in assessing forecasting ability, especially over long horizons, because cointegration captures the long-run comovement of the variables. Several studies have investigated the forecasting properties of cointegrated models. Engle and Yoo (1987) make a small Monte Carlo study where they compare mean-squared forecast errors from a VAR in levels, which *does not* impose cointegration, to forecasts from a correctly specified error-correction model (ECM), which *does* impose cointegration, and they find that longer-run forecasts from the ECM are more accurate. This result supports the above intuition that imposing cointegration gives better long-horizon forecasts for variables that are tied together in the long run. However, subsequent research has somewhat questioned and modified this – at first glance appealing – conclusion.

According to Christoffersen and Diebold (1998), the doubts on the usefulness of cointegrating restrictions on the long-run forecasts are related to the following conjecture. The improved predictive power of cointegrating systems comes from the fact that deviations from the cointegrating relations tend to be eliminated. Thus, these deviations contain useful information on the likely future evolution of the cointegrated system which can be exploited to produce superior forecasts when compared to those made from models that omit cointegrating restrictions. However, since the long-run forecast of the cointegrating term is always zero, this information is only likely to be effective when the focus lies on producing the short-run forecasts. Hence, at least from this point of view, the usefulness of imposing cointegrating relations for producing long-run superior forecasts can be questioned.

Clements and Hendry (1995) compare mean-squared error forecasts from a correctly specified ECM to forecasts from both an unrestricted VAR in levels and a misspecified VAR in first-differences (DVAR) which omits cointegrating restrictions present among the variables. They find that the forecasting superiority of the model that correctly imposes these cointegrating restrictions hinges crucially on whether the forecasts are for the levels of the variables, their first-differences, or the cointegrating relationship between the variables. They show that this difference in rankings for alternative yet isomorphic representations of the variables is due to the mean-squared forecast error (MSFE) criterion not being invariant to nonsingular, scale-preserving linear transformations of the model.[1] In particular, they show that the forecasts from the ECM model are not superior to those made from the DVAR model at all but the shortest forecast horizons when the first-differences of I(1) variables are forecasted.

Christoffersen and Diebold (1998) compare mean-squared error forecasts of the levels of I(1) variables

[1] Clements and Hendry (1993) suggest an alternative measure that is invariant to scale-preserving linear transformations of the data: the *generalized forecast error second moment* (GFESM) measure.

from a true cointegrated VAR to forecasts from correctly specified univariate representations, and they similarly find that imposing cointegration does not improve long-horizon forecast accuracy. Thus, it appears that the simple univariate forecasts are just as accurate as the multivariate forecasts when judged using the loss function based on the MSFE criterion. They argue that this apparent paradox is due to the fact that the standard MSFE criterion fails to value the long-run forecasts' hanging together correctly. Long-horizon forecasts from the cointegrated VAR always satisfy the cointegrating restrictions exactly, whereas the long-horizon forecasts from the univariate models do so only on average, but this distinction is ignored in the MSFE criterion. Christoffersen and Diebold suggest an alternative criterion that explicitly accounts for this feature. The criterion is based on the *triangular representation* of cointegrated systems (see Campbell and Shiller, 1987, and Phillips, 1991). The virtue of this criterion is that it assesses forecast accuracy in the conventional "small MSFE" sense, but at the same time it makes full use of the information in the cointegrating relationships amongst the variables. Using this new forecast criterion, they indeed find that at long horizons the forecasts from the cointegrated VAR are superior to the univariate forecasts. Christoffersen and Diebold (1998) demonstrate that the reason for Engle and Yoo's (1987) Monte Carlo experiment to turn out favorable to a model with cointegrating restrictions is not due to the fact that such long-run relations are imposed but rather that the *correct* number of unit roots is imposed.

The purpose of the present paper is twofold. First, we extend the analysis of Christoffersen and Diebold to the case where the variables under study not only obey cointegrating relationships, but also obey certain *multi*cointegrating restrictions. Multicointegration was originally defined by Granger and Lee (1989, 1991) and refers to the case where the underlying I(1) variables are cointegrated in the usual sense *and* where, in addition, the cumulated cointegration errors cointegrate with the original I(1) variables. Thus, essentially there are two levels of cointegration amongst the variables and hence a common feature in the form of a stochastic trend will exist at different levels of the multiple time series.

Multicointegration is a very convenient way of modeling the interactions between stock and flow variables. Granger and Lee consider the case where the two I(1) variables production, y_t, and sales, x_t, cointegrate, such that inventory investments, s_t, are stationary, $s_t \equiv y_t - \beta x_t \sim \mathrm{I}(0)$, but where the cumulation of inventory investment, $I_t \equiv \Sigma_{j=1}^{t} s_j$, i.e. the level of inventories (which is then an I(1) stock variable), in turn cointegrates with either y_t or x_t, or both of them. Another example, analyzed by Lee (1992) and Engsted and Haldrup (1999), is where y_t is new housing units started, x_t is new housing units completed, s_t is uncompleted starts, and hence I_t is housing units under construction. Leachman (1996), and Leachman and Francis (2000) provide examples of multicointegrated systems with government revenues and expenditures, and a country's export and import, respectively. Here the stock variable is defined as the government debt and the country's external debt, such that each variable is the cumulated series of past government and trade deficits, respectively. Yet another example is provided by Siliverstovs

(2001) who analyze consumption and income, and where cumulated savings (i.e. the cumulation of the cointegrating relationship between income and consumption) constitutes wealth, which cointegrates with consumption and income. In general, multicointegration captures the notion of *integral control* in dynamic systems, see, for example, Hendry and von Ungern Sternberg (1981).

We investigate how the presence of multicointegration affects long-run forecasting comparisons. In particular, we set up a model that contains both cointegrating and multicointegrating restrictions, and we examine how forecasts from this multicointegrated system compare to univariate forecasts. The comparison is done in terms of the ratio of the (trace) mean-squared forecast errors, but we follow Christoffersen and Diebold (1998) in using both a standard loss function and a loss function based on the triangular representation of the cointegrated system. For a model with multicointegrating restrictions the standard trace mean-squared forecast error criterion entails a loss function that penalizes forecast errors associated with the levels of flow variables whereas the loss function associated with the triangular representation penalizes forecast errors of changes in both the flow *and* the stock variables.

Secondly, we are concerned with the fact that, when the loss function of Christoffersen and Diebold (1998) is applied to the multicointegrated systems, it focuses exclusively on the maintenance of the cointegrating restrictions while ignoring multicointegrating restrictions present in the data. This corresponds to ignoring how the levels of both stock and flow variables are related. To this end, we propose a new loss function that is based on the triangular representation of the multicointegrating variables. This extended loss function explicitly acknowledges the maintenance of the multicointegrating restrictions in the data. We also investigate the implications of using this new loss function in assessing the forecast accuracy between the system and univariate forecasts.

Our most important results can be summarized as follows. First we find that the general result of Christoffersen and Diebold (1998) derived for a standard cointegration model carries over to multicointegrated models, that is, based on a standard MSFE criterion, long-horizon forecasts of the levels of I(1) (flow) variables from the multicointegrated system are found not to be superior to simple univariate forecasts. However, based on the triangular MSFE criterion (accounting for changes in both stocks and flows), the system forecasts are clearly superior to the univariate forecasts. This result demonstrates that as long as the comparison is between the standard MSFE loss function and the triangular MSFE loss function, multicointegration will have no influence on the conclusions drawn by Christoffersen and Diebold. Hence, if the loss function reflects changes in the flow variables, or changes in both the flow and stock variables, then there is really no new insights to be gained from multicointegration in terms of the forecasting properties. However, in stock-flow models one will typically prefer a loss function that also values forecast errors associated with the linkage between the levels of stock and flow variables. Our suggested loss function is doing just that. As a second important result, it is shown that the new loss function presented in this paper reflects increasing forecasting gains (for the forecast horizon tending to infinity) when mean squared forecast errors from a multicointegrated system are compared to univariate forecasts. These results illustrate the importance of carefully selecting loss functions for systems involving stock and flow variables.

Testing for multicointegration, and estimation of models with multicointegrating restrictions, are most naturally conducted within an I(2) cointegration framework, see Engsted, Gonzalo and Haldrup (1997), Haldrup (1998), Engsted and Johansen (1999), and Engsted and Haldrup (1999). However, since our primary interest is on the particular dynamic characteristics of multicointegration with respect to forecasting, we abstract from estimation issues and hence assume known parameters.

The rest of the paper is organized as follows. In Section 2 we set up the multicointegrated systems used in the subsequent analysis. Also, we derive the corresponding univariate representations of the system variables. Section 3 derives the expressions for system and univariate forecasts and the associated forecasting errors. In Section 4 we demonstrate the implications on model ranking using various loss functions. Section 5 illustrates our findings using a numerical example and the final section concludes.

2 Multivariate and univariate representations of the multicointegrating variables

In this section we define multicointegrated models and derive the corresponding univariate representations of the system variables. In order to ease the exposition we employ the simplest models with relevant multicointegrating restrictions. Our bivariate setup is motivated by the fact that all applications of multicointegration in the literature have been performed for systems of just two variables.

2.1 Multicointegrated system.

Consider the two I(1) variables, x_t and y_t, that obey a cointegrating relation

$$y_t - \lambda x_t \sim I(0), \tag{1}$$

such that the cumulated cointegration error

$$\sum_{j=1}^{t} (y_j - \lambda x_j) \sim I(1)$$

is an I(1) variable by construction[2]. We refer to the system as *multicointegrated* when there exists a stationary linear combination of the cumulated cointegrating error and the original variables, e.g.

$$\sum_{j=1}^{t} (y_j - \lambda x_j) - \alpha x_t \sim I(0). \tag{2}$$

As discussed in Granger and Lee (1989, 1991), the multicointegrating restrictions are likely to occur in stock-flow models, where both cointegrating relations have an appealing interpretation. The first cointegrating relation (1) is formed between the original flow variables, for example, production and sales, income and expenditures, export and import, etc. The second cointegrating relation (2) represents the relation between the cumulated past discrepancies between the flow variables, for instance: the stock of inventories, the stock of wealth, the stock of external debt, and all or some flow variables present in the system. It implies that the equilibrium path of the system is maintained not only through the flow variables alone, but also through additional forces tying together the stock and flow series and in so doing provides a second layer of equilibrium.

It is convenient to represent the system of multicointegrating variables in the triangular form

$$\begin{bmatrix} (1-L) & 0 \\ -\lambda(1-L)^{-1} - \alpha & (1-L)^{-1} \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}, \tag{3}$$

where L is the lag operator, $(1-L)$ is the difference operator, and $(1-L)^{-1}$ is the summation operator, such that when the latter operator is applied to an I(1) time series the resulting time series is I(2) by construction, i.e. $(1-L)^{-1}x_t = \sum_{j=1}^{t} x_j$. For simplicity, it is assumed that the disturbances are uncorrelated at all leads and lags, i.e. $E(e_{1t-j}e_{2t-i}) = 0, \forall\, j \neq i$ for $j = 0, \pm 1, \pm 2, ...$ and $i = 0, \pm 1, \pm 2, ...$, and the variances of the disturbances e_{1t} and e_{2t} are given by σ_1^2 and σ_2^2, respectively, for all t. Hence x_t is considered a strictly exogenous variable.

If we denote the generated I(2) variables by capital letters, i.e. $Y_t = \sum_{j=1}^{t} y_j$ and $X_t = \sum_{j=1}^{t} x_j$, then we are able to write the system above as

$$\begin{aligned} \Delta x_t &= e_{1t} \\ Y_t &= \lambda X_t + \alpha x_t + e_{2t}. \end{aligned}$$

Observe that it closely resembles the socalled polynomially cointegrating system where original I(2) variables cointegrate with their own first differences, see Rahbek, Kongsted, and Jørgensen (1999), and Banerjee, Cockerell, and Russell (2001) for examples. The only difference between multi- and polynomially cointegrated models is that in the former case the I(2) variables are generated from the original I(1) variables, whereas in the latter case I(2) variables are the original time series.

[2]No deterministic components are assumed in the series and hence, by construction, no trend, for instance, is generated in the cumulated series.

Below we provide two equivalent representations of the system in (3). The Vector Error-Correction model (VECM) can be represented as follows:

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} [y_{t-1} - \lambda x_{t-1}] + \begin{bmatrix} 0 \\ -1 \end{bmatrix} [Y_{t-1} - \lambda X_{t-1} - \alpha x_{t-1}] + \begin{bmatrix} e_{1t} \\ (\lambda + \alpha) e_{1t} + e_{2t} \end{bmatrix}.$$

As seen, the VECM explicitly incorporates both cointegration levels, see equations (1) and (2) , that are present in the multicointegrated system. Alternatively, the multicointegrated system (3) can be given the moving-average (MA) representation:

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = C(L) e_t = \begin{bmatrix} 1 & 0 \\ [\lambda + (1-L)\alpha] & (1-L)^2 \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}. \tag{4}$$

Granger and Lee (1991) argue that the necessary and sufficient condition for x_t and y_t to be multicointegrated is that the determinant of $C(L)$ should have a root $(1-L)^2$. This condition is clearly satisfied for our simple system.

2.2 Univariate representations.

In this section we derive the implied univariate representations for the I(1) variables x_t and y_t. Of course, for x_t the univariate representation is just

$$x_t = x_{t-1} + e_{1t}.$$

In deriving the implied univariate representation for y_t we follow Christoffersen and Diebold (1998) by matching the autocovariances of the process Δy_t. From the MA-representation of Δy_t we have

$$\begin{aligned} \Delta y_t &= [\lambda + (1-L)\alpha] e_{1t} + (1-L)^2 e_{2t}, \\ y_t &= y_{t-1} + z_t, \end{aligned} \tag{5}$$

where, as shown in the appendix, the process z_t corresponds to the MA(2) process

$$z_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}, \quad u_t \sim \mathsf{IID}\left(0, \sigma_u^2\right).$$

The coefficient θ_2 represents a root of the following fourth order polynomial

$$\begin{gathered} \theta_2^4 + (2-B)\theta_2^3 + \left(A^2 - 2B + 2\right)\theta_2^2 + (2-B)\theta_2 + 1 = 0, \\ \text{where } \; A = [-\alpha(\lambda+\alpha) q - 4], \; B = \left[(\lambda+\alpha)^2 + \alpha^2\right] q + 6, \text{ and } q = \tfrac{\sigma_1^2}{\sigma_2^2}. \end{gathered} \tag{6}$$

and the coefficient θ_1 and the variance term σ_u^2 can be found as follows:

$$\theta_1 = \frac{\theta_2}{(1+\theta_2)} A \quad \text{and} \quad \sigma_u^2 = \frac{\sigma_2^2}{\theta_2}. \tag{7}$$

Observe that the values of the MA coefficients θ_1 and θ_2 are chosen such that they satisfy the invertibility conditions for the MA(2) process z_t.

3 Long-run forecasting in multicointegrated systems.

In this section we derive expressions for forecasts of the levels of I(1) variables as well as the corresponding forecast errors both from the system and univariate representations.

3.1 System forecasts of I(1) variables.

The MA-representation of the multicointegrating variables (4) allows us to write the future values of the system variables in terms of x_t and future innovations e_{1t+h} and e_{2t+h}:

$$\begin{aligned} x_{t+h} &= x_t + \sum_{i=1}^{h} e_{1t+i}, \\ y_{t+h} &= \lambda x_t + \lambda \sum_{i=1}^{h} e_{1t+i} + \alpha e_{1t+h} + \Delta e_{2t+h}. \end{aligned} \tag{8}$$

Correspondingly, the h-steps ahead forecasts for the I(1) variables are given by

$$\begin{aligned} \widehat{x}_{t+h} &= x_t, \\ \widehat{y}_{t+h} &= \lambda x_t \end{aligned} \tag{9}$$

for **all** forecast horizons but $h = 1$. In the latter case we have

$$\begin{aligned} \widehat{x}_{t+1} &= x_t \\ \widehat{y}_{t+1} &= \lambda x_t - e_{2t} = \lambda x_t - [Y_t - \lambda X_t - \alpha x_t]. \end{aligned} \tag{10}$$

In particular, observe that the long-run forecasts from the multicointegrated system maintain the cointegrating relation exactly:

$$\widehat{y}_{t+h} = \lambda \widehat{x}_{t+h}, \text{ for } h > 1. \tag{11}$$

Continuing, the forecast errors are

$$\widehat{\varepsilon}_{x,t+h} = \sum_{i=1}^{h} e_{1t+i} \quad \forall h > 0, \tag{12}$$

$$\widehat{\varepsilon}_{y,t+h} = \begin{cases} \lambda e_{1t+1} + \alpha e_{1t+1} + e_{2t+1} = (\lambda + \alpha) e_{1t+1} + e_{2t+1} & \text{for } h = 1 \\ \lambda \sum_{i=1}^{h} e_{1t+i} + \alpha e_{1t+h} + \Delta e_{2t+h} & \text{for } h > 1. \end{cases} \tag{13}$$

Furthermore, note that the forecast errors and the original system as in (4) follow the same stochastic process, i.e.

$$\begin{bmatrix} \Delta\widehat{\varepsilon}_{x,t+h} \\ \Delta\widehat{\varepsilon}_{y,t+h} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda + \alpha(1-L) & (1-L)^2 \end{bmatrix} \begin{bmatrix} e_{1t+h} \\ e_{2t+h} \end{bmatrix}. \tag{14}$$

3.2 Univariate forecasts of I(1) variables.

Next, we turn to forecasting of the I(1) variables based on the correctly specified implied univariate representations. Future values of x_{t+h} are given in equation (8) above and for y_{t+h}

$$y_{t+h} = \begin{cases} y_t + z_{t+1} = y_t + u_{t+1} + \theta_1 u_t + \theta_2 u_{t-1}, & h = 1 \\ y_t + \sum_{i=1}^{h} z_{t+i} = y_t + u_{t+1} + \theta_1 u_t + \theta_2 u_{t-1} + u_{t+2} + \theta_1 u_{t+1} + \theta_2 u_t + \sum_{i=3}^{h} z_{t+i}, & h > 1. \end{cases}$$

The corresponding $h-$steps ahead forecasts for I(1) variables can now be derived as follows. The forecast for x_t is the same as the system forecast

$$\widetilde{x}_{t+h} = \widehat{x}_{t+h} = x_t,$$

whereas the forecast $\widetilde{y}_{t+h}$ is given by

$$\widetilde{y}_{t+h} = \begin{cases} y_t + \theta_1 u_t + \theta_2 u_{t-1}, & \text{for } h = 1 \\ y_t + \theta_1 u_t + \theta_2 u_{t-1} + \theta_2 u_t = y_t + (\theta_1 + \theta_2)\, u_t + \theta_2 u_{t-1}, & \text{for } h > 1. \end{cases}$$

The forecast error for x_{t+h} is given by

$$\widetilde{\varepsilon}_{x,t+h} = \widehat{\varepsilon}_{x,t+h} = \sum_{i=1}^{h} e_{1t+i}. \tag{15}$$

The corresponding forecast error $\widetilde{\varepsilon}_{y,t+h} = y_{t+h} - \widetilde{y}_{t+h}$ for y_t is

$$\widetilde{\varepsilon}_{y,t+h} = \begin{cases} u_{t+1}, & \text{for } h = 1 \\ u_{t+1} + u_{t+2} + \theta_1 u_{t+1} + \sum_{i=3}^{h} z_{t+i} = & \\ (1 + \theta_1 + \theta_2) \sum_{i=1}^{h-2} u_{t+i} + (1 + \theta_1)\, u_{t+h-1} + u_{t+h}, & \text{for } h > 1. \end{cases} \tag{16}$$

4 Assessing the forecast accuracy.

In this section we investigate the implications of using different specifications of the loss functions on model ranking based on the long-run forecasts. We subsequently explore how the long-run forecasts compare when judged in terms of three different loss functions. The first is the traditional trace MSFE loss function which penalizes forecast errors associated with the flow variables. The second is the triangular trace MSFE loss function based on the triangular representation of the standard I(1) cointegrated system as suggested in Christoffersen and Diebold (1998), which in a multicointegrated model corresponds to penalizing loss associated with forecast errors for the changes in stock variable as well as changes in one of the flow variables. The last loss function, which is suggested in this paper, is novel to the forecasting literature. It is termed the *extended triangular loss function* and it is based on the triangular representation of the multicointegrating variables. Thus, it can be considered an extension of the loss function suggested in Christoffersen and Diebold (1998) to multicointegrated systems. The feature of the suggested loss function is that the link between the levels of stock and flow variables is incorporated.

4.1 Traditional trace MSFE loss function.

First we use the trace MSFE criterion to compare the forecast accuracy of the multivariate and univariate forecast representations. The traditional trace MSFE loss function reads:

$$\text{trace MSFE} = \mathsf{E}\left[\begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}' \begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}\right], \tag{17}$$

where v_{1t+h} and v_{2t+h} are the forecast errors of the I(1) flow variables. As seen only the losses associated with flow variables are penalized in this case.

4.1.1 Trace MSFE for system forecasts.

Using the expressions for the system forecast errors in (12) and (13) we can calculate the following forecast error variances:

$$Var\left(\widehat{\varepsilon}_{x,t+h}\right) = h\sigma_1^2 \sim O\left(h\right), \text{ for } h > 0 \tag{18}$$

$$Var\left(\widehat{\varepsilon}_{y,t+h}\right) = \begin{cases} \left(\lambda+\alpha\right)^2\sigma_1^2 + \sigma_2^2, & \text{for } h = 1 \\ \lambda^2\sigma_1^2 h + \left[\left(\lambda+\alpha\right)^2 - \lambda^2\right]\sigma_1^2 + 2\sigma_2^2 \sim O\left(h\right), & \text{for } h > 1. \end{cases} \tag{19}$$

Notice that the variance of the system forecast error for y_{t+h} and x_{t+h} is growing of order $O\left(h\right)$. Then for the system forecasts we have

$$\text{trace } \widehat{MSFE} = \begin{cases} \sigma_1^2 + \left(\lambda+\alpha\right)^2\sigma_1^2 + \sigma_2^2, & h = 1 \\ \lambda^2\sigma_1^2 h + \left[\left(\lambda+\alpha\right)^2 - \lambda^2\right]\sigma_1^2 + 2\sigma_2^2 + h\sigma_1^2, & h > 1. \end{cases} \tag{20}$$

4.1.2 Trace MSFE for univariate forecasts.

Using expressions (15) and (16) we can calculate the variance of the univariate forecast errors:

$$Var\left(\widetilde{\varepsilon}_{x,t+h}\right) = Var\left(\widehat{\varepsilon}_{x,t+h}\right) = h\sigma_1^2 \sim O\left(h\right). \tag{21}$$

$$Var(\widetilde{\varepsilon}_{y,t+h}) = \begin{cases} \sigma_u^2, & \text{for } h = 1 \\ \left[\left(1+\theta_1+\theta_2\right)^2\left(h-2\right) + \left(1+\theta_1\right)^2 + 1\right]\sigma_u^2 = & \\ = \lambda^2\sigma_1^2\left(h-2\right) + \left[\left(1+\theta_1\right)^2 + 1\right]\sigma_u^2 \sim O\left(h\right), & \text{for } h > 1. \end{cases} \tag{22}$$

Observe that similar to the system forecast errors the variance of the univariate forecast errors grows of order $O(h)$. As a result we have

$$\text{trace } \widetilde{MSFE} = \begin{cases} \sigma_1^2 + \sigma_u^2, & h = 1 \\ \lambda^2\sigma_1^2\left(h-2\right) + \left[\left(1+\theta_1\right)^2 + 1\right]\sigma_u^2 + h\sigma_1^2 \sim O\left(h\right), & h > 1. \end{cases} \tag{23}$$

4.1.3 Trace MSFE ratio.

We can now compare the forecast accuracy of the system- and univariate models using the trace MSFE ratio as the forecast horizon increases:

$$\frac{\text{trace } \widetilde{MSFE}}{\text{trace } \widehat{MSFE}} = \frac{h\sigma_1^2 + \lambda^2 (h-1)\sigma_1^2 - \lambda^2\sigma_1^2 + \left[(1+\theta_1)^2 + 1\right]\sigma_u^2}{h\sigma_1^2 + \lambda^2 (h-1)\sigma_1^2 + (\lambda+\alpha)^2\sigma_1^2 + 2\sigma_2^2} \sim \frac{O(h)}{O(h)} \to 1. \tag{24}$$

As seen, as $h \to \infty$ this ratio approaches 1 since the coefficients to the leading terms both in the nominator and denominator are identical. That is, on the basis of the traditional forecast comparison criterion (trace MSFE ratio) it is impossible to distinguish between the model with imposed multicointegration restrictions and the model that ignores these restrictions completely. Thus, the conclusion of the use of the traditional trace MSFE ratio in assessing long-run system- and univariate forecasts in the multicointegrated systems coincides with that of Christoffersen and Diebold (1998) derived for the standard I(1) cointegrated model.

4.2 Triangular trace MSFE loss function.

In this section we investigate the implications of using the loss function suggested in Christoffersen and Diebold (1998) to long-run forecasts of the multicointegrating variables. Recall that this loss function has been proposed for evaluating long-run forecasts in the standard I(1) cointegrated system. The main point that we want to make is that the motivation for using Christoffersen and Diebold's loss function in the standard I(1) cointegrated systems carries over to the multicointegrating setup in a straightforward manner. This justifies the use of their loss function in multicointegrated models if the forecast evaluator is not concerned with losses associated with the linkage between the levels of stock and flow variables. This loss function has the interpretation of attaching loss to forecast errors associated with changes in stock and flow variables as opposed to the standard trace MSFE criterion which only accommodates losses associated with forecasting levels of flow variables. Below we illustrate this important finding.

First, it is worthwhile reviewing related results of Christoffersen and Diebold (1998) for the long-run forecasts in standard I(1) cointegrated systems. As discussed above, Christoffersen and Diebold (1998) show that when comparing the forecasting performance of models that impose cointegration and correctly specified univariate models in terms of the MSFE ratio, there are no gains of imposing cointegration except at the shortest forecast horizons. The problem is that the MSFE criterion fails to acknowledge the important distinction between long-run system forecasts and univariate forecasts. That is, the intrinsic feature of the long-run system forecasts is that they preserve the cointegrating relations exactly, whereas the long-run forecasts from the univariate models satisfy the cointegrating relations only on average. As a result, the variance of the cointegrating combination of the system forecast errors will always be smaller than that of the univariate forecast errors.

Therefore, if one can define a loss function which recognizes the distinction between system- and univariate forecasts, then it becomes possible to discriminate between the forecasts made from these models. Christoffersen and Diebold (1998) show that such a loss function can be based on the triangular representation of cointegrating variables, see Campbell and Shiller (1987), and Phillips (1991). In its simplest form a standard I(1) cointegrated system reads

$$\begin{bmatrix} 1-L & 0 \\ -\lambda & 1 \end{bmatrix} \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix},$$

where it is assumed that the disturbance terms are uncorrelated at all leads and lags. The corresponding loss function, introduced in Christoffersen and Diebold (1998), looks as follows:

$$\text{trace MSFE}_{tri} = \mathsf{E}\left[\left(\begin{bmatrix} 1-L & 0 \\ -\lambda & 1 \end{bmatrix} \begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}\right)' \left(\begin{bmatrix} 1-L & 0 \\ -\lambda & 1 \end{bmatrix} \begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}\right)\right],$$

such that the forecast accuracy of a given model is judged upon the linear transformations of the corresponding forecast errors v_{1t+h} and v_{2t+h} of the I(1) flow variables. Observe that for multicointegrated series the cointegrating combination of the forecast errors $v_{2,t+h} - \lambda v_{1,t+h}$ corresponds to the forecast errors of changes in the stock variable whereas $(1-L)\, v_{1,t+h}$ is the forecast error of changes in a flow variable. The trace MSFE_{tri} also reads

$$\text{trace MSFE}_{tri} = \mathsf{E}\left[\begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}' K \begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}\right], \text{ where } K = \begin{bmatrix} 1-L & 0 \\ -\lambda & 1 \end{bmatrix}' \begin{bmatrix} 1-L & 0 \\ -\lambda & 1 \end{bmatrix} \tag{25}$$

and it is instructive to compare this with the traditional MSFE used in other studies, see equation (17).

As seen, the traditional MSFE can be regarded as the special case of the trace MSFE_{tri} with K being the identity matrix. The trace MSFE_{tri} criterion values small forecast errors as does the traditional MSFE criterion, but at the same time it also values maintenance of the cointegrating restrictions amongst the generated forecasts. The latter fact proves to be crucial in distinguishing between system- and univariate forecasts. The well-recognized drawback of the trace MSFE criterion is that it fails to value the exact maintenance of cointegrating relations by the long-run forecasts. Hence, the solution is to employ a loss function that recognizes this fact. Recall that, as we have shown above, the long-run forecasts from the multicointegrated system obey the cointegrating relation exactly, see equation (11). Therefore, for our purpose it seems natural to adopt the loss function based on the triangular system.

4.2.1 Triangular trace MSFE for system forecasts.

In order to use the triangular trace MSFE criterion we need to compute the variance of the cointegrating combination of the forecast errors. Using expressions (12) and (13), it follows that:

$$Var\left(\widehat{\varepsilon}_{y,t+h}-\lambda\widehat{\varepsilon}_{x,t+h}\right)=\begin{cases}\alpha^2\sigma_1^2+\sigma_2^2, & \text{for } h=1\\ \alpha^2\sigma_1^2+2\sigma_2^2, & \text{for } h>1\end{cases} \tag{26}$$

which is finite for all forecast horizons. Observe that in this simple model the variance of the cointegrating combination of the forecast errors is the same for all forecast horizons except for $h=1$. The reason for the difference that occurs when $h=1$ can be seen from equations (9) and (10) which show that the multicointegrating term is in the information set for $h=1$ and it has expectation zero for $h>1$. Using the loss function (25), we have for the system forecasts

$$\text{trace }\widehat{\text{MSFE}}_{tri}=E\left\{\left(\begin{array}{c}(1-L)\,\widehat{\varepsilon}_{x,t+h}\\ \widehat{\varepsilon}_{y,t+h}-\lambda\widehat{\varepsilon}_{x,t+h}\end{array}\right)'\left(\begin{array}{c}(1-L)\,\widehat{\varepsilon}_{x,t+h}\\ \widehat{\varepsilon}_{y,t+h}-\lambda\widehat{\varepsilon}_{x,t+h}\end{array}\right)\right\}$$

and it follows that

$$\text{trace }\widehat{\text{MSFE}}_{tri}=\begin{cases}\alpha^2\sigma_1^2+\sigma_2^2+\sigma_1^2, & \text{for } h=1\\ \alpha^2\sigma_1^2+2\sigma_2^2+\sigma_1^2, & \text{for } h>1.\end{cases} \tag{27}$$

4.2.2 Triangular trace MSFE for univariate forecasts.

Next we derive the variance of the cointegrating combination of the univariate forecast errors:

$$Var\left(\widetilde{\varepsilon}_{y,t+h}-\lambda\widetilde{\varepsilon}_{x,t+h}\right)=Var\left(\widetilde{\varepsilon}_{y,t+h}\right)+\lambda^2Var\left(\widetilde{\varepsilon}_{x,t+h}\right)-2\lambda cov(\widetilde{\varepsilon}_{y,t+h},\widetilde{\varepsilon}_{x,t+h}),$$

using expressions (21) and (22) and the following expression for the covariance term

$$cov(\widetilde{\varepsilon}_{y,t+h},\widetilde{\varepsilon}_{x,t+h})=\lambda h\sigma_1^2+\alpha\sigma_1^2.$$

The variance of the cointegrating combination of the univariate forecast errors is

$$Var\left(\widetilde{\varepsilon}_{y,t+h}-\lambda\widetilde{\varepsilon}_{x,t+h}\right)=\begin{cases}\sigma_u^2-\lambda^2\sigma_1^2-2\lambda\alpha\sigma_1^2, & \text{for } h=1\\ -2\lambda^2\sigma_1^2-2\lambda\alpha\sigma_1^2+\left[(1+\theta_1)^2+1\right]\sigma_u^2, & \text{for } h>1.\end{cases} \tag{28}$$

This implies that the variance of the cointegrating combination of forecasts from the implied univariate representations is *finite* as well.

For the forecasts from the univariate models we have

$$\text{trace }\widetilde{\text{MSFE}}_{tri}=E\left\{\left(\begin{array}{c}(1-L)\,\widetilde{\varepsilon}_{x,t+h}\\ \widetilde{\varepsilon}_{y,t+h}-\lambda\widetilde{\varepsilon}_{x,t+h}\end{array}\right)'\left(\begin{array}{c}(1-L)\,\widetilde{\varepsilon}_{x,t+h}\\ \widetilde{\varepsilon}_{y,t+h}-\lambda\widetilde{\varepsilon}_{x,t+h}\end{array}\right)\right\}$$

and thus

$$\text{trace } \widetilde{\text{MSFE}}_{tri} = \begin{cases} \sigma_u^2 - \lambda^2\sigma_1^2 - 2\lambda\alpha\sigma_1^2 + \sigma_1^2, & \text{for } h = 1 \\ \left[(1+\theta_1)^2 + 1\right]\sigma_u^2 - 2\lambda^2\sigma_1^2 - 2\lambda\alpha\sigma_1^2 + \sigma_1^2, & \text{for } h > 1. \end{cases}$$

Further simplification results in

$$\text{trace } \widetilde{\text{MSFE}}_{tri} = \begin{cases} \sigma_u^2 - \lambda^2\sigma_1^2 - 2\lambda\alpha\sigma_1^2 - \alpha^2\sigma_1^2 + \alpha^2\sigma_1^2 + \sigma_1^2, & \text{for } h = 1 \\ \left[(1+\theta_1)^2 + 1\right]\sigma_u^2 - 2\lambda^2\sigma_1^2 - 2\lambda\alpha\sigma_1^2 - \alpha^2\sigma_1^2 + \alpha^2\sigma_1^2 + \sigma_1^2, & \text{for } h > 1, \end{cases}$$

$$\text{trace } \widetilde{\text{MSFE}}_{tri} = \begin{cases} \sigma_u^2 - (\lambda+\alpha)^2\sigma_1^2 - \sigma_2^2 + \alpha^2\sigma_1^2 + \sigma_2^2 + \sigma_1^2, & \text{for } h = 1 \\ \left[(1+\theta_1)^2 + 1\right]\sigma_u^2 - (\lambda+\alpha)^2\sigma_1^2 - 2\sigma_2^2 + 2\sigma_2^2 + \alpha^2\sigma_1^2 + \sigma_1^2, & \text{for } h > 1. \end{cases}$$

Then, using expressions (19) and (22) we have

$$\text{trace } \widetilde{\text{MSFE}}_{tri} = \begin{cases} Var(\widetilde{\varepsilon}_{y,t+h}) - Var(\widehat{\varepsilon}_{y,t+h}) + \alpha^2\sigma_1^2 + \sigma_2^2 + \sigma_1^2, & \text{for } h = 1 \\ Var(\widetilde{\varepsilon}_{y,t+h}) - Var(\widehat{\varepsilon}_{y,t+h}) + 2\sigma_2^2 + \alpha^2\sigma_1^2 + \sigma_1^2, & \text{for } h > 1. \end{cases} \tag{29}$$

4.2.3 Triangular trace MSFE ratio.

Using expressions (27) and (29) we can now compute the trace MSFE ratio's

$$\frac{\text{trace } \widetilde{\text{MSFE}}_{tri}^{h=1}}{\text{trace } \widehat{\text{MSFE}}_{tri}^{h=1}} = 1 + \frac{Var(\widetilde{\varepsilon}_{y,t+h}) - Var(\widehat{\varepsilon}_{y,t+h})}{\alpha^2\sigma_1^2 + \sigma_2^2 + \sigma_1^2} > 1, \tag{30}$$

$$\frac{\text{trace } \widetilde{\text{MSFE}}_{tri}^{h>1}}{\text{trace } \widehat{\text{MSFE}}_{tri}^{h>1}} = 1 + \frac{Var(\widetilde{\varepsilon}_{y,t+h}) - Var(\widehat{\varepsilon}_{y,t+h})}{\alpha^2\sigma_1^2 + 2\sigma_2^2 + \sigma_1^2} > 1. \tag{31}$$

The trace MSFE_{tri} ratio is constant and greater than one as the system forecasts based on the full information is more accurate than the univariate forecasts based on the partial information, i.e. $[Var(\widetilde{\varepsilon}_{y,t+h}) - Var(\widehat{\varepsilon}_{y,t+h})] > 0$ for all $h > 0$. Expressed in terms of the model parameters, expressions (30) and (31) read:

$$\frac{\text{trace } \widetilde{\text{MSFE}}_{tri}^{h=1}}{\text{trace } \widehat{\text{MSFE}}_{tri}^{h=1}} = \frac{\sigma_u^2 - \lambda^2\sigma_1^2 - 2\lambda\alpha\sigma_1^2 + \sigma_1^2}{\alpha^2\sigma_1^2 + \sigma_2^2 + \sigma_1^2} > 1, \tag{32}$$

$$\frac{\text{trace } \widetilde{\text{MSFE}}_{tri}^{h>1}}{\text{trace } \widehat{\text{MSFE}}_{tri}^{h>1}} = \frac{\left[(1+\theta_1)^2 + 1\right]\sigma_u^2 - 2\lambda^2\sigma_1^2 - 2\lambda\alpha\sigma_1^2 + \sigma_1^2}{\alpha^2\sigma_1^2 + 2\sigma_2^2 + \sigma_1^2} > 1. \tag{33}$$

In summary, several of the results in Christoffersen and Diebold (1998) derived for standard cointegrated systems carry over to models that obey multicointegrating restrictions. First, long-run forecasts generated from the multicointegrated system preserve the cointegrating relations exactly, see (11). Second, the system forecast errors follow the same stochastic process as the original variables, as depicted in (14) . Third, the variance of the cointegrating combination of the system forecast errors is finite (see (26)) even though the variance of the system forecast errors of the individual variables grow of order $O(h)$, as seen in expressions (18) and (19). Fourth, the variance of the cointegrating combination of the univariate forecast errors is finite too, see expression (28), even though the variance of the univariate forecast errors grows of order $O(h)$, see expressions (21) and (22). Fifth, imposing the multicointegrating restrictions does not lead to improved long-run forecast performance over the univariate models when compared in terms of the ratio of the traditional mean squared forecast error criterion, as shown in (24). Finally, adoption of a loss function based on the triangular representation of the standard I(1) cointegrated system leads to the superior ranking of the system forecasts over their univariate competitors, see expressions (30) and (31).

4.3 An extended triangular trace MSFE loss function.

As shown above adoption of the loss function suggested in Christoffersen and Diebold (1998) leads to clear superior ranking of the system long-run forecasts over the univariate long-run forecasts. Observe that this loss function incorporates only the first layer of cointegration while ignoring the second – the multicointegrating restriction. In this section we suggest a solution to this issue by proposing a new loss function that is based on the triangular representation of the multicointegrating system given in equation (3).

The suggested loss function looks as follows

$$\text{trace MSFE}^{\star}_{tri} =$$
$$= \mathsf{E}\left[\left(\begin{bmatrix} (1-L) & 0 \\ -\lambda(1-L)^{-1}-\alpha & (1-L)^{-1} \end{bmatrix}\begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}\right)'\left(\begin{bmatrix} (1-L) & 0 \\ -\lambda(1-L)^{-1}-\alpha & (1-L)^{-1} \end{bmatrix}\begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}\right)\right],$$

where v_{1t+h} and v_{2t+h} are the forecast errors of the I(1) flow variables. This can be rewritten as

$$\text{trace MSFE}^{\star}_{tri} = \mathsf{E}\left[\begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}' K \begin{bmatrix} v_{1t+h} \\ v_{2t+h} \end{bmatrix}\right] \tag{34}$$

with the K matrix given by

$$K = \begin{bmatrix} (1-L) & 0 \\ -\lambda(1-L)^{-1}-\alpha & (1-L)^{-1} \end{bmatrix}'\begin{bmatrix} (1-L) & 0 \\ -\lambda(1-L)^{-1}-\alpha & (1-L)^{-1} \end{bmatrix}. \tag{35}$$

Again, the suggested loss function can be considered a generalization of the traditional trace MSFE loss function presented in equation (17) where the K matrix is the identity matrix. The loss function (34) reflects the costs of deviating from the multicointegrating relation and hence explicitly accounts for the fact that the levels of stock and flow variables are directly interrelated.

Next, we illustrate the implications of using the new loss function in model ranking.

4.3.1 Trace $\text{MSFE}^{\star}_{tri}$ for system forecasts.

First we need to calculate the following expression

$$\text{trace}\ \widehat{\text{MSFE}}^{\star}_{tri} = \mathsf{E}\left[\begin{bmatrix} \widehat{\varepsilon}_{x,t+h} \\ \widehat{\varepsilon}_{y,t+h} \end{bmatrix}' K \begin{bmatrix} \widehat{\varepsilon}_{x,t+h} \\ \widehat{\varepsilon}_{y,t+h} \end{bmatrix}\right]$$

with the K matrix given in equation (35).

Using the results from Section 3 we can compute the following transformations of the system forecast errors:

$$\begin{aligned}
\widehat{\varepsilon}_{x,t+h} &= \sum_{i=1}^{h} e_{1t+i} \\
\widehat{\varepsilon}_{X,t+h} &= (1-L)^{-1}\widehat{\varepsilon}_{x,t+h} = \sum_{q=1}^{h}\widehat{\varepsilon}_{x,t+q} = \sum_{q=1}^{h}\sum_{i=1}^{q} e_{1t+i} = \sum_{i=1}^{h}(h+1-i)\, e_{1t+i} \\
\widehat{\varepsilon}_{Y,t+h} &= (1-L)^{-1}\widehat{\varepsilon}_{y,t+h} = \sum_{q=1}^{h}\widehat{\varepsilon}_{y,t+q} = \sum_{i=1}^{h}\left[\lambda(h+1-i)+\alpha\right] e_{1t+i} + e_{2t+h}.
\end{aligned}$$

Observe that we denote the cumulative forecast errors as $\widehat{\varepsilon}_{X,t+h}$ and $\widehat{\varepsilon}_{Y,t+h}$. This is because they effectively are the forecast errors of the levels of the generated I(2) variables X_{t+h} and Y_{t+h}, respectively.

The variance of the multicointegrating combination of the forecast errors is

$$Var\left(\widehat{\varepsilon}_{Y,t+h} - \lambda\widehat{\varepsilon}_{X,t+h} - \alpha\widehat{\varepsilon}_{x,t+h}\right) = \sigma_2^2. \tag{36}$$

This is finite, and for our simple model it is constant for all forecast horizons $h > 0$ as there is no short-run dynamics. The finding of a finite variance of the multicointegrating combination of the forecast errors is similar to that of Christoffersen and Diebold (1998), and Engle and Yoo (1987) for I(1) systems with the standard cointegrating restrictions. This is due to the fact that the forecast errors follow the same stochastic process as the forecasted time series. As a consequence, the forecasts are integrated of the same order and share the multicointegrating properties of the system dynamics as well.

In addition, observe that the corresponding forecast error variances of the transformed forecast errors are of the order $O\left(h^3\right)$ as seen below:

$$\begin{aligned}
Var\left(\widehat{\varepsilon}_{X,t+h}\right) &= Var\left(\sum_{q=1}^{h}\sum_{i=1}^{q} e_{1t+i}\right) = \frac{h(h+1)(2h+1)}{6}\sigma_1^2 \sim O\left(h^3\right) \\
Var\left(\widehat{\varepsilon}_{Y,t+h}\right) &= \frac{h(h+1)(2h+1)}{6}\lambda^2\sigma_1^2 + 2\alpha\lambda\frac{h(h+1)}{2}\sigma_1^2 + h\alpha^2\sigma_1^2 + \sigma_2^2 \sim O\left(h^3\right).
\end{aligned}$$

Using the expression $Var((1-L)\widehat{\varepsilon}_{x,t+h}) = \sigma_1^2$ we can calculate the trace $\text{MSFE}^{\star}_{tri}$ for the system forecasts

$$\text{trace}\ \widehat{\text{MSFE}}^{\star}_{tri} = \sigma_2^2 + \sigma_1^2. \tag{37}$$

4.3.2 Trace MSFE$^{\star}_{tri}$ for univariate forecasts.

Next we calculate the trace $\widetilde{\text{MSFE}}^{\star}_{tri}$ for the corresponding univariate forecasts

$$\text{trace } \widetilde{\text{MSFE}}^{\star}_{tri} = \mathsf{E}\left[\left[\begin{array}{c}\widetilde{\varepsilon}_{x,t+h}\\ \widetilde{\varepsilon}'_{y,t+h}\end{array}\right] K \left[\begin{array}{c}\widetilde{\varepsilon}_{x,t+h}\\ \widetilde{\varepsilon}'_{y,t+h}\end{array}\right]\right]$$

with the K matrix given in equation (35).

The transformation of the forecast errors yields the following

$$\begin{aligned}
\widetilde{\varepsilon}_{x,t+h} &= \sum_{i=1}^{h} e_{1t+i} \\
\widetilde{\varepsilon}_{X,t+h} &= (1-L)^{-1}\widetilde{\varepsilon}_{x,t+h} = \sum_{q=1}^{h}\sum_{i=1}^{q} e_{1t+i} = \sum_{i=1}^{h}(h+1-i)\, e_{1t+i}
\end{aligned}$$

with the variances

$$Var\left(\widetilde{\varepsilon}_{X,t+h}\right) = Var\left(\widehat{\varepsilon}_{X,t+h}\right) = \frac{h\left(h+1\right)\left(2h+1\right)}{6}\sigma_1^2 \sim O\left(h^3\right) \tag{38}$$

$$Var\left(\widetilde{\varepsilon}_{x,t+h}\right) = Var\left(\widehat{\varepsilon}_{x,t+h}\right) = h\sigma_1^2 \sim O\left(h\right). \tag{39}$$

The corresponding transformation of the forecast errors for y_{t+h} reads

$$\begin{aligned}
\widetilde{\varepsilon}_{Y,t+1} &= (1-L)^{-1}\widetilde{\varepsilon}_{y,t+1} = u_{t+1} \\
\widetilde{\varepsilon}_{Y,t+h} &= (1-L)^{-1}\widetilde{\varepsilon}_{y,t+h} = \\
&= \sum_{i=1}^{h-2}\left\{\left(1+\theta_1+\theta_2\right)\left(h-2-i+1\right)+\left(1+\theta_1\right)+1\right\}u_{t+i} + \left(\left(1+\theta_1\right)+1\right)u_{t+h-1} + u_{t+h}
\end{aligned}$$

with the corresponding forecast error variances

$$\begin{aligned}
Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) &= \left(1+\theta_1+\theta_2\right)^2\frac{\left(h-2\right)\left(h-2+1\right)\left(2\left(h-2\right)+1\right)}{6}\sigma_u^2 \\
&\quad +2\left(\left(1+\theta_1\right)+1\right)\left(1+\theta_1+\theta_2\right)\frac{\left(h-2\right)\left(h-2+1\right)}{2}\sigma_u^2 + \\
&\quad +\left(\left(1+\theta_1\right)+1\right)^2\left(h-1\right)\sigma_u^2+\sigma_u^2 \sim O\left(h^3\right).
\end{aligned} \tag{40}$$

Next, we calculate the variance of the polynomially cointegrating combination of the forecast errors from the univariate representation. Straightforward but tedious algebra relegated to the appendix yields the following relation

$$Var\left(\widetilde{\varepsilon}_{Y,t+h} - \lambda\widetilde{\varepsilon}_{X,t+h} - \alpha\widetilde{\varepsilon}_{x,t+h}\right) = \left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) - Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right] + \sigma_2^2. \tag{41}$$

Thus, using the result $Var((1-L)\widetilde{\varepsilon}_{x,t+h}) = \sigma_1^2$ the resulting trace $\widetilde{\text{MSFE}}^{\star}_{tri}$ for the univariate forecasts reads

$$\text{trace } \widetilde{\text{MSFE}}^{\star}_{tri} = \left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) - Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right] + \sigma_2^2 + \sigma_1^2. \tag{42}$$

It is of further interest to find the order of growth of the trace $\widetilde{\mathrm{MSFE}}^{\star}_{tri}$ for the univariate forecasts. As shown above both $Var\left(\widetilde{\varepsilon}_{Y,t+h}\right)$ and $Var\left(\widehat{\varepsilon}_{Y,t+h}\right)$ are $O(h^3)$. Thus, we want to find the growth order of the expression $\left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) - Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right]$. A little algebra and use of the expression $\lambda^2\sigma_1^2 = \left(1+\theta_1+\theta_2\right)^2\sigma_u^2$ yield the following result

$$\begin{aligned}
&\left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) - Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right] = \\
&= -h^2\lambda^2\sigma_1^2 - (h-1)^2\lambda^2\sigma_1^2 + 2\left(2+\theta_1\right)\left(1+\theta_1+\theta_2\right)\tfrac{(h-2)(h-2+1)}{2}\sigma_u^2 + \\
&+ \left(2+\theta_1\right)^2(h-1)\sigma_u^2 + \sigma_u^2 - 2\alpha\lambda\tfrac{h(h+1)}{2}\sigma_1^2 - h\alpha^2\sigma_1^2 - \sigma_2^2.
\end{aligned}$$

As seen, although each of the terms $Var\left(\widetilde{\varepsilon}_{Y,t+h}\right)$ and $Var\left(\widehat{\varepsilon}_{Y,t+h}\right)$ are $O(h^3)$, their difference $\left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) - -Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right]$ is $O(h^2)$. Therefore we have the following result

$$\text{trace } \widetilde{\mathrm{MSFE}}^{\star}_{tri} = \left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) - Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right] + \sigma_2^2 + \sigma_1^2 \sim O(h^2). \tag{43}$$

4.3.3 Ratio trace $\mathbf{MSFE}^{\star}_{tri}$ of the univariate to system forecasts.

Using equations (37) and (43) we can compute the trace $\mathrm{MSFE}^{\star}_{tri}$ ratio of the univariate to system forecasts

$$\frac{trace\ \widetilde{\mathrm{MSFE}}^{\star}_{tri}}{trace\ \widehat{\mathrm{MSFE}}^{\star}_{tri}} = \frac{\left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) - Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right] + \sigma_2^2 + \sigma_1^2}{\sigma_2^2 + \sigma_1^2} = 1 + \frac{\left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) - Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right]}{\sigma_2^2 + \sigma_1^2} > 1.$$

Intuitively, this inequality holds as the forecasts that utilize all the information in the system (system forecasts) will produce a smaller forecast error variance than the ones that are based on the partial information (univariate forecasts). It also resembles the trace MSFE_{tri} ratio in (30) and (31).

Using the results from above we can write

$$\frac{trace\ \widetilde{\mathrm{MSFE}}^{\star}_{tri}}{trace\ \widehat{\mathrm{MSFE}}^{\star}_{tri}} = \frac{O\left(h^2\right)}{O\left(1\right)} \to \infty \quad \text{as} \quad h \to \infty. \tag{44}$$

This means that we would prefer the model with multicointegrating restrictions using this criterion. In fact, there are high (increasing) gains to be achieved in using the new loss function both over the traditional MSFE loss function and the triangular MSFE loss function suggested in Christoffersen and Diebold (1998). The result (44) emphasizes that if in fact the forecast evaluator is concerned with losses associated with the stocks and flows not deviating too much from their steady state level, then this should be reflected in the loss function. As seen, huge gains can be achieved from the system forecast when compared to using simple univariate forecasts.

5 Example.

We illustrate the findings of the previous sections using the model (3) with the following values of the parameters $\lambda = 2, \alpha = 1, \sigma_1^2 = \sigma_2^2 = 1$. Such parameter combination leads to the following MA(2) process

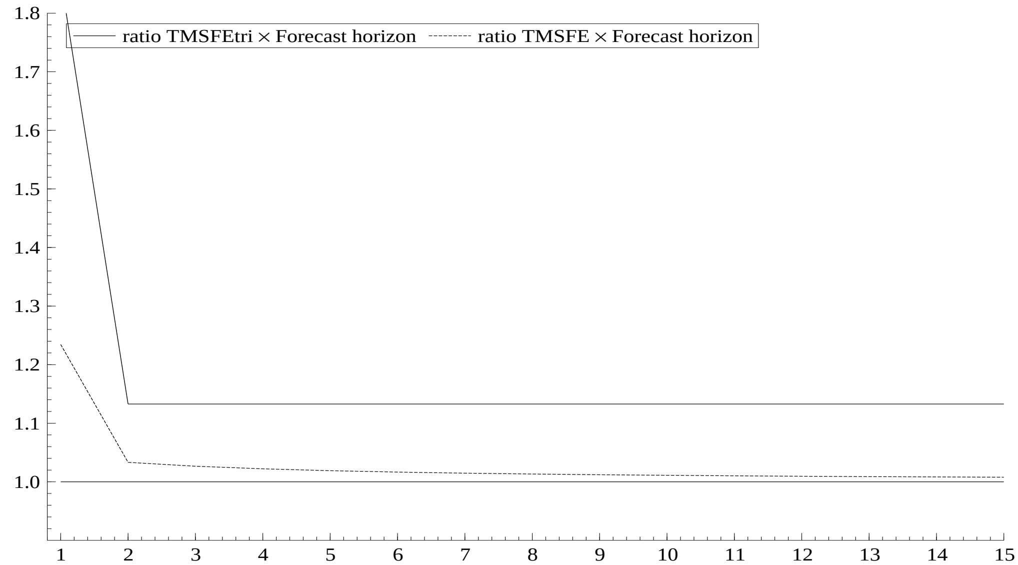

Figure 1: Trace MSFE ratio and Trace MSFE_{tri} ratio of univariate versus system forecasts of multicointegrating I(1) variables.

for Δy_t : $\theta_1 = -0.5155$, $\theta_2 = 0.0795$, and $\sigma_u^2 = 12.578$. Figures 1 and 2 are plotted using these true coefficient parameters.

Figure 1 displays the ratios (24) , (33) and (32). From Figure 1 it can be seen that for the standard trace MSFE ratio criterion there are no gains as h tends to infinity when system and univariate forecasts are compared. On the other hand, a specification with a loss function based on the triangular representation will induce persistent gains in system forecasting as compared to a univariate forecast model. These results are entirely in line with those reported by Christoffersen and Diebold (1998) for standard cointegration models.

Similarly, Figure 2 corresponds to the results given in expression (44). A loss function incorporating the multicointegrating restrictions across the levels of stock and flow variables is given in expression (34) and when the system forecasts are compared to the univariate forecasts, using this criterion, the (relative) explosive behavior of the univariate forecast errors is rather apparent as can be seen from Figure 2.

6 Conclusions.

In this paper we have extended the analysis of Christoffersen and Diebold (1998) to multicointegrated systems. The motivation was that in multicointegrated systems a complicated dynamic interaction of flow and stock variables may take place and in forecasting such variables a range of loss functions are

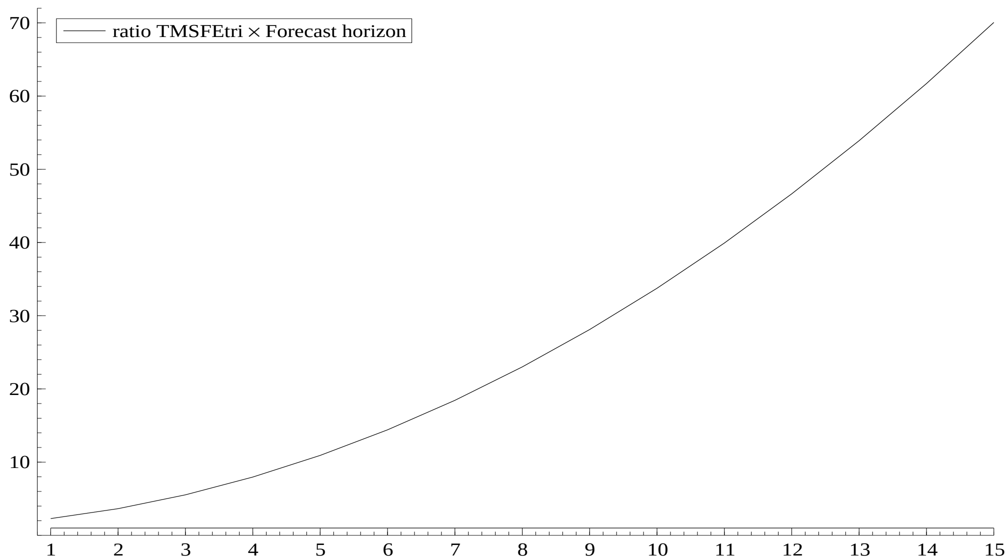

Figure 2: Trace $\mathrm{MSFE}^{\star}_{tri}$ ratio of univariate versus system forecasts of multicointegrating I(1) variables.

available when evaluating and comparing forecasts from different models. Christoffersen and Diebold's analysis can be conducted by assuming multicointegrated series rather than cointegrated series in the usual I(1) sense. When this occurs the variables can be given a particular interpretation. A loss function based on a standard trace mean square forecast error criterion corresponds to forecast evaluation where the forecast errors associated with the flows of the variables enter the loss function. On the other hand, the loss function of Christoffersen and Diebold, based on the triangular form representation of cointegrated systems, can be expressed in terms of losses associated with forecast errors of the changes in both the flow and stock variables. Although this loss function penalizes the deviations from the cointegrating relation only and completely ignores the multicointegrating restrictions, when applied to the forecasts from multicointegrated models, it clearly favours those over the univariate model forecasts. Notwithstanding, if the desired loss function should reflect the multicointegrating nature of the forecasted variables a new loss function is required. This function can be derived from the triangular representation of a multicointegrated system and we show that such a function will penalize deviations from a long-run stock and flow relation. In fact, the suggested loss function appears to have huge gains when compared to forecasts of the implied univariate models.

We do not want to take a strong stand upon which loss function to use in practice when evaluating different models. In this paper we have compared model forecasts from a correctly specified univariate model with that of a correctly specified system forecast. In model selection based on forecasting performance, one may prefer choosing a loss function which favors models which incorporate stronger (multicointegrating) restrictions on the variables that models which do not (i.e. the univariate models). Ultimately, however, the loss function to be chosen will reflect the preferences of the analyst.

The paper highlights the importance of carefully selecting loss functions when evaluating forecasts from cointegrated systems, and it shows how different loss functions based on a MSFE criterion help selection of competing models of increasing complexity. Comparing competing models, some of which are potentially incorrectly specified, is a different, though extremely relevant, issue. Deriving new results for multicointegrated systems along these lines, for instance by extending the work of Clements and Hendry (1995) to multicointegrated systems, is a topic for future research.

In this paper we used a simple bivariate, low-order multicointegrated model in order to establish the results. Naturally, it is of interest to derive the corresponding results for more general models that obey multicointegrating restrictions. Also, the consequences of introducing deterministic components are important as are estimation issues. These extensions will follow in future work.

References

Banerjee, A., L. Cockerell, and B. Russell (2001): "An I(2) Analysis of Inflation and the Markup," *Journal of Applied Econometrics*, 16(3), 221–40.

Campbell, J. Y., and R. J. Shiller (1987): "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95, 1052–1088.

Christoffersen, P. F., and F. X. Diebold (1998): "Cointegration and Long-Run Forecasting," *Journal of Business and Economic Statistics*, 16(4), 450–458.

Clements, M. P., and D. F. Hendry (1995): "Forecasting in Cointegrating Systems," *Journal of Applied Econometrics*, 10(2), 127–146.

Engle, R. F., and B. S. Yoo (1987): "Forecasting and Testing in Co-Integrated Systems," *Journal of Econometrics*, 35, 143–159.

Engsted, T., J. Gonzalo, and N. Haldrup (1997): "Testing for Multicointegration," *Economics Letters*, 56, 259–266.

Engsted, T., and N. Haldrup (1999): "Multicointegration in Stock-Flow Models," *Oxford Bulletin of Economics and Statistics*, 61, 237–254.

Engsted, T., and S. Johansen (1999): "Granger's Representation Theorem and Multicointegration," in *Cointegration, Causality and Forecasting, Festschrift in Honour of Clive Granger*, ed. by R. Engle, and H.White, Oxford. Oxford University Press.

Granger, C. W. J., and T. H. Lee (1989): "Investigation of Production, Sales and Inventory Relations Using Multicointegration and Non-Symmetric Error Correction Models," *Journal of Applied Econometrics*, 4, S145–S159.

Granger, C. W. J., and T. H. Lee (1991): "Multicointegration," in *Long-Run Economic Relationships. Reading in Cointegration*, ed. by R. F. Engle, and C. W. J. Granger, Advanced Texts in Econometrics, Oxford. Oxford University Press.

Haldrup, N. (1998): "An Econometric Analysis of I(2) Variables," *Journal of Economic Surveys*, 12(5), 595–650.

Hendry, D. F., and T. von Ungern-Sternberg (1981): "Liquidity and Inflation Effects on Consumers' Expenditure," in *Essay in the Theory and Measurement of Consumers' Behaviour*, ed. by A. S. Deaton, Cambridge. Cambridge University Press.

Leachman, L. L. (1996): "New Evidence on the Ricardian Equivalence Theorem: A Multicointegration Approach," *Applied Economics*, 28(6), 695–704.

Leachman, L. L., and B. B. Francis (2000): "Multicointegration Analysis of the Sustainability of Foreign Debt," *Journal of Macroeconomics*, 22(2), 207–27.

Lee, T. H. (1992): "Stock-Flow Relationships in US Housing Construction," *Oxford Bulletin of Economics and Statistics*, 54, 419–430.

Phillips, P. C. B. (1991): "Optimal Inference in Cointegrating Systems," *Econometrica*, 59, 283–306.

Rahbek, A., H. C. Kongsted, and C. Jørgensen (1999): "Trend-Stationarity in the I(2) Cointegration Model," *Journal of Econometrics*, 90, 265–289.

Siliverstovs, B. (2001): "Multicointegration in US Consumption Data," Aarhus University, Department of Economics, Working paper 2001-6.

7 Appendix.

7.1 Derivation of the implied univariate representation for Δy_t.

$$
\begin{aligned}
z_t &= [\lambda + (1-L)\,\alpha]\, e_{1t} + (1-L)^2\, e_{2t} \\
z_t &= \lambda e_{1t} + \alpha e_{1t} - \alpha e_{1t-1} + e_{2t} - 2e_{2t-1} + e_{2t-2}. \\
z_t &= u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}
\end{aligned}
$$

The autocovariance structure for z_t reads

$$
\begin{aligned}
\gamma_z(0) &= \left[(\lambda+\alpha)^2 + \alpha^2\right]\sigma_1^2 + 6\sigma_2^2 \\
\gamma_z(1) &= -\alpha\,(\lambda+\alpha)\,\sigma_1^2 - 4\sigma_2^2 \\
\gamma_z(2) &= \sigma_2^2 \\
\gamma_z(\tau) &= 0, \quad |\tau| \geq 3.
\end{aligned}
$$

This is a MA(2) process with the non-zero first and second autocorrelations. The first autocorrelation coefficient is

$$
\begin{aligned}
\rho_z(1) &= \frac{-\alpha\,(\lambda+\alpha)\,\sigma_1^2 - 4\sigma_2^2}{\left[(\lambda+\alpha)^2+\alpha^2\right]\sigma_1^2 + 6\sigma_2^2} = \frac{-\alpha\,(\lambda+\alpha)\,q - 4}{\left[(\lambda+\alpha)^2+\alpha^2\right]q + 6} \\
\rho_z(2) &= \frac{\sigma_2^2}{\left[(\lambda+\alpha)^2+\alpha^2\right]\sigma_1^2 + 6\sigma_2^2} = \frac{1}{\left[(\lambda+\alpha)^2+\alpha^2\right]q + 6},
\end{aligned}
$$

where

$$
q = \frac{\sigma_1^2}{\sigma_2^2}
$$

is the signal-to-noise ratio.

¿From this we can try to infer values for the parameters θ_1 and θ_2. By denoting

$$
A = [-\alpha\,(\lambda+\alpha)\,q - 4] \qquad B = \left[(\lambda+\alpha)^2+\alpha^2\right]q + 6
$$

and after some algebra we have that

$$
\theta_1 = \frac{\theta_2}{(1+\theta_2)}A
$$

and θ_2 is one of the root of the fourth-order polynomial

$$
\theta_2^4 + (2-B)\,\theta_2^3 + \left(A^2 - 2B + 2\right)\theta_2^2 + (2-B)\,\theta_2 + 1 = 0.
$$

Observe that the coefficient values θ_1 and θ_2 should satisfy the invertibility conditions for the MA(2) process z_t. The variance σ_u^2 is found from the following expression

$$\sigma_u^2 = \frac{\left[(\lambda+\alpha)^2 + \alpha^2\right]\sigma_1^2 + 6\sigma_2^2}{(1+\theta_1^2+\theta_2^2)} \quad \text{or} \quad \sigma_u^2 = \frac{\sigma_2^2}{\theta_2}.$$

Furthermore, the following relation holds

$$\frac{(1+\theta_1+\theta_2)^2}{(1+\theta_1^2+\theta_2^2)} = \frac{\lambda^2\sigma_1^2}{\left[(\lambda+\alpha)^2+\alpha^2\right]\sigma_1^2+6\sigma_2^2},$$

which further leads to

$$\lambda^2\sigma_1^2 = (1+\theta_1+\theta_2)^2\,\sigma_u^2.$$

7.2 Variance of the multicointegrating combination of univariate forecast errors.

Here, we calculate the variance of the multicointegrating combination of the forecast errors from the univariate representation:

$$\begin{aligned}
&Var\left(\widetilde{\varepsilon}_{Y,t+h} - \lambda\widetilde{\varepsilon}_{X,t+h} - \alpha\widetilde{\varepsilon}_{\Delta X,t+h}\right) = \\
&= Var\left(\widetilde{\varepsilon}_{Y,t+h} - \lambda\widetilde{\varepsilon}_{X,t+h}\right) + \alpha^2 Var\left(\widetilde{\varepsilon}_{\Delta X,t+h}\right) - 2\alpha Cov\left(\widetilde{\varepsilon}_{Y,t+h} - \lambda\widetilde{\varepsilon}_{X,t+h}, \widetilde{\varepsilon}_{\Delta X,t+h}\right) = \\
&= Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) + \lambda^2 Var\left(\widetilde{\varepsilon}_{X,t+h}\right) - 2\lambda Cov\left(\widetilde{\varepsilon}_{Y,t+h}, \widetilde{\varepsilon}_{X,t+h}\right) + \alpha^2 Var\left(\widetilde{\varepsilon}_{\Delta X,t+h}\right) - \\
&-2\alpha Cov\left(\widetilde{\varepsilon}_{Y,t+h}\widetilde{\varepsilon}_{\Delta X,t+h}\right) + 2\alpha\lambda Cov\left(\widetilde{\varepsilon}_{X,t+h}, \widetilde{\varepsilon}_{\Delta X,t+h}\right).
\end{aligned}$$

Thus, in order to calculate the variance of the multicointegrating combination of the forecast errors we need to derive the following expressions:

$$\begin{aligned}
Var\left(\widetilde{\varepsilon}_{Y,t+h}\right) &= (1+\theta_1+\theta_2)^2\,\frac{(h-2)(h-2+1)(2(h-2)+1)}{6}\sigma_u^2 \\
&\quad +2\left((1+\theta_1)+1\right)(1+\theta_1+\theta_2)\,\frac{(h-2)(h-2+1)}{2}\sigma_u^2 + \\
&\quad +\left((1+\theta_1)+1\right)^2(h-1)\,\sigma_u^2 + \sigma_u^2 \\
Var\left(\widetilde{\varepsilon}_{X,t+h}\right) &= Var\left(\sum_{q=1}^{h}\sum_{i=1}^{q} e_{1t+i}\right) = \left(h^2+(h-1)^2+..+1\right)\sigma_1^2 = \frac{h(h+1)(2h+1)}{6}\sigma_1^2 \\
Var\left(\widetilde{\varepsilon}_{\Delta X,t+h}\right) &= h\sigma_1^2
\end{aligned}$$

$$\begin{aligned}
Cov\left(\widetilde{\varepsilon}_{Y,t+h}, \widetilde{\varepsilon}_{X,t+h}\right) &= \lambda\frac{h(h+1)(2h+1)}{6}\sigma_1^2 + \alpha\frac{h(h+1)}{2}\sigma_1^2 \\
Cov\left(\widetilde{\varepsilon}_{Y,t+h}, \widetilde{\varepsilon}_{\Delta X,t+h}\right) &= \lambda\frac{h(h+1)}{2}\sigma_1^2 + \alpha h\sigma_1^2 \\
Cov\left(\widetilde{\varepsilon}_{X,t+h}, \widetilde{\varepsilon}_{\Delta X,t+h}\right) &= Cov\left(\sum_{q=1}^{h}\sum_{i=1}^{q} e_{1t+i}, \sum_{i=1}^{h} e_{1t+i}\right) = \frac{h(h+1)}{2}\sigma_1^2
\end{aligned}$$

These expressions lead to the following result:

$$\begin{aligned}
&Var\left(\widetilde{\varepsilon}_{Y,t+h}-\lambda\widetilde{\varepsilon}_{X,t+h}-\alpha\widetilde{\varepsilon}_{\Delta X,t+h}\right)=\\
&=Var\left(\widetilde{\varepsilon}_{Y,t+h}\right)-\lambda^2\tfrac{h(h+1)(2h+1)}{6}\sigma_1^2-2\alpha\lambda\tfrac{h(h+1)}{2}\sigma_1^2-\alpha^2h\sigma_1^2=\\
&=Var\left(\widetilde{\varepsilon}_{Y,t+h}\right)-\left[Var\left(\widehat{\varepsilon}_{Y,t+h}\right)-\sigma_2^2\right]=\left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right)-Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right]+\sigma_2^2.
\end{aligned}$$

Using the earlier result that $\lambda^2\sigma_1^2=\left(1+\theta_1+\theta_2\right)^2\sigma_u^2$ we get the following:

$$\begin{aligned}
&\left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right)-Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right]=\\
&=\left(1+\theta_1+\theta_2\right)^2\tfrac{(h-2)(h-2+1)(2(h-2)+1)}{6}\sigma_u^2+2\left(2+\theta_1\right)\left(1+\theta_1+\theta_2\right)\tfrac{(h-2)(h-2+1)}{2}\sigma_u^2+\left(2+\theta_1\right)^2\left(h-1\right)\sigma_u^2+\sigma_u^2-\\
&-\tfrac{h(h+1)(2h+1)}{6}\lambda^2\sigma_1^2-2\alpha\lambda\tfrac{h(h+1)}{2}\sigma_1^2-h\alpha^2\sigma_1^2-\sigma_2^2=\\
&=\tfrac{h(h+1)(2h+1)}{6}\lambda^2\sigma_1^2-\tfrac{4h(h+1)}{6}\lambda^2\sigma_1^2-\tfrac{2h(2h-3)}{6}\lambda^2\sigma_1^2-\tfrac{2(h-1)(2h-3)}{6}\lambda^2\sigma_1^2+2\left(2+\theta_1\right)\left(1+\theta_1+\theta_2\right)\tfrac{(h-2)(h-2+1)}{2}\sigma_u^2+\\
&+\left(2+\theta_1\right)^2\left(h-1\right)\sigma_u^2+\sigma_u^2-\tfrac{h(h+1)(2h+1)}{6}\lambda^2\sigma_1^2-2\alpha\lambda\tfrac{h(h+1)}{2}\sigma_1^2-h\alpha^2\sigma_1^2-\sigma_2^2=\\
&=-\tfrac{4h(h+1)}{6}\lambda^2\sigma_1^2-\tfrac{2h(2h-3)}{6}\lambda^2\sigma_1^2-\tfrac{2(h-1)(2h-3)}{6}\lambda^2\sigma_1^2+2\left(2+\theta_1\right)\left(1+\theta_1+\theta_2\right)\tfrac{(h-2)(h-2+1)}{2}\sigma_u^2+\\
&+\left(2+\theta_1\right)^2\left(h-1\right)\sigma_u^2+\sigma_u^2-2\alpha\lambda\tfrac{h(h+1)}{2}\sigma_1^2-h\alpha^2\sigma_1^2-\sigma_2^2.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&-\tfrac{4h(h+1)}{6}\lambda^2\sigma_1^2-\tfrac{2h(2h-3)}{6}\lambda^2\sigma_1^2-\tfrac{2(h-1)(2h-3)}{6}\lambda^2\sigma_1^2=\\
&=\tfrac{-4h^2-4h-4h^2+6h-4h^2+10h-6}{6}\lambda^2\sigma_1^2=\tfrac{-12h^2+12h-6}{6}\lambda^2\sigma_1^2=\left(-2h^2+2h-1\right)\lambda^2\sigma_1^2=\\
&=-h^2\lambda^2\sigma_1^2-\left(h-1\right)^2\lambda^2\sigma_1^2.
\end{aligned}$$

Then,

$$\begin{aligned}
&\left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right)-Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right]=-h^2\lambda^2\sigma_1^2-\left(h-1\right)^2\lambda^2\sigma_1^2+2\left(2+\theta_1\right)\left(1+\theta_1+\theta_2\right)\tfrac{(h-2)(h-2+1)}{2}\sigma_u^2+\\
&+\left(2+\theta_1\right)^2\left(h-1\right)\sigma_u^2+\sigma_u^2-2\alpha\lambda\tfrac{h(h+1)}{2}\sigma_1^2-h\alpha^2\sigma_1^2-\sigma_2^2.
\end{aligned}$$

The upshot is that even each of the expressions $Var\left(\widetilde{\varepsilon}_{Y,t+h}\right)$ and $Var\left(\widehat{\varepsilon}_{Y,t+h}\right)$ are $O(h^3)$, their difference $\left[Var\left(\widetilde{\varepsilon}_{Y,t+h}\right)-Var\left(\widehat{\varepsilon}_{Y,t+h}\right)\right]$ is $O(h^2)$.

Working Paper

2002-3: Bo Sandemann Rasmussen: Credibility, Cost of Reneging and the Choice of Fixed Exchange Rate Regime.

2002-4: Bo William Hansen and Lars Mayland Nielsen: Can Nominal Wage and Price Rigidities Be Equivalent Propagation Mechanisms? The Case of Open Economies.

2002-5: Anna Christina D'Addio and Michael Rosholm: Left-Censoring in Duration Data: Theory and Applications.

2002-6: Morten Ørregaard Nielsen: Efficient Inference in Multivariate Fractionally Integration Models.

2002-7: Morten Ørregaard Nielsen: Optimal Residual Based Tests for Fractional Cointegration and Exchange Rate Dynamics.

2002-8: Morten Ørregaard Nielsen: Local Whittle Analysis of Stationary Fractional Cointegration.

2002-9: Effrosyni Diamantoudi and Licun Xue: Coalitions, Agreements and Efficiency.

2002-10: Effrosyni Diamantoudi and Eftichios S. Sartzetakis: International Environmental Agreements - The Role of Foresight.

2002-11: N.E. Savin and Allan H. Würtz: Testing the Semiparametric Box–Cox Model with the Bootstrap.

2002-12: Morten Ø. Nielsen, Spectral Analysis of Fractionally Cointegrated Systems

2002-13: Anna Christina D'Addio and Bo E. Honoré, Duration Dependence and Timevarying Variables in Discrete Time Duration Models.

2002-14: Anna Christina D'Addio and Michael Rosholm, Labour Market Transitions of French Youth.

2002-15: Boriss Siliverstovs, Tom Engsted and Niels Haldrup, Long-run forecasting in multicointegrated systems.