FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall
Philadelphia, Pennsylvania 19106-1574
(215) 574-6428, www.phil.frb.org

Working Papers

Research Department

WORKING PAPER NO. 97-14

COINTEGRATION AND LONG-HORIZON FORECASTING

Peter F. Christoffersen
Research Department
International Monetary Fund

Francis X. Diebold
Department of Economics
University of Pennsylvania

June 1996
This Print: October 7, 1997

WORKING PAPER NO. 97-14

Cointegration and Long-Horizon Forecasting

Peter F. Christoffersen
Research Department
International Monetary Fund

Francis X. Diebold
Department of Economics
University of Pennsylvania

June 1996
This Print: October 7, 1997

Correspondence to:

F.X. Diebold
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104
fdiebold@mail.sas.upenn.edu

Acknowledgments: Helpful discussion was provided by Dave DeJong, Rob Engle, Clive Granger, Bruce Hansen, Dennis Hoffman, Laura Kodres, Jim Stock, Ruey Tsay, Ken Wallis, Mark Watson, Chuck Whiteman, Mike Wickens, Tao Zha, and participants at the July 1996 NBER/NSF conference on Forecasting and Empirical Methods in Macroeconomics. All remaining inadequacies are ours alone. We thank the National Science Foundation, the Sloan Foundation and the University of Pennsylvania Research Foundation for support.

1. Introduction

Cointegration implies restrictions on the low-frequency dynamic behavior of multivariate time series. Thus, imposition of cointegrating restrictions has immediate implications for the behavior of long-horizon forecasts, and it is widely believed that imposition of cointegrating restrictions, when they are in fact true, will produce superior long-horizon forecasts. Stock (1995, p. 1), for example, provides a nice distillation of the consensus when he asserts that "if the variables are cointegrated, their values are linked over the long run, and imposing this information can produce substantial improvements in forecasts over long horizons." The consensus stems from the theoretical result that long-horizon forecasts from cointegrated systems satisfy the cointegrating relationships exactly and the related result that only the cointegrating combinations of the variables can be forecast with finite long-horizon error variance. Moreover, it appears to be supported by a number of independent Monte Carlo analyses (e.g., Engle and Yoo, 1987; Reinsel and Ahn, 1992; Clements and Hendry, 1993; Lin and Tsay, 1996).

This paper grew out of an attempt to reconcile the popular intuition sketched above, which seems reasonable, with a competing conjecture, which also seems sensible. Forecast enhancement from exploiting cointegration comes from using information in the current deviations from the cointegrating relationships. That is, knowing whether and by how much the cointegrating relations are violated today is valuable in assessing where the variables will go tomorrow, because deviations from cointegrating relations tend to be eliminated. However, although the current value of the error-correction term clearly provides information about the likely *near*-horizon evolution of the system, it seems unlikely that it provides information about the *long*-horizon evolution of the system, because the long-horizon forecast of the error-correction term is always zero. (The error-correction term, by construction, is covariance stationary with a zero mean.) From this perspective, it seems unlikely that cointegration could be exploited to improve long-horizon forecasts.

Motivated by this apparent paradox, we provide a precise characterization of the implications of cointegration for long-horizon forecasting. In Section 2 we show that, contrary to popular belief, nothing is lost by ignoring cointegration when long-horizon forecasts are evaluated using standard accuracy measures; in fact, even *univariate* Box-Jenkins forecasts are equally accurate. Insection 3 we illustrate our results with a simple bivariate cointegrated system. In section 4, we address a potentially important deficiency of standard forecast accuracy measures highlighted by our analysis—they fail to value the maintenance of cointegrating relationships among variables—and we suggest alternative measures of accuracy that explicitly do so. In section 5, we consider forecasting from models with estimated parameters, and we use our results to clarify the interpretation of a number of well-known Monte Carlo studies. We conclude in section 6.

2. Multivariate and Univariate Forecasts of Cointegrated Variables

Assume that the Nx1 vector process of interest is generated by

$$(1-L)x_t = \mu + C(L)\varepsilon_t,$$

where C(L) is an NxN matrix lag operator polynomial of possibly infinite order. Then, under regularity conditions, the existence of r linearly independent cointegrating vectors is equivalent to rank(C(1)) = N-r, and the cointegrating vectors are given by the rows of the rxN matrix α', where $\alpha' C(1) = \alpha'\mu = 0$. That is, $z_t = \alpha' x_t$ is an r-dimensional *stationary* zero-mean time series. We will assume that the system is in fact cointegrated, with 0<rank(C(1))<N. For future reference, note

that following Stock and Watson (1988) we can use the decomposition $C(L) = C(1) + (1-L)C^*(L)$,

where $C_j^* = -\sum_{i=j+1}^{\infty} C_i$, to write the system in "common-trends" form,

$$x_t = \mu t + C(1)\xi_t + C^*(L)\varepsilon_t,$$

where $\xi_t = \sum_{i=1}^{t} \varepsilon_i$.

We will compare the accuracy of two forecasts of a multivariate cointegrated system that are polar extremes in terms of cointegrating restrictions imposed: first, forecasts from the multivariate model, and second, forecasts from the implied univariate models. Both forecasting models are correctly specified from a univariate perspective, but one imposes the cointegrating restrictions and one does not.[1]

We will make heavy use of a ubiquitous measure of forecast accuracy, mean squared error, the multivariate version of which is

$$MSE = E(e_{t+h}' K e_{t+h}),$$

where K is an NxN positive definite symmetric matrix and e_{t+h} is the vector of h-step-ahead forecast errors. MSE, of course, depends on the weighting matrix K. It is standard to set K=I, in which case

$$MSE = E(e_{t+h}' e_{t+h}) = trace(\Sigma_h),$$

where $\Sigma_h = var(e_{t+h})$. We call this the "trace MSE" accuracy measure. To compare the accuracy of two forecasts, say 1 to 2, it is standard to examine the ratio $\frac{trace(\Sigma_h^1)}{trace(\Sigma_h^2)}$, which we call

[1] The work of Clements and Hendry (1994, 1995) is closely related to ours. They compare forecasts from the true VAR to forecasts from a misspecified VAR in differences, whereas we compare forecasts from the true VAR to exact forecasts from correctly specified univariate representations. More important, our motivation and focus are very different from Clements and Hendry's, as will become clear shortly.

the “trace MSE ratio.”

Forecasts from the Multivariate Cointegrated System[2]

From the moving average representation, we can unravel the process recursively from time t+h to time 1 and write

$$x_{t+h} = (t+h)\mu + \sum_{i=1}^{t}\sum_{j=0}^{t+h-i} C_j\varepsilon_i + \sum_{i=1}^{h}\sum_{j=0}^{h-i} C_j\varepsilon_{t+i},$$

from which the h-step-ahead forecasts are easily calculated as

$$\hat{x}_{t+h} = (t+h)\mu + \sum_{i=1}^{t}\sum_{j=0}^{t+h-i} C_j\varepsilon_i.$$

From the fact that

$$\lim_{h\to\infty} \sum_{j=0}^{t+h-i} C_j = C(1),$$

we get that

$$\lim_{h\to\infty} \alpha' \hat{x}_{t+h} = 0,$$

so that the cointegrating relationship is satisfied exactly by the long-horizon system forecasts. This is the sense in which long-horizon forecasts from cointegrated systems hang together correctly.

We define the h-step-ahead forecast error from the multivariate system as

$$\hat{e}_{t+h} = x_{t+h} - \hat{x}_{t+h}.$$

The forecast errors from the multivariate system satisfy

---

[2] Many results on forecasting in cointegrated systems relevant for our purposes, and cataloged in this subsection, are contained in the lucid and insightful paper of Engle and Yoo (1987).

$$\hat{e}_{t+h} = \sum_{i=1}^{h} \sum_{j=0}^{h-i} C_j \varepsilon_{t+i},$$

so the variance of the h-step-ahead forecast error is

$$\mathrm{var}[\hat{e}_{t+h}] = \sum_{i=1}^{h} \left[\left(\sum_{j=0}^{h-i} C_j \right) \Omega \left(\sum_{j=0}^{h-i} C_j' \right) \right],$$

where Ω is the variance of ε_t.

From the definition of $\hat{e}_{t+h}$ we can also see that the system forecast errors satisfy

$$\hat{e}_{t+h} - \hat{e}_{t+h-1} = \sum_{i=1}^{h} C_{h-i} \varepsilon_{t+i} = C(L)\varepsilon_{t+h},$$

where the last equality holds if we take ε_j=0 for all j<t. That is, when we view the system forecast error process as a function of the forecast horizon, h, it has the same stochastic structure as the original process, x_t, and therefore is integrated and cointegrated. Consequently, the variance of the h-step ahead forecast errors from the cointegrated system grows like h,

$$\mathrm{var}[\hat{e}_{t+h}] = O(h).$$

In contrast, the cointegrating combinations of the system forecast errors, just as the error-correction process z_t, will have finite variance for large h,

$$\mathrm{var}[\alpha' \hat{e}_{t+h}] = \alpha' Q \alpha < \infty,$$

where the matrix Q is a constant function of the stationary component of the forecast error. Although individual series can be forecast only with increasingly wide confidence intervals, the cointegrating combination has a confidence interval of finite width, even as the forecast horizon goes to infinity.

<u>Forecasts from the Implied Univariate Representations</u>

Now consider ignoring the multivariate features of the system, forecasting instead using the implied univariate representations. We can use Wold's decomposition theorem and write for any series (the n-th, say),

$$(1-L)x_{n,t} = \mu_n + \sum_{j=0}^{\infty} \theta_{n,j} u_{n,t-j},$$

where $\theta_{n,0} = 1$ and $u_{n,t}$ is white noise. It follows from this expression that the univariate time-t forecast for period t+h is,

$$\tilde{x}_{n,t+h} = h\mu_n + x_{n,t} + \left(\sum_{i=1}^{h} \theta_{n,i} \right) u_{n,t} + \left(\sum_{i=2}^{h+1} \theta_{n,i} \right) u_{n,t-1} + \ldots$$

Using obvious notation we can write

$$\tilde{x}_{n,t+h} = h\mu_n + x_{n,t} + \theta_n(L) u_{n,t},$$

and stacking the N series we have

$$\tilde{x}_{t+h} = h\mu + x_t + \Theta(L) u_t,$$

where $\Theta(L)$ is a diagonal matrix polynomial with the individual $\theta_n(L)$'s on the diagonal.

Now let us consider the errors from the univariate forecasts. We will rely on the following convenient decomposition

$$\tilde{e}_{t+h} \equiv x_{t+h} - \tilde{x}_{t+h} = (x_{t+h} - \hat{x}_{t+h}) + (\hat{x}_{t+h} - \tilde{x}_{t+h}).$$

Recall that the system forecast is

$$\hat{x}_{t+h} = \mu(t+h) + \sum_{i=1}^{t} \sum_{j=0}^{t+h-i} C_j \varepsilon_i \approx \mu(t+h) + C(1)\xi_t,$$

where the approximation holds as h gets large. Using univariate forecasts, the decomposition for $\tilde{e}_{t+h}$, and the approximate long-horizon system forecast, we get

$$\tilde{e}_{t+h} \approx \hat{e}_{t+h} + \mu(t+h) + C(1)\xi_t - (x_t + \mu h + \Theta(L)u_t).$$

Now insert the common trends representation for x_t to get

$$\tilde{e}_{t+h} \approx \hat{e}_{t+h} + \mu(t+h) + C(1)\xi_t - (\mu t + C(1)\xi_t + C^*(L)\varepsilon_t + \mu h + \Theta(L)u_t),$$

and finally cancel terms to get

$$\tilde{e}_{t+h} \approx \hat{e}_{t+h} - (C^*(L)\varepsilon_t + \Theta(L)u_t).$$

Notice that the ε_t's are serially uncorrelated and the u_t's depend only on current and past ε_t's; thus, the two terms in the expression are orthogonal. Notice also that the second term is just a sum of stationary series and is therefore stationary; furthermore, its variance is constant as the forecast horizon h changes. We can therefore write the long-horizon variance of the univariate forecasts as

$$\mathrm{var}(\tilde{e}_{t+h}) = \mathrm{Var}(\hat{e}_{t+h}) + O(1) = O(h) + O(1) = O(h),$$

which is of the same order of magnitude as the variance of the *system* forecast errors. Furthermore, the trace MSE ratio goes to one. Thus, when comparing accuracy using the trace MSE ratio, the univariate forecasts perform as well as the cointegrated system forecasts as the horizon gets large. This is the opposite of the folk wisdom—it turns out that imposition of cointegrating restrictions helps at short, but not long, horizons. Quite simply, when accuracy is evaluated with the trace MSE ratio, there is no long-horizon benefit from imposing cointegration; all that matters is getting the level of *integration* right. We summarize the result as:

Proposition 1

$$\lim_{h\to\infty} \frac{\mathrm{trace}(\mathrm{var}(\tilde{e}_{t+h}))}{\mathrm{trace}(\mathrm{var}(\hat{e}_{t+h}))} = 1.$$

Proposition 1 provides the theoretical foundation for the results of Hoffman and Rasche (1996), who find in an extensive empirical application that imposing cointegration does little to enhance long-horizon forecast accuracy, and Brandner and Kunst (1990), who suggest that when in doubt about how many unit roots to impose in a long-horizon forecasting model, it's less harmful to impose too many than to impose too few.

Now let's consider the variance of cointegrating combinations of univariate forecast errors. Above we recounted the Engle-Yoo (1987) result that the cointegrating combinations of the system forecast errors have finite variance as the forecast horizon gets large. Now we want to look at the same cointegrating combinations of the *univariate* forecast errors. From our earlier derivations it follows that

$$\alpha' \tilde{e}_{t+h} \approx \alpha' \hat{e}_{t+h} - (\alpha' C^*(L)\varepsilon_t + \alpha' \Theta(L) u_t).$$

Again we can rely on the orthogonality of the two terms. The first term has finite variance, as discussed above. So too does the second, because it is a linear combination of stationary processes. Thus we have

Proposition 2

$$\mathrm{var}(\alpha' \tilde{e}_{t+h}) = \alpha' Q \alpha + \alpha' \mathrm{var}(C^*(L)\varepsilon_t + \Theta(L) u_t)\alpha = O(1).$$

The cointegrating combinations of the long-horizon errors from the univariate forecasts, which completely ignore cointegration, *also have finite variance*. Thus, it is, in fact, not imposition of cointegration on the forecasting system that yields the finite variance of the cointegrating combination of the errors; rather, it is the cointegration property inherent in the system itself.

3. A Simple Example

To illustrate our results in a transparent way, we consider the simple bivariate cointegrated system,

$$x_t = \mu + x_{t-1} + \varepsilon_t$$

$$y_t = \lambda x_t + v_t,$$

where the disturbances are orthogonal at all leads and lags. The moving average representation is

$$(1-L)\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \mu \\ \lambda\mu \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ \lambda & 1-L \end{pmatrix}\begin{pmatrix} \varepsilon_t \\ v_t \end{pmatrix} \equiv \begin{pmatrix} \mu \\ \lambda\mu \end{pmatrix} + C(L)\begin{pmatrix} \varepsilon_t \\ v_t \end{pmatrix},$$

and the error-correction representation is

$$(1-L)\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \mu \\ \lambda\mu \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(\lambda \quad -1)\begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \lambda\varepsilon_t + v_t \end{pmatrix}.$$

The system's simplicity allows us to compute exact formulae that correspond to the qualitative results derived in the previous section.

<u>Univariate Representations</u>

Let us first derive the implied univariate representations for x and y. The univariate representation for x is, of course, a random walk with drift, exactly as given in the first equation of the system,

$$x_t = \mu + x_{t-1} + \varepsilon_t.$$

Derivation of the univariate representation for y is a bit more involved. From the moving-average representation of the system, rewrite the process for y_t as a univariate two-shock process,

$$y_t = \lambda\mu + y_{t-1} + (1-L)v_t + \lambda\varepsilon_t$$

$$= \lambda\mu + y_{t-1} + z_t,$$

where $z_t = (1-L)v_t + \lambda\varepsilon_t$. The autocovariance structure for z_t is

$$\gamma_z(0) = 2\sigma_v^2 + \lambda^2\sigma_\varepsilon^2$$

$$\gamma_z(1) = \gamma_z(-1) = -\sigma_v^2$$

$$\gamma_z(\tau) = 0, \ |\tau| \geq 2.$$

The only non-zero positive autocorrelation is therefore

$$\rho_z(1) = -\frac{\sigma_v^2}{2\sigma_v^2 + \lambda^2\sigma_\varepsilon^2} = -\frac{1}{2+\lambda^2 q},$$

where $q = \sigma_\varepsilon^2/\sigma_v^2$ is the signal to noise ratio. This is exactly the autocorrelation structure of an MA(1) process, so we write $z_t = \theta u_{t-1} + u_t$. To find the value for θ, match autocorrelations at lag 1, yielding

$$\frac{\theta}{1+\theta^2} = -\frac{1}{2+\lambda^2 q}.$$

This gives a second-order polynomial in θ, with invertible solution

$$\theta = (1/2)[\sqrt{\lambda^4 q^2 + 4\lambda^2 q} - 2 - \lambda^2 q].$$

Finally, we find the variance of the univariate innovation by matching the variances, yielding

$$(1+\theta^2)\sigma_u^2 = 2\sigma_v^2 + \lambda^2\sigma_\varepsilon^2 = \sigma_v^2(2+\lambda^2 q),$$

or

$$\sigma_u^2 = \frac{\sigma_v^2(2+\lambda^2 q)}{(1+\theta^2)}.$$

Forecasts from the Multivariate Cointegrated System

First consider forecasting from the multivariate cointegrated system. Write the time t+h values as

$$x_{t+h} = \mu h + x_t + \sum_{i=1}^{h} \varepsilon_{t+i}$$

$$y_{t+h} = \lambda(\mu h + x_t) + \lambda \sum_{i=1}^{h} \varepsilon_{t+i} + v_{t+h}.$$

The h-step-ahead forecasts are

$$\hat{x}_{t+h} = \mu h + x_t$$

$$\hat{y}_{t+h} = \lambda\mu h + \lambda x_t,$$

and the h-step-ahead forecast errors are

$$\hat{e}_{x,t+h} = \sum_{i=1}^{h} \varepsilon_{t+i}$$

$$\hat{e}_{y,t+h} = \lambda \sum_{i=1}^{h} \varepsilon_{t+i} + v_{t+h}.$$

Note that the forecast errors follow the same stochastic process as the original system (aside from the drift term),

$$(1-L)\hat{e}_{t+h} = \begin{pmatrix} \varepsilon_{t+h} \\ \lambda\varepsilon_{t+h}+(1-L)v_{t+h} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda & 1-L \end{pmatrix}\begin{pmatrix} \varepsilon_{t+h} \\ v_{t+h} \end{pmatrix} = C(L)\begin{pmatrix} \varepsilon_{t+h} \\ v_{t+h} \end{pmatrix}.$$

Finally, the corresponding forecast error variances are

$$\mathrm{var}(\hat{e}_{x,t+h}) = h\sigma_{\varepsilon}^2$$

$$\mathrm{var}(\hat{e}_{y,t+h}) = h\lambda^2\sigma_{\varepsilon}^2 + \sigma_v^2.$$

Both forecast error variances are O(h). As for the variance of the cointegrating combination, we have

$$\mathrm{var}[\hat{e}_{y,t+h}-\lambda\hat{e}_{x,t+h}] = \mathrm{var}\left[\left(\lambda\sum_{i=1}^{h} \varepsilon_{t+i}+v_{t+h}\right) - \lambda\left(\sum_{i=1}^{h} \varepsilon_{t+i}\right)\right] = \sigma_v^2,$$

for all h, because there are no short-run dynamics.[3] Similarly, because we have no short-run dynamics, the forecasts satisfy the cointegrating relationship at all horizons, not just in the limit. That is,

---

[3] We are using the fact that var[(λ, -1)e_{t+h}] = var[- (λ, -1)e_{t+h}].

$$\hat{y}_{t+h} - \lambda \hat{x}_{t+h} = 0, \ \forall \ h = 1, 2, \ldots$$

Forecasts from the Implied Univariate Representations

Now consider forecasting from the implied univariate models. Immediately, the univariate forecast for x is the same as the system forecast,

$$\tilde{x}_{t+h} = \mu h + x_t.$$

Thus,

$$\tilde{e}_{x,t+h} = \hat{e}_{x,t+h} = \sum_{i=1}^{h} \varepsilon_{t+i},$$

so that

$$\text{var}(\tilde{e}_{x,t+h}) = \text{var}(\hat{e}_{x,t+h}) = h\sigma_{\varepsilon}^2 = O(h).$$

To form the univariate forecast for y, write

$$y_{t+h} = h\lambda\mu + y_t + \theta u_t + u_{t+1} + \sum_{i=2}^{h} z_{t+i}.$$

The forecast is

$$\tilde{y}_{t+h} = h\lambda\mu + y_t + \theta u_t,$$

and the corresponding forecast error is

$$\tilde{e}_{y,t+h} = u_{t+1} + \sum_{i=2}^{h} z_{t+i} = (1+\theta)\sum_{i=1}^{h-1} u_{t+i} + u_{t+h},$$

yielding the forecast error variance

$$\text{var}(\tilde{e}_{y,t+h}) = [(1+\theta)^2(h-1) + 1]\sigma_u^2.$$

Notice in particular that the univariate forecast error variance is O(h), as is the system forecast error variance.

Now let's compute the variance of the cointegrating combination of univariate forecast errors. We have

$$var[\tilde{e}_{y,t+h} - \lambda \tilde{e}_{x,t+h}] = [2\lambda^2 qh + 2 + \theta]\sigma_v^2 - 2\lambda \ cov\left(\tilde{e}_{y,t+h}, \tilde{e}_{x,t+h}\right).$$

To evaluate the covariance term, use the fact that

$$\tilde{y}_{t+h} = h\lambda\mu + y_t + \theta u_t = h\lambda\mu + \lambda x_t + v_t + \theta u_t = \hat{y}_{t+h} + v_t + \theta u_t$$

to write

$$\tilde{e}_{y,t+h} = \lambda \sum_{i=1}^{h} \varepsilon_{t+i} + v_{t+h} - v_t - \theta u_t.$$

Now recall the formula for the forecast error of x and the fact that future values of ε are uncorrelated with future and current values of v, and with current values of u, so that

$$cov\left(\tilde{e}_{y,t+h}, \tilde{e}_{x,t+h}\right) = E\left[\left(\sum_{i=1}^{h} \varepsilon_{t+i}\right)\left(\lambda \sum_{i=1}^{h} \varepsilon_{t+i} + v_{t+h} - v_t - \theta u_t\right)\right] = \lambda h \sigma_\varepsilon^2.$$

Armed with this result, we have that[4]

$$var[\tilde{e}_{y,t+h} - \lambda \tilde{e}_{x,t+h}] = (2+\theta)\sigma_v^2 < \infty \ \forall h,$$

which, of course, accords with our general result derived earlier: that the variance of the cointegrating combination of *univariate* forecast errors is finite.

Forecast Accuracy Comparison

Finally, compare the forecast error variances from the multivariate and univariate

[4] Keep in mind that θ is a function of $q = \sigma_\varepsilon^2 / \sigma_v^2$.

representations. Of course, x has the same representation in both, so the comparison hinges on y. We must compare

$$\mathrm{var}(\hat{e}_{y,t+h}) = h\lambda^2\sigma_\varepsilon^2 + \sigma_v^2 = \sigma_v^2[h\lambda^2 q + 1]$$

to

$$\mathrm{var}(\tilde{e}_{y,t+h}) = [(1+\theta)^2(h-1) + 1]\sigma_u^2.$$

Expanding the product in the expression for $\mathrm{var}(\tilde{e}_{y,t+h})$ yields

$$\mathrm{var}(\tilde{e}_{y,t+h}) = [(1+\theta)^2 h - \theta(2+\theta)]\sigma_u^2.$$

Substituting for σ_u^2, and using the fact that

$$\frac{\theta}{1+\theta^2} = -\frac{1}{2+\lambda^2 q} \quad \Rightarrow \quad \frac{(1+\theta)^2}{1+\theta^2} = \frac{\lambda^2 q}{2+\lambda^2 q},$$

we get

$$\mathrm{var}(\tilde{e}_{y,t+h}) = \left[\frac{(1+\theta)^2}{1+\theta^2}h - \theta\frac{2+\theta}{1+\theta^2}\right]\sigma_v^2(2+\lambda^2 q) = [\lambda^2 qh + 2 + \theta]\sigma_v^2.$$

Thus,

$$\mathrm{var}(\tilde{e}_{y,t+h}) = \mathrm{var}(\hat{e}_{y,t+h}) + (1+\theta)\sigma_v^2.$$

The error variance of the univariate forecast is greater than that of the system forecast, but it grows at the same rate.

Assembling all of the results, we have immediately that

$$\frac{\mathrm{trace}(\mathrm{var}(\tilde{e}_{t+h}))}{\mathrm{trace}(\mathrm{var}(\hat{e}_{t+h}))} = \frac{\mathrm{var}(\tilde{e}_{x,t+h}) + \mathrm{var}(\tilde{e}_{y,t+h})}{\mathrm{var}(\hat{e}_{x,t+h}) + \mathrm{var}(\hat{e}_{y,t+h})} = \frac{qh + \lambda^2 qh + 2 + \theta}{qh + 1 + \lambda^2 hq}.$$

In Figure 1 we show the values of this ratio as h gets large, for $q = \lambda = 1$. Note in particular the speed with which the limiting result,

$$\lim_{h\to\infty} \frac{\text{trace}(\text{var}(\tilde{e}_{t+h}))}{\text{trace}(\text{var}(\hat{e}_{t+h}))} = 1,$$

obtains.

In closing this section, we note that in spite of the fact that the trace MSE ratio approaches 1, the ratio of the variances of the cointegrating combinations of the forecast errors does *not* approach 1 in this simple model; rather,

$$\frac{\text{var}[\tilde{e}_{y,t+h} - \lambda \tilde{e}_{x,t+h}]}{\text{var}[\hat{e}_{y,t+h} - \lambda \hat{e}_{x,t+h}]} = (2+\theta) > 1, \quad \forall\ h.$$

This observation turns out to hold quite generally, and it forms the basis for an improved class of accuracy measures, to which we now turn.

4. Accuracy Measures and Cointegration

Accuracy Measures I: Trace MSE

We have seen that long-horizon univariate forecasts of cointegrated variables (which completely ignore cointegrating restrictions) are just as accurate as their system counterparts (which explicitly impose cointegrating restrictions), when accuracy is evaluated using the standard trace MSE criterion. So on traditional grounds there is no reason to prefer long-horizon forecasts from the cointegrated system.

One might argue, however, that the system forecasts are nevertheless more appealing because "... the forecasts of levels of co-integrated variables will 'hang together' in a way likely to be viewed as sensible by an economist, whereas forecasts produced in some other way, such as by a group of individual, univariate Box-Jenkins models, may well not do so" (Granger and Newbold, 1986, p. 226). But as we have seen, univariate Box-Jenkins forecasts *do* hang together if the variables are cointegrated—the cointegrating combinations, and only the cointegrating combinations, of univariate forecast errors have finite variance.

Accuracy Measures II: Trace MSE in Forecasting the Cointegrating Combinations of Variables

But all is not lost. The long-horizon system forecasts do a better job of satisfying the cointegrating restrictions than do the univariate forecasts—the long-horizon system forecasts *always* satisfy the cointegrating restrictions, whereas the long-horizon univariate forecasts do so only on average. That's what's responsible for our earlier result in our bivariate system: that, although the cointegrating combinations of both the univariate and system forecast errors have finite variance, the variance of the cointegrating combination of the univariate errors is larger.

Such effects are lost on standard accuracy measures like trace MSE, however, because the loss functions that underlie them don't value long-run forecasts' hanging together. The solution is obvious: if we value maintenance of the cointegrating relationship, then so, too, should the loss functions underlying our forecast accuracy measures. One approach, in the spirit of Granger (1996), is to focus on forecasting the cointegrating combinations of the variables and to evaluate forecasts in terms of the variability of the cointegrating combinations of the errors, $\alpha' e_{t+h}$.

Accuracy measures based on cointegrating combinations of the forecast errors require that the cointegrating vector be known. Fortunately, such is often the case. Horvath and Watson (1995, pp. 984-985), for example, note that[5]

> "Economic models often imply that variables are cointegrated with simple and known cointegrating vectors. Examples include the neoclassical growth model, which implies that income, consumption, investment, and the capital stock will grow in a balanced way, so that any stochastic growth in one of the series must be matched by corresponding growth in the others. Asset pricing models with stable

[5] See also Watson (1994) and Zivot (1996).

risk premia imply corresponding stable differences in spot and forward prices, long- and short-term interest rates, and the logarithms of stock prices and dividends. Most theories of international trade imply long-run purchasing power parity, so that long-run movements in nominal exchange rates are matched by countries' relative price levels. Certain monetarist propositions are centered around the stability of velocity, implying cointegration among the logarithms of money, prices and income. Each of these theories has distinct implications for the properties of economic time series under study: First, the series are cointegrated, and second, the cointegrating vector takes on a specific value. For example, balanced growth implies that the logarithms of income and consumption are cointegrated and that the cointegrating vector takes on the value of (1, -1)."

Thus, although the assumption of a known cointegrating vector certainly involves a loss of generality, it is nevertheless legitimate in a variety of empirically and economically relevant cases. This is fortunate because of problems associated with identification of cointegrating vectors in estimated systems, as stressed in Wickens (1996). We will maintain the assumption throughout this paper.

Interestingly, evaluation of accuracy in terms of the trace MSE *of the cointegrating combinations of forecast errors* is a special case of the general mean squared error measure. To see this, consider the general N-variate case with r cointegrating relationships, and consider again the mean squared error,

$$E(e_{t+h}'Ke_{t+h}) = E\ \text{trace}(e_{t+h}'Ke_{t+h}) = E\ \text{trace}(Ke_{t+h}e_{t+h}') = \text{trace}(K\Sigma_h),$$

where Σ_h is the variance of e_{t+h}. Evaluating accuracy in terms of trace MSE of the cointegrating combinations of the forecast errors amounts to evaluating

$$E\Big((\alpha' e_{t+h})'(\alpha' e_{t+h})\Big) = \text{trace}\ E\Big((\alpha' e_{t+h})'(\alpha' e_{t+h})\Big) = \text{trace}(K\Sigma_h),$$

where $K = \alpha\alpha'$. Thus the trace MSE of the cointegrating combinations of the forecast errors is in fact a particular variant of MSE formulated on the raw forecast errors, $E(e'Ke) = \text{trace}(K\Sigma_h)$, where the weighting matrix $K = \alpha\alpha'$ is of (deficient) rank r, the cointegrating rank of the system.

Accuracy Measures III: Trace MSE from the Triangular Representation

The problem with the traditional $E(e'Ke)$ approach with $K = I$ is that, although it values small MSE, it fails to value the long-run forecasts' hanging together correctly. Conversely, a problem with the $E(e'Ke)$ approach with $K = \alpha\alpha'$ is that it values *only* the long-run forecasts' hanging together correctly, whereas both pieces seem clearly relevant. The challenge is to incorporate both pieces into an overall accuracy measure in a natural way, and an attractive approach for doing so follows from the triangular representation of cointegrated systems exploited by Campbell and Shiller (1987) and Phillips (1991).

From the fact that α' has rank r, it is possible to rewrite the system so that the N left-hand-side variables are the r error-correction terms followed by the differences of N-r integrated but not cointegrated variables. That is, we rewrite the system in terms of $\begin{pmatrix} x_{1t} - \Gamma' x_{2t} \\ (1-L)x_{2t} \end{pmatrix}$, where the variables have been rearranged and partitioned into $x_t = (x_{1t}', x_{2t}')'$, where $\Gamma = \Gamma(\alpha)$ and the variables in x_{2t} are integrated but not cointegrated. We then evaluate accuracy in terms of the trace MSE of forecasts *from the triangular system*,

$$E\left\{ \begin{pmatrix} e_{1,t+h} - \Gamma' e_{2,t+h} \\ (1-L)e_{2,t+h} \end{pmatrix}' \begin{pmatrix} e_{1,t+h} - \Gamma' e_{2,t+h} \\ (1-L)e_{2,t+h} \end{pmatrix} \right\} = E\left\{ \left(\begin{bmatrix} I_r & -\Gamma' \\ 0 & (1-L) \end{bmatrix} e_{t+h} \right)' \left(\begin{bmatrix} I_r & -\Gamma' \\ 0 & (1-L) \end{bmatrix} e_{t+h} \right) \right\},$$

which we denote trace MSE_{tri}. Notice that the trace MSE_{tri} accuracy measure is also of $E(e'Ke)$ form, with

$$K = K(L) = \begin{bmatrix} I_r & -\Gamma' \\ 0 & (1-L) \end{bmatrix}' \begin{bmatrix} I_r & -\Gamma' \\ 0 & (1-L) \end{bmatrix}.$$

Recall Proposition 1, which says that under trace MSE, long-horizon forecast accuracy from the cointegrated system is no better than that from univariate models. We now show that under trace MSE_{tri}, long-horizon forecast accuracy from the cointegrated system is *always* better than that from univariate models.

Proposition 3

$$\lim_{h\to\infty} \frac{\text{trace } \tilde{MSE}_{tri}}{\text{trace } \hat{MSE}_{tri}} > 1.$$

Proof: Consider a cointegrated system in triangular form, that is, a system such that $\alpha' = [\, I \ \ -\Gamma' \,]$. We need to show that for large h,

$$\sum_{i=1}^{r} \text{var}[\alpha_i' \hat{e}_{t+h}] + \sum_{j=r+1}^{N} \text{var}[(1-L)\hat{e}_{j,t+h}] < \sum_{i=1}^{r} \text{var}[\alpha_i' \tilde{e}_{t+h}] + \sum_{j=r+1}^{N} \text{var}[(1-L)\tilde{e}_{j,t+h}]$$

and

$$\sum_{i=1}^{r} \text{var}[\alpha_i' \hat{e}_{t+h}] + \sum_{j=r+1}^{N} \text{var}[(1-L)\hat{e}_{j,t+h}] < \infty.$$

To establish the first inequality it is sufficient to show that

$$\sum_{i=1}^{r} \text{var}[\alpha_i' \hat{e}_{t+h}] < \sum_{i=1}^{r} \text{var}[\alpha_i' \tilde{e}_{t+h}].$$

We showed earlier that for large h,

$$\text{var}(\alpha' \tilde{e}_{t+h}) = \alpha' Q \alpha + \alpha' \text{var}(C^*(L)\varepsilon_t + \Theta(L)u_t)\alpha \equiv \alpha'(Q+S)\alpha,$$

where $\alpha' Q \alpha = \text{var}(\alpha' \hat{e}_{t+h})$, from which it follows that

$$\sum_{i=1}^{r} \text{var}[\alpha_i' \tilde{e}_{t+h}] - \sum_{i=1}^{r} \text{var}[\alpha_i' \hat{e}_{t+h}] = \text{tr}(\alpha' S \alpha) > 0,$$

because S is positive definite. To establish the second inequality, recall that

$$\hat{e}_{t+h} - \hat{e}_{t+h-1} = \sum_{i=1}^{h} C_{h-i}\varepsilon_{t+i} = C(L)\varepsilon_{t+h},$$

so that

$$\mathrm{var}[(1-L)\hat{e}_{t+h}] = \left(\sum_{j=0}^{h-1} C_j\right)\Omega\left(\sum_{j=0}^{h-1} C_j^{\prime}\right) \rightarrow C(1)\Omega C(1)^{\prime} \quad \text{as } h\rightarrow\infty.$$

Let $C_{N\text{-}r}(1)$ be the last N-r rows of C(1); then altogether we have

$$\sum_{i=1}^{r} \mathrm{var}[\alpha_i^{\prime}\hat{e}_{t+h}] + \sum_{j=r+1}^{N} \mathrm{var}[(1-L)\hat{e}_{j,t+h}] = \mathrm{tr}(\alpha^{\prime}Q\alpha) + \mathrm{tr}(C_{N-r}(1)\Omega C_{N-r}(1)^{\prime}) < \infty,$$

and the proof is complete.

<u>The Bivariate Example, Revisited</u>

In our simple bivariate example all we have to do to put the system in the triangular form sketched above is to switch x and y in the autoregressive representation, yielding

$$\begin{pmatrix} 1 & -\lambda \\ 0 & 1-L \end{pmatrix}\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} 0 \\ \mu \end{pmatrix} + \begin{pmatrix} v_t \\ \varepsilon_t \end{pmatrix}.$$

For the system forecasts we have

$$\text{trace } M\hat{S}E_{tri} = E\left\{\begin{pmatrix} \hat{e}_{y,t+h} - \lambda\hat{e}_{x,t+h} \\ (1-L)\hat{e}_{x,t+h} \end{pmatrix}^{\prime}\begin{pmatrix} \hat{e}_{y,t+h} - \lambda\hat{e}_{x,t+h} \\ (1-L)\hat{e}_{x,t+h} \end{pmatrix}\right\} = \sigma_v^2 + \sigma_\varepsilon^2.$$

For the univariate forecasts we have

$$\text{trace } M\tilde{S}E_{tri} = E\left\{\begin{pmatrix} \tilde{e}_{y,t+h} - \lambda\tilde{e}_{x,t+h} \\ (1-L)\tilde{e}_{x,t+h} \end{pmatrix}^{\prime}\begin{pmatrix} \tilde{e}_{y,t+h} - \lambda\tilde{e}_{x,t+h} \\ (1-L)\tilde{e}_{x,t+h} \end{pmatrix}\right\} = (2+\theta)\sigma_v^2 + \sigma_\varepsilon^2.$$

Thus, we see that the trace MSE_{tri} ratio does not approach one as the horizon increases; in particular, it is constant and above one for all h,

$$\frac{\text{trace } \tilde{\text{MSE}}_{\text{tri}}}{\text{trace } \hat{\text{MSE}}_{\text{tri}}} = \frac{1+(2+\theta)q}{1+q} = 1 + \frac{(1+\theta)q}{1+q} > 1, \quad \forall h.$$

In Figure 2 we plot the trace MSE_{tri} ratio vs. h, for λ=1 and $\sigma_\varepsilon^2 = \sigma_v^2 = q = 1$.

In summary, although the long-horizon performances of the system and univariate forecasts are identical under the conventional trace MSE accuracy measure, they differ under trace MSE_{tri}. The system forecast is superior to the univariate forecast under trace MSE_{tri}, because the system forecast is accurate in the conventional "small MSE" sense *and* it hangs together correctly.

5. Understanding Earlier Monte Carlo Studies

Here we clarify the interpretation of earlier influential Monte Carlo work, in particular Engle and Yoo (1987), as well as Reinsel and Ahn (1992), Clements and Hendry (1993), and Lin and Tsay (1996), among others. We do so by performing a Monte Carlo analysis of our own, which reconciles our theoretical results and the apparently conflicting Monte Carlo results reported in the literature, and shows how the existing Monte Carlo analyses have been misinterpreted. Throughout, we use our simple bivariate system (which is very similar to the one used by Engle and Yoo), with parameters set to λ=1, μ=0 and $\sigma_\varepsilon^2 = \sigma_v^2 = 1$. We use a sample size of 100 and perform 4000 Monte Carlo replications. In keeping with our earlier discussion, we assume a known cointegrating vector, but we estimate all other parameters.[6]

Let us first consider an analog of our theoretical results except we now estimate parameters instead of assuming them known. In Figure 3 we plot the trace MSE ratio and the

[6] This simple design allows us to make our point forcefully and with a minimum of clutter. The results are robust to changes in parameter values and sample size.

trace MSE_{tri} ratio vs. h. Using estimated parameters changes none of the theoretical results reached earlier under the assumption of known parameters. Use of the trace MSE ratio obscures the long-horizon benefits of imposing cointegration, whereas use of trace MSE_{tri} reveals those benefits clearly.

How then can we reconcile our results with those of Engle and Yoo (1987) and the many subsequent authors who conclude that imposing cointegration produces superior long-horizon forecasts? The answer is two-part: Engle and Yoo make a different and harder-to-interpret comparison than we do, and they misinterpret the outcome of their Monte Carlo experiments.

First, consider the forecast comparison. We have thus far compared forecasts from univariate models (which impose integration) to forecasts from the cointegrated system (which impose both integration and cointegration). Thus, a comparison of the forecasting results isolates the effects of imposing cointegration. Engle and Yoo, in contrast, compare forecasts from a VAR in levels (which impose *neither* integration nor cointegration) to forecasts from the cointegrated system (which impose *both* integration and cointegration). Thus, differences in forecasting performance in the Engle-Yoo setup cannot necessarily be attributed to the imposition of cointegration; instead, they may simply be due to imposition of integration, irrespective of whether cointegration is imposed.

Now consider the interpretation of the results. The VAR in levels is of course integrated, but in finite samples, the well-known Dickey-Fuller-Hurwitz bias tends to produce parameter estimates in the covariance stationary region. Thus, it’s no surprise that forecasts from the VAR estimated in levels perform poorly, with performance worsening with horizon, as shown in Figure 4. It is tempting to attribute the poor performance of the VAR in levels to its failure to impose cointegration, as do Engle and Yoo. The fact is, however, that the VAR in levels performs poorly because it fails to impose *integration*, not because it fails to impose *cointegration*—estimation of the cointegrated system simply imposes the correct level of integration *a priori*. To see this, consider Figure 5, in which we compare the forecasts from an estimated VAR in *differences* to the forecasts from the estimated cointegrated system. At long horizons, the forecasts from the VAR in differences, which impose integration but completely ignore cointegration, perform just as well.[7] In contrast, if we instead evaluate forecast accuracy with the trace MSE_{tri} ratio that we have advocated, the forecasts from the VAR in differences compare poorly at all horizons to those from the cointegrated system, as shown in Figure 6.

In our simple bivariate system, we are restricted to studying models with exactly one unit root and one cointegration relationship. It is also of interest to examine richer systems; conveniently, the literature already contains relevant (but unnoticed) evidence, which is entirely consistent with our theoretical results. Reinsel and Ahn (1992) and Lin and Tsay (1996), in particular, provide Monte Carlo evidence on the comparative forecasting performance of competing estimated models. Both study a four-variable VAR(2), with two unit roots and two cointegrating relationships. Their results clearly suggest that under the trace MSE accuracy measure, one need only worry about imposing enough unit roots on the system. Imposing three (one too many) unit roots is harmless at *any* horizon, and imposing four unit roots (two too many, so that the VAR is in differences) is harmless at long horizons.

[7] Figure 5 corresponds to our simple bivariate system. We have also duplicated the result on the Engle-Yoo system. Moreover, the result is apparent (but not noticed or discussed) in the Monte Carlo results of Reinsel and Ahn (1992), Clements and Hendry (1993), and Lin and Tsay (1996).

6. Summary and Concluding Remarks

First, we have shown that imposing cointegration does not improve the accuracy of long-horizon forecasts when forecasts of cointegrated variables are evaluated using the standard trace MSE ratio. Ironically enough, although cointegration implies restrictions on low-frequency dynamics, imposing cointegration is helpful for short- but not long-horizon forecasting, in contrast to the impression created in the literature. Imposition of cointegration on an estimated system, when the system is in fact cointegrated, helps the accuracy of long-horizon forecasts relative to those from systems estimated in levels with no restrictions, but that is because of the imposition of integration, not cointegration. Univariate forecasts in differences do just as well!

Second, we have shown that the variance of the cointegrating combination of the long-horizon forecast errors is finite regardless of whether cointegration is imposed. The variance of the error in forecasting the cointegrating combination *is* smaller, however, for the cointegrated system forecast errors. This suggests that accuracy measures that value long-run forecasts' hanging together correctly should be defined, in part, on the cointegrating combinations of the forecast errors. We explored one such accuracy measure based on the triangular representation of the cointegrated system.

Third, we showed that our theoretical results are entirely consistent with several well-known Monte Carlo analyses, whose interpretation we clarified. The existing Monte Carlo results are correct, but their widespread interpretation is not. Imposition of *integration*, not cointegration, is responsible for the repeated finding that the long-horizon forecasting performance of cointegrated systems is better than that of VARs in levels.

We hasten to add that the message of this paper is *not* that cointegration is of no value in forecasting. First, even under the conventional trace MSE accuracy measure, imposing cointegration *does* improve forecasts. Our message is simply that, under the conventional accuracy measure, it does so at short and moderate, not long, horizons, in contrast to the folk wisdom. Second, in our view, imposing cointegration certainly *may be* of value in long-horizon forecasting; the problem is simply that standard forecast accuracy measures don't reveal it. The upshot is that in forecast evaluation we need to think hard about which characteristics make a good forecast good and how best to measure those characteristics.[8] Seemingly omnibus measures such as trace MSE, although certainly useful in many situations, are inadequate in others.

In closing, we emphasize that the particular alternative to trace MSE that we examine in this paper, trace MSE_{tri}, is but one among many possibilities, and we look forward to exploring variations in future research. The key insight, it seems to us, is that if we value preservation of cointegrating relationships in long-horizon forecasts, then so, too, should our accuracy measures, and trace MSE_{tri} is a natural loss function that does so.

Interestingly, it is possible to process the trace MSE differently to obtain an accuracy measure that ranks the system forecasts as superior to the univariate forecasts, even as the forecast horizon goes to infinity. One obvious candidate is the trace MSE *difference*, as opposed to the trace MSE *ratio*. It follows from the results of section 2 that the trace MSE difference is

---

[8] In that respect this paper is in the tradition of our earlier work, such as Diebold and Mariano (1995), Diebold and Lopez (1996), and Christoffersen and Diebold (1996a, 1996b), in which we argue the virtues of tailoring accuracy measures in applied forecasting to the specifics of the problem at hand.

positive and does not approach zero as the forecast horizon grows.[9] As stressed above, however, it seems more natural to work with alternatives to trace MSE that explicitly value preservation of cointegrating relationships, rather then simply processing the trace MSE differently. As the forecast horizon grows, the trace MSE difference becomes negligible *relative* to either the system or the univariate trace MSE, so that the trace MSE difference would appear to place too little value on preserving cointegrating relationships.

[9] The idea of checking whether the trace MSE difference is positive follows immediately from the more general idea of checking whether $\Sigma_1 - \Sigma_2$ is positive definite, as advocated in Wallis (1995).

References

Brandner, P. and Kunst, R.M. (1990), "Forecasting Vector Autoregressions - The Influence of Cointegration: A Monte Carlo Study," Research Memorandum No. 265, Institute for Advanced Studies, Vienna.

Campbell, J.Y. and Shiller, R.J. (1987), "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95, 1062-1088.

Christoffersen, P.F. and Diebold, F.X. (1996a), "Optimal Prediction Under Asymmetric Loss," *Econometric Theory*, forthcoming.

Christoffersen, P.F. and Diebold, F.X. (1996b), "Further Results on Forecasting and Model Selection Under Asymmetric Loss," *Journal of Applied Econometrics*, 11, 561-571.

Clements, M.P. and Hendry, D.F. (1993), "On the Limitations of Comparing Mean Square Forecast Errors," *Journal of Forecasting*, 12, 617-637.

Clements, M.P. and Hendry, D.F. (1994), "Towards a Theory of Economic Forecasting," in C.P. Hargreaves (ed.), *Nonstationary Time Series and Cointegration*. Oxford: Oxford University Press.

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

Diebold, F.X. and Lopez, J. (1996), "Forecast Evaluation and Combination," in G.S. Maddala and C.R. Rao (eds.), *Handbook of Statistics.* Amsterdam: North-Holland, 241-268.

Diebold, F.X. and Mariano, R.S. (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253-265.

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

Granger, C.W.J. (1996), "Can We Improve the Perceived Quality of Economic Forecasts?," *Journal of Applied Econometrics*, 11, 455-473.

Granger, C.W.J., and Newbold, P. (1986), *Forecasting Economic Time Series*, Second Edition. New York: Academic Press.

Hoffman, D.L. and Rasche, R.H. (1996), "Assessing Forecast Performance in a Cointegrated System," *Journal of Applied Econometrics*, 11, 495-517.

Horvath, M.T.K. and Watson, M.W. (1995), "Testing for Cointegration When Some of the Cointegrating Vectors are Known," *Econometric Theory*, 11, 984-1014.

Lin, J.-L. and Tsay, R.S. (1996), "Cointegration Constraints and Forecasting: An Empirical Examination," *Journal of Applied Econometrics*, 11, 519-538.

Phillips, P.C.B. (1991), "Optimal Inference in Cointegrated Systems," *Econometrica*, 59, 283-306.

Reinsel, G.C. and Ahn, S.K. (1992), "Vector Autoregressive Models with Unit Roots and Reduced Rank Structure: Estimation, Likelihood Ratio Test, and Forecasting," *Journal of Time Series Analysis*, 13, 353-375.

Stock, J.H. (1995), "Point Forecasts and Prediction Intervals for Long Horizon Forecasts," Manuscript, J.F.K. School of Government, Harvard University.

Stock, J.H. and Watson, M.W. (1988), "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097-1107.

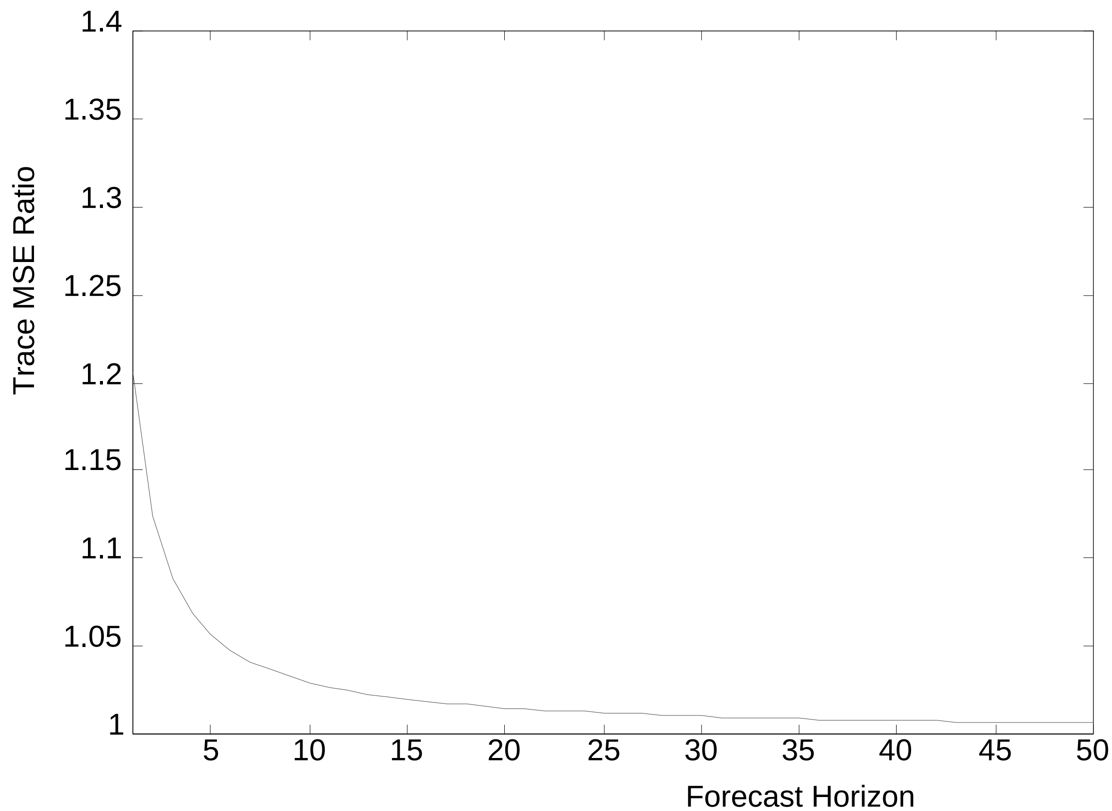
Wallis, K.F. (1995), "Large-Scale Macroeconomic Modeling," in M.H. Pesaran and M.R. Wickens (eds.), *Handbook of Applied Econometrics*, Chapter 6. Oxford: Blackwell.

Watson, M.W. (1994), "Vector Autoregressions and Cointegration," in R.F. Engle and D. McFadden (eds.), *Handbook of Econometrics, Vol. IV*, Chapter 47. Amsterdam: North-Holland.

Wickens, M.R. (1996), "Interpreting Cointegrating Vectors and Common Stochastic Trends," *Journal of Econometrics*, 74, 255-271.

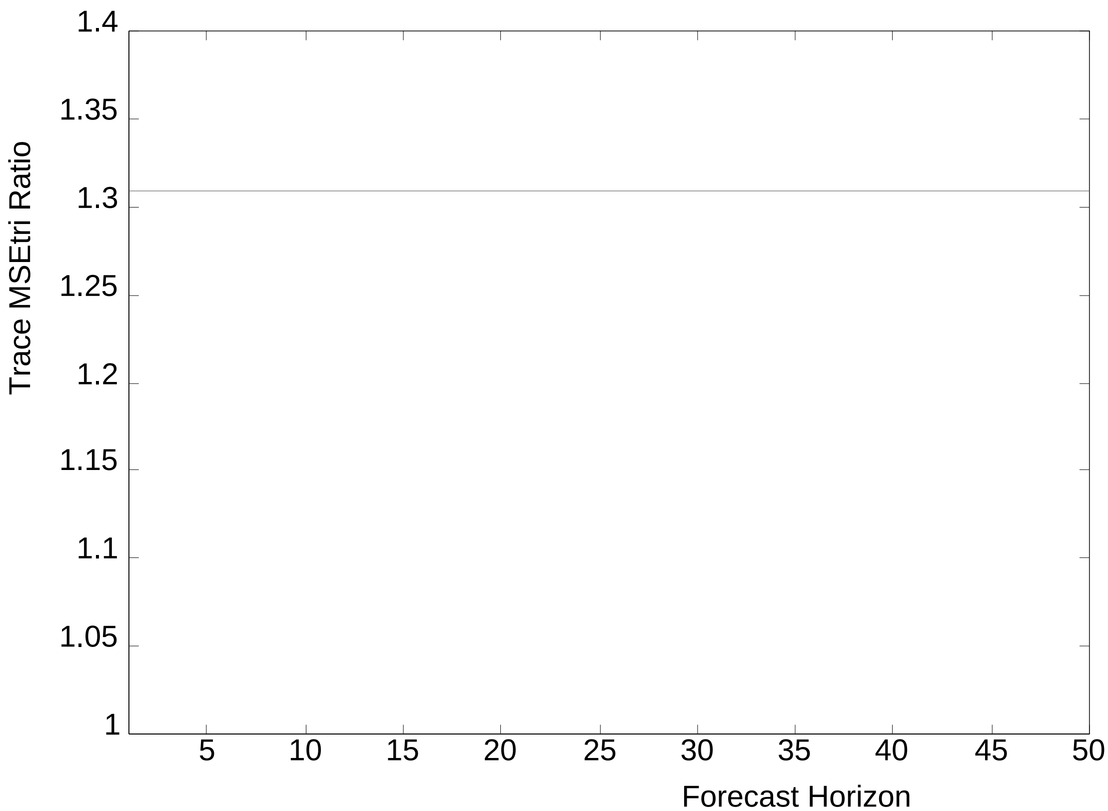
Zivot, E. (1996), "The Power of Single Equation Tests for Cointegration when the Cointegrating Vector is Prespecified," Manuscript, Department of Economics, University of Washington.

Figure 1
Trace MSE Ratio
Univariate vs. System Forecasts
Bivariate Example, λ=q=1



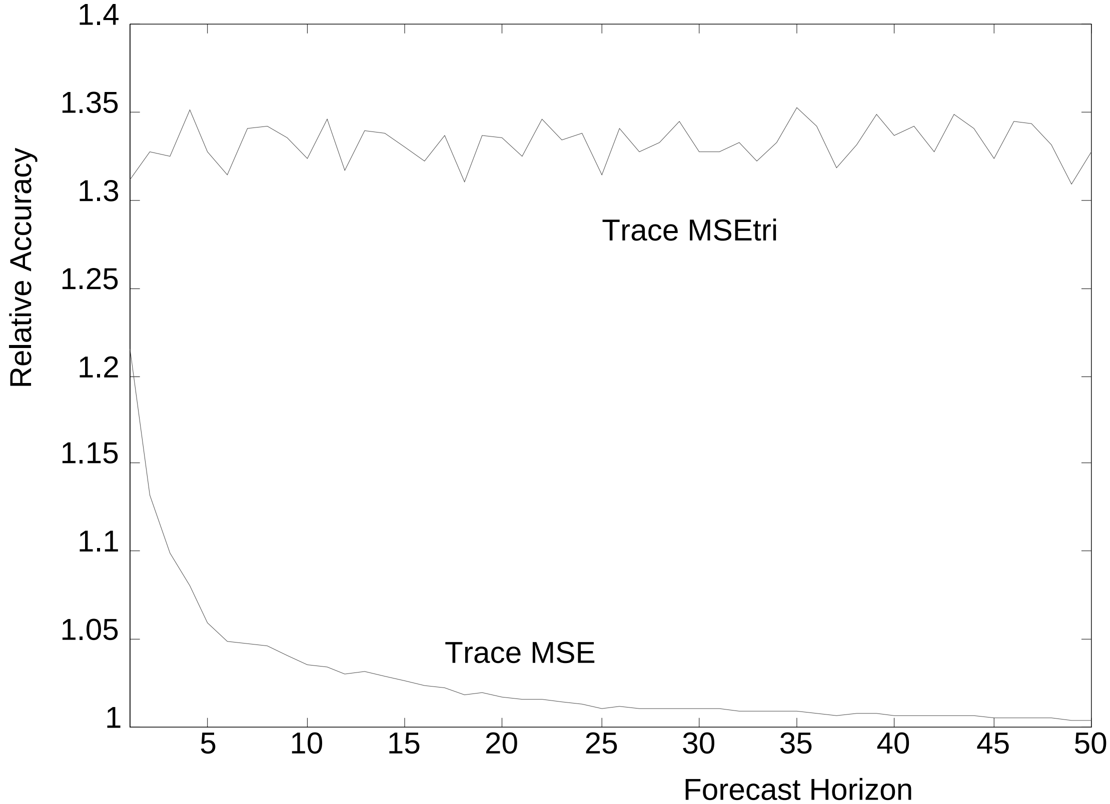
Notes to Figure: We plot the trace MSE ratio (univariate / cointegrated system) against the forecast horizon.

Figure 2
Trace MSE_{tri} Ratio
Univariate vs. System Forecasts
Bivariate Example, λ=q=1



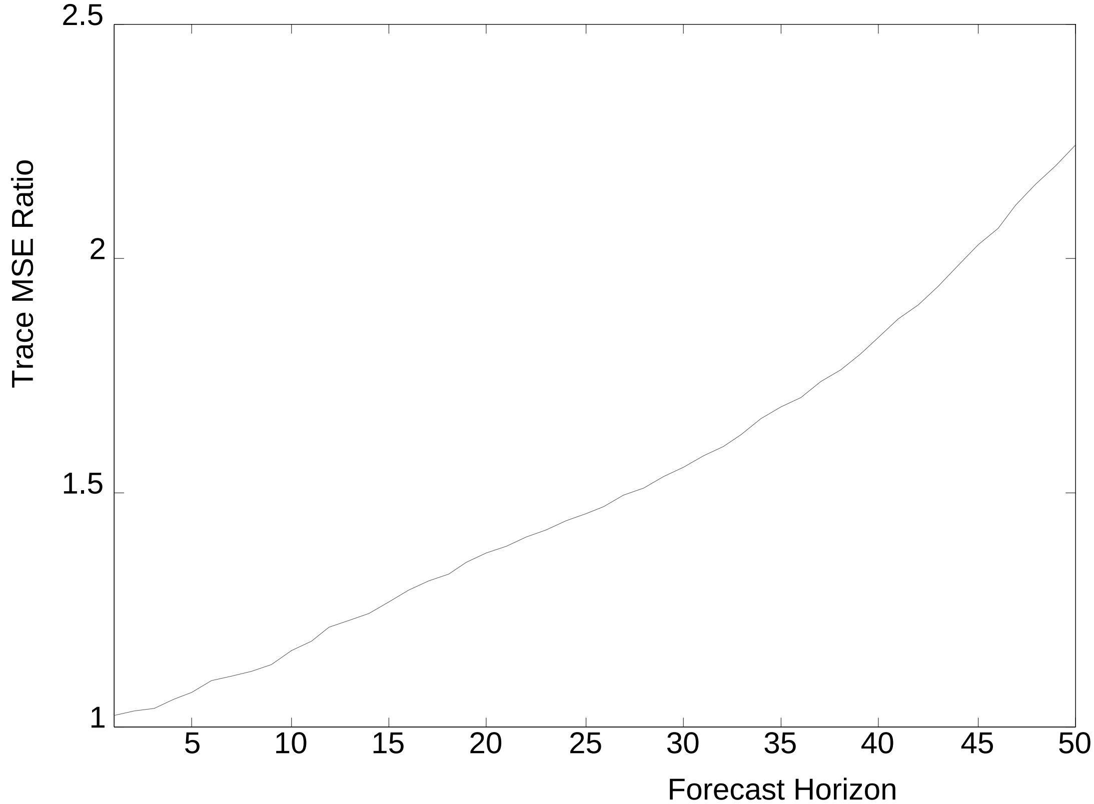
Notes to Figure: We plot the trace MSE_{tri} ratio (univariate / cointegrated system) against the forecast horizon.

Figure 3
Trace MSE Ratio and Trace MSE_{tri} Ratio
Univariate vs. System Forecasts, Estimated Parameters
Bivariate Example, λ=q=1



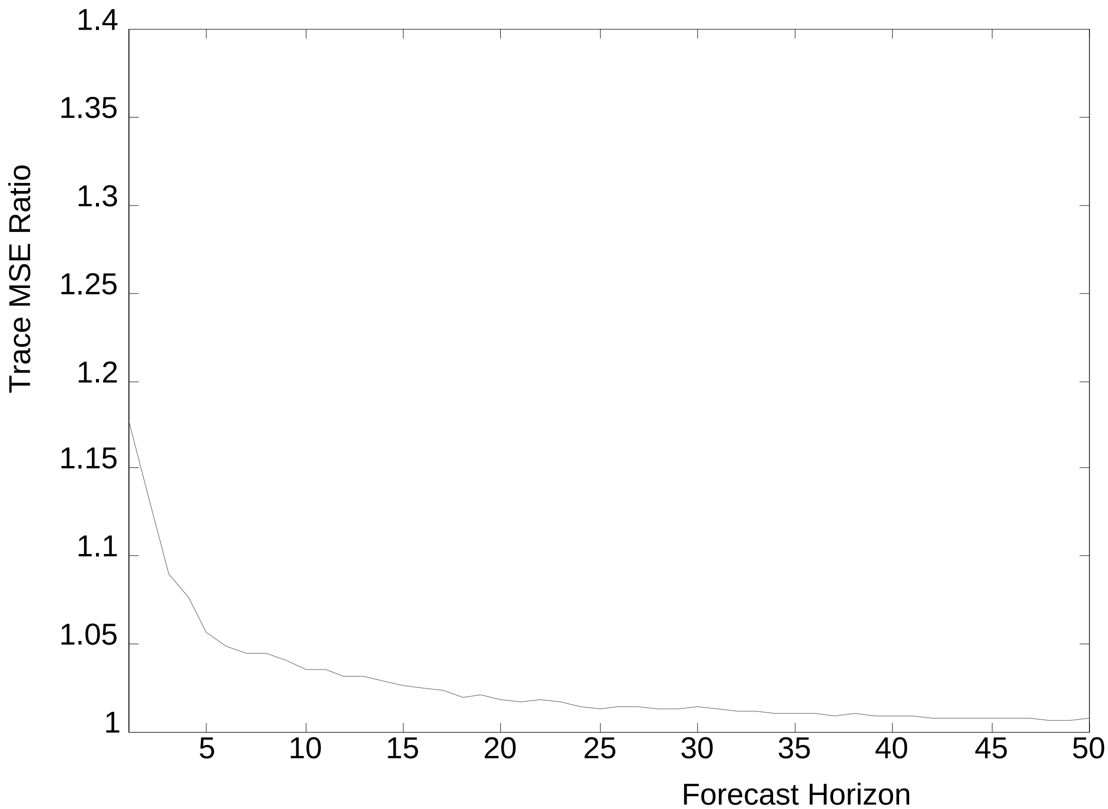
Notes to Figure: We plot the trace MSE ratio and the trace MSE_{tri} ratio (univariate / cointegrated system) against the forecast horizon

Figure 4
Trace MSE Ratio
Levels VAR vs. Cointegrated System Forecasts, Estimated Parameters
Bivariate Example, λ=q=1



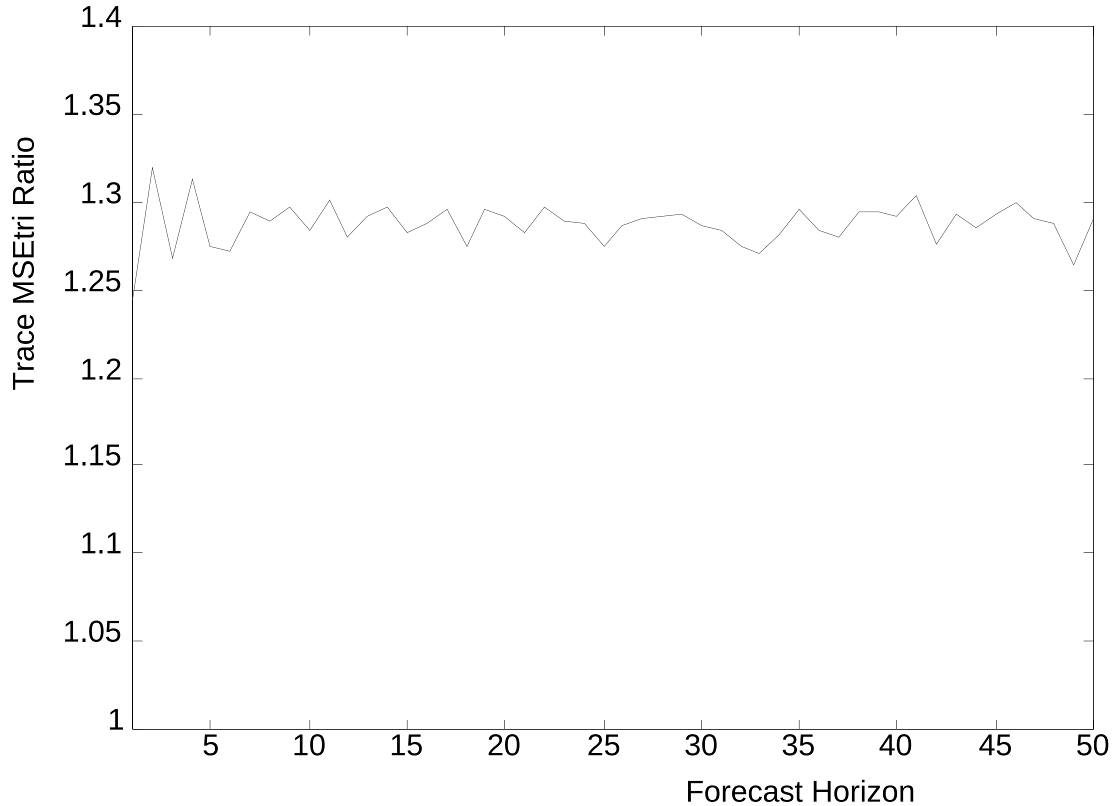
Notes to Figure: We plot the trace MSE ratio for (VAR in levels / cointegrated system) against the forecast horizon.

Figure 5
Trace MSE Ratio
Differenced VAR vs. Cointegrated System Forecasts, Estimated Parameters
Bivariate Example, λ=q=1



Notes to Figure: We plot the trace MSE ratio (VAR in differences / cointegrated system) against the forecast horizon.

Figure 6
Trace MSE_{tri} Ratio
Differenced VAR vs. Cointegrated System Forecasts, Estimated Parameters
Bivariate Example, λ=q=1



Notes to Figure: We plot the trace MSE_{tri} ratio (VAR in differences / cointegrated system) against the forecast horizon.