Conditional Inference in Nearly Cointegrated Vector Error-Correction Models with Small Signal-to-Noise Ratio*

Nikolay Gospodinov[†]
Federal Reserve Bank of Atlanta

Alex Maynard[‡]
University of Guelph

Elena Pesavento[§]
Emory University

October 2017

Abstract

This paper studies inference in cointegrating regessions and vector error correction (VEC) models when the cointegrating errors are a nearly integrated process with a low signal-to-noise-ratio. This combination of persistent, yet low variance error terms characterizes most "carry" regressions of different asset classes, including for example exchange rates and gold futures prices. We develop rates of convergence and asymptotic distributions for estimators of the VEC and conditional VEC model when the error correction term is characterized by a dampened near unit root process, thereby combining the literatures on near cointegration and near zero variance

Abstract

regressors. We find that the estimator in the conventional VEC model suffer from a reduced rate of convergence, a large bias, and a highly dispersed asymptotic distribution. Its conditional counterpart is found to have better asymptotic properties.

JEL Classification: C12, C15, C22.

Keywords: Cointegration; Vector error correction models; Local-to-unity asymptotics; Local-to-zero variance.

---

*The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

[†]Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. N.E., Atlanta, Georgia, 30309–4470, USA; Email: nikolay.gospodinov@atl.frb.org.

[‡]Department of Economics, University of Guelph, Guelph, Ontario, N1G 2W1, Canada; Email: maynarda@uoguelph.ca.

[§]Department of Economics, Emory University, Atlanta, Georgia, 30322–2240, USA; Email: epesave@emory.edu.

1 Introduction and Motivation

This paper considers some important inference issues that arise in the analysis of nearly cointegrated processes is the presence of highly persistent cointegrating errors whose variability is only a small fraction of the variance of the original variables. Equivalently, in the vector error correction (VEC) representation of the cointegrated system, the error correction term is near-integrated with low signal-to-noise ratio. Typical examples of this setup include models that study the unbiasedness of forward and futures prices (exchange rates, interest rates, commodity prices) for the expected future spot values. For instance, while spot and forward exchange rates trace each other very closely, a formal unit root test on their difference (forward premium) often cannot reject the null hypothesis of a lack of cointegration. The heuristic reason for this is that the forward premium has a tiny variance compared to the variability of the individual variables and prevents its near random walk component to force spot and forward rates to drift apart in the long run. To illustrate this phenomenon, Figure 1 plots the dynamics of the spot and forward exchange rate for British Pound / US dollar along with the forward premium. The left graphs illustrates the extremely small signal-to-noise ratio while the right graph visualizes the high persistence in the forward premium which, tested in isolation for a unit root, would suggest that the spot and forward rates are not cointegrated. This problem is not an isolated problem specific to exchange rates. The same feature is presented in Figure 2 for the gold price which gives rise to an unbalanced regression when the difference between the spot and forward (futures) price is used as a regressor. In fact, this appears to be a common problem for most "carry" regressions of different asset classes (Koijen, Moskowitz, Pedersen and Vrugt, 2013).

The paper derives the theoretical implications of the simultaneous presence of high persistence, low variability and endogeneity of the cointegrating errors for the concept of cointegration, the properties of cointegrating regressions, estimation and testing in vector error-correction models, power of cointegration tests etc. More specifically, we develop the appropriate theory (rate of convergence and asymptotic distributions) for the estimators in VEC and conditional VEC models (Phillips, 1991; Johansen, 1992; Boswijk, 1994) when the error correction term is parameterized as a dampened near unit root process (local-to-unity process with local-to-zero variance). In doing this, we combine the literatures on near cointegration (Zivot, 2000; Jansson and Haldrup, 2002; Pesavento, 2004; Elliott, Jansson and Pesavento, 2005) and near zero variance regressors (Torous and Valkanov, 2000; Moon, Robia and Valkanov, 2004; Gospodinov, 2009; Deng, 2014). This double local parameterization of the persistence and variance of the cointegration errors provides a powerful tool for deriving limiting results by capturing the salient features of the data in the empirical examples. One important result that emerges from our analysis is that the estimator in the conventional VEC models is characterized by large bias, reduced rate of convergence and highly dispersed asymptotic distribution while its conditional counterpart enjoys a substantially improved asymptotic behavior. The paper provides a detailed investigation of the numerical properties of the estimators in unconditional and conditional VEC models and the empirical size and power of tests for cointegration based on the corresponding test statistics. The practical importance of the analytical results is demonstrated in the context of exchange rate models.

2 Model

Suppose that the vector $z_t = (y_t, x_t')'$ is generated by the triangular system

$$\begin{aligned} y_t &= \mu_y + \tau_y t + \gamma' x_t + u_{y,t} \\ x_t &= \mu_x + \tau_x t + u_{x,t} \end{aligned}$$

and

$$A(L)\begin{pmatrix} (1-\rho L)u_{y,t} \\ (1-L)u_{x,t} \end{pmatrix} = \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix}.$$

Assumption A. *Assume that $A(L) = I - \sum_{i=1}^{p} A_i L^i$ is a p^{th}-order matrix polynomial with roots that lie outside the unit circle and $(\varepsilon_{y,t},\ \varepsilon_{x,t})' \sim iid(0, \Sigma)$.*

For presentational simplicity, we further assume that $A(L) = I$, x_t is a scalar and the deterministic terms are ignored. Let $\{B_1(r), B_2(r) : r \in [0,1]\}$ denote a bivariate Brownian motion with correlation δ.

Assumption B. *Assume that $u_{y,t}$ is generated by*

$$u_{y,t} = \phi_T u_{y,t-1} + \tau_T \varepsilon_{y,t}, \tag{1}$$

where u_0 is $o_p(T^{1/2})$, $\phi_T = 1 + c/T$ for some fixed constant $c \leq 0$ and $\tau_T = \lambda/\sqrt{T}$ for some fixed constant $\lambda > 0$.

Assumption B follows Gospodinov (2009) and reparameterizes ϕ_T and τ_T as local-to-unity and local-to-zero sequences to account for possibility of high persistent errors (near-cointegration) and low (local-to-zero) signal-to-noise ratio. The normalization factors T and $T^{1/2}$ for the local-to-unity and local-to-zero parameterizations are chosen to match the asymptotics of the estimators of ϕ_T and τ_T. The local-to-zero parameterization has been used in a predictive regression framework by

Torous and Valkanov (2000), Moon, Rubia and Valkanov (2004) and Deng (2014). In a different context, Ng and Perron (1997) adopt a similar parameterization to study the effect of low signal-to-noise ratio of the regressor on the sampling properties of cointegrating vector estimators.

The dual localization proves to be instrumental in producing a process that is stochastically bounded and hence consistent with both statistical and economic theory. Unlike regular near-unit root processes that are of order $O_p(T^{1/2})$, the local-to-zero variance localization dampens the stochastic trend behavior of $u_{y,t}$ and keeps it stochastically bounded ($O_p(1)$). More specifically,

$$\begin{aligned} u_{y,t} &= \lambda T^{-1/2} \sum_{i=1}^{t} (1 + c/T)^{t-i} v_i \\ &\Rightarrow \lambda J_c(r), \end{aligned}$$

where $J_c(r) = \exp(cr) \int_0^r \exp(-cs) dB_2(s)$; i.e., $u_{y,t}$ converges weakly to an Ornstein-Uhlenbeck process without any normalization that depends on the sample size. The dual localization removes the economically unappealing possibility that the errors $u_{y,t}$ can wander off and preserves the cointegration between y_t and x_t. Also, the error-correction representation is now balanced as both the dependent variable $\triangle y_t$ and the regressor $u_{y,t-1}$ are stochastically bounded.

3 Main Results

Using a control variable approach, an efficient estimator of γ can be obtained from the regression

$$y_t = \gamma x_t + \omega \triangle x_t + e_t,$$

where ω is the regression coefficient of $u_{y,t}$ on $u_{x,t}$ and e_t are the residuals from this regression. Note that the OLS estimator of γ from this regression is equivalent to the MLE (Phillips, 1991)

The estimator $\hat{\gamma}$ is asymptotically distributed as

$$\sqrt{T} (\hat{\gamma} - \gamma_0) \Rightarrow \lambda \frac{\int_0^1 J_c(s) B_1(s) ds}{\int_0^1 B_1(s)^2 ds}.$$

Interestingly, (i) the estimator $\hat{\gamma}$ is consistent but has a slower rate of convergence and (ii) the conventional t-statistic of $H_0 : \gamma = \gamma_0$ diverges at rate $T^{1/2}$ as in spurious regressions.

Now consider the VEC representation given by

$$\begin{pmatrix} \triangle y_t \\ \triangle x_t \end{pmatrix} = (\rho - 1) \begin{pmatrix} 1 & -\gamma \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix}.$$

The single-equation conditional VECM has the form

$$\triangle y_t = \beta u_{y,t-1} + \varphi \triangle x_t + \varepsilon_{y,t},$$

where $\beta = \rho - 1$. In this model, the test for cointegration is $H_0 : \beta = 0$.

The asymptotic behavior of the estimator $\tilde{\beta}$ is the same as in Hansen (1995) and Zivot (2000)

$$T\left(\tilde{\beta} - \beta\right) \Rightarrow \delta \frac{\int_0^1 J_c(s) dW_1(s)}{\int_0^1 J_c(s)^2 ds} + (1-\delta)^{1/2} \frac{\int_0^1 J_c(s) dB_2(s)}{\int_0^1 J_c(s)^2 ds}$$

and

$$t_{\tilde{\beta}} \Rightarrow \delta z + (1-\delta)^{1/2} \frac{\int_0^1 J_c(s) dB_2(s)}{\left(\int_0^1 J_c(s)^2 ds\right)^{1/2}},$$

where z is a standard normal random variable and W_1 is a standard Brownian motion independent of B_2. Importantly, the limiting distributions of the estimator and the t-statistic do not depend on the signal-to-noise ratio through the localizing constant λ.

It is often the case that the VECM is defined (for predictive purposes, for instance) as

$$\triangle y_t = \beta u_{y,t-1} + \xi_t,$$

where $\beta = \rho - 1$ and $\xi_t = \varphi \triangle x_t + \varepsilon_{y,t}$. In this case,

$$\sqrt{T}\left(\hat{\beta} - \beta\right) \Rightarrow \frac{1}{\lambda}\left[\delta \frac{\int_0^1 J_c(s) dB_2(s)}{\int_0^1 J_c(s)^2 ds} + (1-\delta)^{1/2} \frac{\int_0^1 J_c(s) dW_1(s)}{\int_0^1 J_c(s)^2 ds}\right]$$

and

$$t_{\hat{\beta}} \Rightarrow \delta \frac{\int_0^1 J_c(s) dB_2(s)}{\left(\int_0^1 J_c(s)^2 ds\right)^{1/2}} + (1-\delta^2)^{1/2} z.$$

Unlike the estimator $\tilde{\beta}$ in the conditional VECM, the estimator $\hat{\beta}$ has a slower rate of convergence and its limiting distribution depends inversely on λ so that low values of λ make the estimator highly volatile.

In summary, the results for VECMs suggest that there are different rates of convergence (T versus root-T) for the conditional and unconditional VECM. The asymptotic analysis shows that the variance of the estimator in unconditional VECM could be much larger than its conditional counterpart. When $\delta \to 0$, the asymptotic distribution of the t-statistic in unconditional VECM approaches the standard normal distribution while in conditional VECM, it approaches the DF-type distribution. The opposite is true when $\delta \to 1$.

4 Simulation Results

References

[1] Boswijk, H. P., (1994), "Testing for an unstable root in conditional and structural error correction models," *Journal of Econometrics* 63, 37–60.

[2] Deng, A. (2014), "Understanding spurious regression in financial economics," *Journal of Financial Econometrics* 12, 122–150.

[3] Elliott, G, M. Jansson, and E. Pesavento (2005), "Optimal power for testing potential cointegrating vectors with known parameters for nonstationarity," *Journal of Business and Economic Statistics* 23, 34–48.

[4] Gospodinov, N. (2009), "A new look at the forward premium puzzle," *Journal of Financial Econometrics* 7, 312–338.

[5] Hansen, B. E. (1995), "Rethinking the univariate approach to unit root tests: How to use covariates to increase power," *Econometric Theory* 11, 1148–1171.

[6] Jansson, M., and N. Haldrup, (2002), "Regression theory for nearlycointegrated time series," *Econometric Theory* 18, 1309–1335.

[7] Johansen, S., (1992), "Cointegration in partial systems and the efficiency of single-equation analysis," *Journal of Econometrics* 52, 389–402.

[8] Koijen, R., T. Moskowitz, L. H. Pedersen, and E. Vrugt, (2013), "Carry," Fama-Miller Working Paper.

[9] Maynard, A., and P. C. B. Phillips. (2001), "Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly." *Journal of Applied Econometrics* 16, 671–708.

[10] Moon, R., A. Rubia, and R. Valkanov. (2004). "Long-horizon regressions when the predictor is slowly varying." Unpublished manuscript, University of Southern California.

[11] Ng, S., and P. Perron. (1997), "Estimation and inference in nearly unbalanced nearly cointegrated systems." *Journal of Econometrics* 79, 53–81.

[12] Pesavento, E. (2004), "Analytical evaluation of the power of tests for the absence of cointegration, " *Journal of Econometrics* 122, 349–384.

[13] Phillips, P. C. B. (1991), "Optimal inference in cointegrated systems." *Econometrica* 59, 283–306.

[14] Phillips, P. C. B., and J. H. Lee (2013), "Predictive regression under various degrees of persistence and robust long-horizon regression," *Journal of Econometrics* 177, 250–264.

[15] Torous, W., and R. Valkanov. (2000), "Boundaries of predictability: Noisy predictive regressions." Unpublished manuscript, UCLA.

[16] Zivot, E. (2000), "The power of single equation tests for cointegration when the cointegrating vector is prespecified," *Econometric Theory* 16, 407–439.

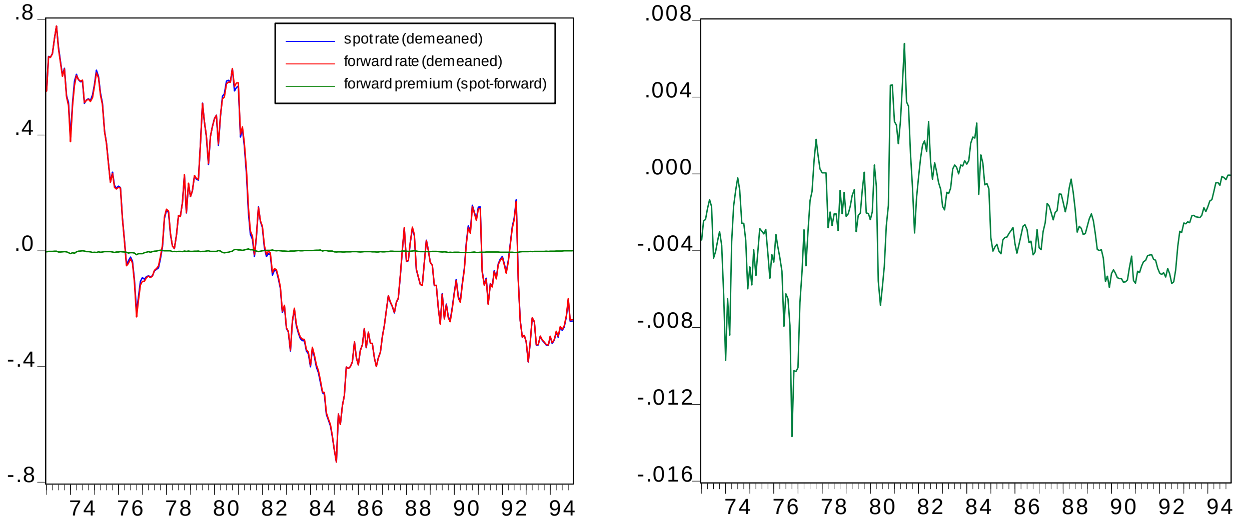

Figure 1. British pound / US dollar demeaned spot exchange rate (blue), demeaned forward exchange rate (red) and forward premium computed as forward rate - spot rate (green).

Figure 1:

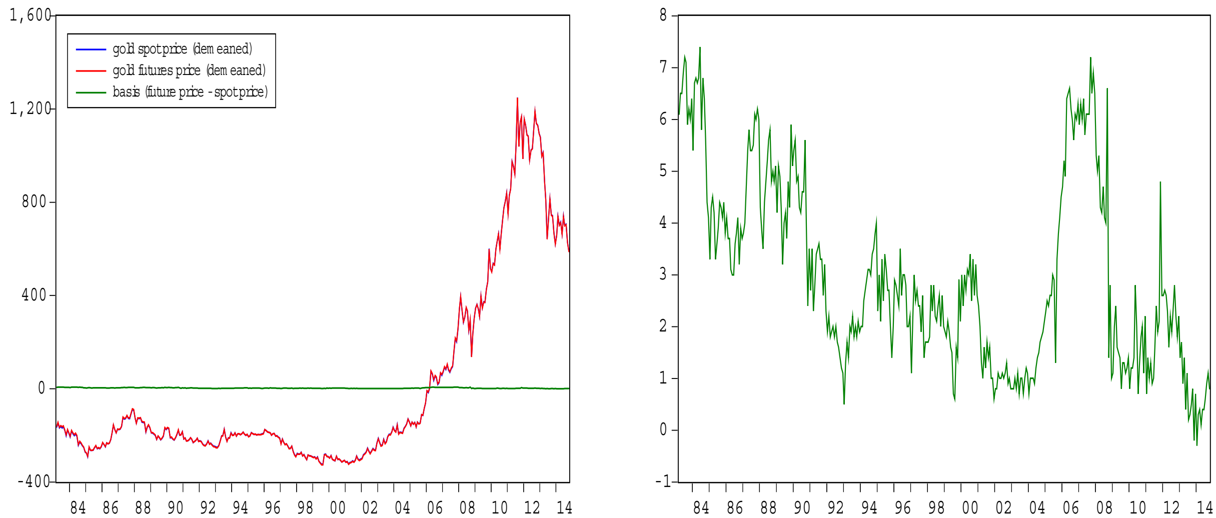

Figure 2. Gold demeaned spot price (blue), demeaned futures price (red) and basis computed as futures price - spot price (green).

Table 1. $\lambda = 0.1$, 5000 replications

	$\delta = 0$			$\delta = 0.9999$		
	bias	s.d.	t-stat	bias	s.d.	t-stat
$\rho = 1$						
$T = 200$						
UVECM	-0.028	2.351	0.054	-3.857	3.271	0.056
CVECM	-0.026	0.022	0.050	0.000	0.0002	0.052
$T = 1000$						
UVECM	-0.012	1.053	0.051	-1.728	1.454	0.052
CVECM	-0.005	0.005	0.049	0.000	0.0001	0.048
$\rho = 0.97$						
$T = 200$						
UVECM	-0.068	3.378	0.057	-3.273	3.751	0.012
CVECM	-0.023	0.027	0.143	0.000	0.0003	1.000
$T = 1000$						
UVECM	0.007	2.588	0.047	-1.225	2.748	0.007
CVECM	-0.004	0.009	0.994	0.000	0.0001	1.000

Notes: The numbers for the t-statistics are rejection rates.

Table 2. $\lambda = 0.05$, 5000 replications

	$\delta = 0$			$\delta = 0.9999$		
	bias	s.d.	t-stat	bias	s.d.	t-stat
$\rho = 1$						
$T = 200$						
UVECM	-0.053	4.645	0.053	-7.660	6.529	0.054
CVECM	-0.026	0.022	0.050	0.000	0.0002	0.057
$T = 1000$						
UVECM	-0.009	2.097	0.050	-3.372	2.815	0.049
CVECM	-0.005	0.005	0.049	0.000	0.0001	0.055
$\rho = 0.97$						
$T = 200$						
UVECM	0.035	6.734	0.053	-6.590	7.690	0.014
CVECM	-0.024	0.028	0.154	0.000	0.0003	1.000
$T = 1000$						
UVECM	0.054	5.239	0.050	-2.585	5.531	0.007
CVECM	-0.004	0.009	0.994	0.000	0.0001	1.000

Notes: The numbers for the t-statistics are rejection rates.