A Re-examination of the Fragility of Evidence from Cointegration-Based Tests of Foreign Exchange Market Efficiency

Authors: John Barkoulas, Christopher Baum

This work is posted on eScholarship@BC, Boston College University Libraries.

Boston College Working Papers in Economics, 1996

Originally posted on: http://ideas.repec.org/p/boc/bocoec/311.html
A Re-examination of the Fragility of Evidence from Cointegration-Based Tests of Foreign Exchange Market Efficiency

by

John Barkoulas *

Department of Economics, P.O. Box 6025, West Virginia University

Morgantown, WV 26505

and

Christopher F. Baum

Department of Economics, Boston College, Chestnut Hill, MA 02167

Abstract

We re-examine Sephton and Larsen's (1991) conclusion that cointegration-based tests for market efficiency suffer from temporal instability. We improve upon their research by i) including a drift term in the vector error correction model (VECM) in the Johansen procedure, ii) correcting the likelihood ratio test statistic for finite-sample bias, and iii) fitting the model over longer data sets. We show that instability of the Johansen cointegration tests mostly disappears after accounting for these two factors. The evidence is even more stable in favor of no cointegration when we apply our analysis to longer data sets.

* We acknowledge the helpful comments of an anonymous referee. Correspondence to Christopher F. Baum, Department of Economics, Boston College, Chestnut Hill, MA 02167, Tel.: (617) 552 3673, Fax: (617) 552 2308, e-mail: baum@bc.edu.
Introduction

Sephton and Larsen (1991), SL henceforth, showed that inference based on Johansen cointegration tests of foreign exchange market efficiency suffers from structural instability. More specifically, they showed that cointegration-based tests of foreign exchange market efficiency using the Johansen procedure are sensitive to model specification (bivariate versus multivariate analysis) as well as to the choice of the sample period under examination. After performing iterated Johansen tests over different systems of exchange rates and sample periods, SL concluded that "These results lead to the indictment of the cointegration approach to testing market efficiency ..." (p. 565).

Cointegration methodology has been extensively used as a convenient way of testing for the weak-form of asset market efficiency, which states that no asset price should be forecastable from the prices of other assets.\(^1\) The rationale derives from the Granger Representation Theorem (Engle and Granger (1987)) which, applied to a system of asset prices, states that the presence of cointegration in the system directly implies the existence of Granger-causal orderings among the cointegrated asset prices. Such orderings enable one to predict one asset price on the basis of the others, which contradicts market efficiency.

The purpose of this paper is to re-examine the evidence of temporal instability from cointegration-based tests of market efficiency using similar data sets and empirical design as in SL. We improve upon their methodology in two respects. First, we include a drift term in the vector error correction model (VECM) in the Johansen procedure, and second, we correct the Johansen trace test statistic to account for finite-sample bias. Both considerations are important in drawing valid inferences. To strengthen our evidence, we apply our analysis to longer data sets than those utilized by SL. We find that inference based on the Johansen
cointegration tests of foreign exchange market efficiency is far more stable than suggested by SL, especially over longer sample periods. The evidence points toward absence of cointegration among major foreign exchange rates.

The next section presents the Johansen cointegration method and elaborates on our suggested improvements. The empirical results follow. Final remarks and suggestions conclude the paper.

I. Econometric Methodology

I.A. The Johansen Cointegration Method

Following SL we employ the Johansen cointegration method (Johansen (1988), Johansen and Juselius (1990)) to determine the existence of common trends in systems of currency spot and forward rates. We first briefly describe the Johansen cointegration procedure.

Without any loss of generality, a \( p \)-dimensional vector autoregressive (VAR) process of \( k \)-th order can be written as follows

\[
\Delta X_t = \Theta_1 \Delta X_{t-1} + \ldots + \Theta_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + \varepsilon_t
\]

where \( \Delta \) is the first-difference lag operator, \( X_t \) is a \((p \times 1)\) random vector of time-series variables with order of integration of at most one denoted by \( I(1) \), \( \varepsilon_t \) is a sequence of zero-mean \( p \)-dimensional white noise vectors, \( \Theta_i \) are \((p \times p)\) matrices of parameters, and \( \Pi \) is a \((p \times p)\) matrix of parameters the rank of which contains information about long-run relationships among the variables in the VAR.

Expression (1) is referred to as the vector error correction model (VECM). If \( \Pi \) has full rank \( p \), all elements in \( X_t \) are stationary. If the rank of \( \Pi \) is zero, the model
reduces to a VAR in first-differences. The interesting case occurs when $0 < r < p$ which suggests the existence of $r$ cointegrating relationships. In this case there exist $(p \times r)$ matrices $\alpha$ and $\beta$ such that $\Pi = \alpha \beta'$. $\beta$ is the matrix of cointegrating vectors and has the property that $\beta' X_t$ is stationary even though $X_t$ may be individually $I(1)$ processes.

To test the hypothesis that the number of cointegrating vectors is at most $r$, the trace statistic is calculated

$$\text{tr}(r) = - T \sum_{i=r+1}^{p} \ln(1 - \hat{\lambda}_i)$$

where $\hat{\lambda}_{r+1}, \ldots, \hat{\lambda}_p$ are the $(p-r)$ smallest eigenvalues to the generalized eigenvalue problem

$$|\lambda_k S_{kk} - S_{k0} S_{00}^{-1} S_{0k}| = 0$$

The $S_{ij}$ matrices are residual moment matrices from the VECM in (1). The asymptotic distribution for the trace test statistic is non-standard and depends only on $(p-r)$. Critical values obtained from Monte Carlo simulations of the limiting distribution are given in Johansen and Juselius (1990) and Osterwald-Lenum (1992).

I.B. Suggested Improvements

The first suggestion we propose relates to the VECM used. Instead of the VECM in (1) estimated in SL, we propose to estimate the following VECM

$$\Delta X_t = \mu + \Theta_1 \Delta X_{t-1} + \ldots + \Theta_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + \varepsilon_t$$
where $\mu$ is a $(p \times 1)$ matrix of constants.

The VECM in (4) differs from SL's in that it allows for a drift term $\mu$ in the VAR process. The drift term should be included unless a strong prior exists for its exclusion. Such a prior may not be justified for exchange rates and estimation of the inappropriate VECM could potentially lead to erroneous inference. Diebold, Gardeazabal, and Yilmaz (1994) also used the VECM in (4) to analyze the evidence by Baillie and Bollerslev (1989) on spot foreign exchange market efficiency, which they were able to overturn. They argued that "... it is now generally agreed that drift should be included, unless there is irrefutable prior information to the contrary" (p. 6). However, Diebold, Gardeazabal, and Yilmaz’s investigation was focused on a particular sample period and did not address the issue of temporal instability, which is the primary objective of our analysis. Baillie and Bollerslev (1994) also argued for the inclusion of a drift term in a regression with seven currency spot rates.

The second suggestion we propose deals with the finite-sample bias in the Johansen test statistics. Empirical evidence by Cheung and Lai (1993) points to significant finite-sample bias in the performance of the Johansen test statistics when asymptotic critical values are used for inference in finite samples. They showed that "...tests are biased toward finding cointegration too often when asymptotic critical values are used" (p. 319). To correct for the finite-sample bias toward over-rejection of the no cointegration hypothesis (spurious cointegration), Cheung and Lai proposed to estimate finite-sample critical values for the Johansen tests using response surface analysis in Monte Carlo experiments. Instead of correcting the asymptotic critical values in finite samples, Reinsel and Ahn (1992) and Reimers (1991) suggest to correct the Johansen test statistics for finite-sample bias and then
compare them with their asymptotic critical values. Their finite-sample correction multiplies the Johansen test statistic by the scale factor $\frac{T-pk}{T}$. This is the approach we follow in this paper to draw inferences. The corrected test statistic will be referred to as the modified Johansen test statistic.

**II. Empirical Results**

To address the issue of temporal instability, we perform the Johansen cointegration tests on the same data set and in the same way as in SL, except for the suggested improvements. They used data on foreign exchange rates of both daily and monthly frequencies and performed cointegration tests on systems of seven, four, and two foreign exchange rates. In addition, we expand the data sets for a longer period in order to obtain further insight regarding the robustness of our results.\(^2\)

Before proceeding with the cointegration tests, it must be established that the currency rates are integrated processes of the same order. All currency rates were subjected to the Augmented Dickey-Fuller test (ADF, Dickey and Fuller (1979, 1981)) (a constant and a time trend was included in the ADF regression) and the null hypothesis of a single unit root could not be rejected at the 5% level for all rates.\(^3\) The unit-root test results as well as any subsequent results not reported here are available upon request.

In implementing the Johansen cointegration test, the critical values depend upon whether the true data generating process (DGP) contains a drift term. Since the DGP is not known a priori, we compare the test statistics to the critical values corresponding to both $\mu \neq 0$ and $\mu = 0$ in the DGP. Table 1, from Osterwald-Lenum (1992), provides the critical values for both cases. The number of lags used in the VECM in (4) is two, that is, $k = 2$, which was found to be optimal according to the multivariate Schwartz Information Criterion.\(^4,5\)
II.A. Daily Data

The first data set studied by SL is that of Baillie and Bollerslev (1989) which consists of daily observations covering the period March 1, 1980 to January 28, 1985 for the spot and 30-day forward rates for the seven major currencies: the Canadian dollar (CD), Deutsche mark (DM), British pound (BP), French franc (FF), Swiss franc (SF), Italian lira (IL), and Japanese yen (JY). We expand this data set to cover the period through October 31, 1993. We repeat the experiment conducted by SL: we perform the trace test on an initial sample of the first 80 observations and then on samples generated by adding 75 observations. Figure 1 summarizes the Johansen cointegration evidence for the no-cointegration ($r=0$) null hypothesis over different sample periods for the system of the seven spot currency rates. This and subsequent figures graph the values for the trace test statistics, with and without correction for finite-sample bias, calculated over different subsamples as well as the 5%, 2.5%, and 1% asymptotic critical values for the case of $\mu \neq 0$ in the DGP. The critical values for the case of $\mu = 0$ in the DGP are not included in the figures in order to preserve their readability. All inference hereafter is based on the values for the modified test statistic and the critical values for the case of $\mu \neq 0$ in the DGP unless otherwise indicated. The evidence for the null hypotheses of at most one ($r\leq1$) up to at most six ($r\leq6$) cointegrating vectors is not reported since the corresponding trace statistics are well below the 5% critical value across all subsamples (the null hypothesis is never rejected). At the 5% significance level, there are only three rejections of the null hypothesis of no cointegration: at the short subsamples consisting of 230 and 380 observations and the subsample of 3155 observations. This evidence against the no-cointegration hypothesis is extremely weak indeed since, at the 2.5% significance level, the no-cointegration null
hypothesis is rejected only for the subsample consisting of 380 observations. The same inference holds true when the critical values for $\mu = 0$ in the DGP are used. Clearly the evidence points toward no cointegration across time.

The same experiment was performed on the 30-day forward rates for the seven major currencies. Again the cointegration test results corresponding to the null hypotheses of at most one ($r \leq 1$) up to at most six ($r \leq 6$) cointegrating vectors are not reported since no rejection of these null hypotheses was obtained at the 5% significance level across all subsamples. Figure 2 reports the Johansen cointegration test results for the no-cointegration ($r = 0$) null hypothesis, which can only be rejected in the subsamples consisting of 305, 380, 455, 530, and 3155 observations at which we find evidence of a single cointegrating vector. However, with the exception of the subsample consisting of 3155 observations, the sizes for these subsamples are too small to base any inferences regarding long-run relationships. With the exception of the subsample of 380 observations, we fail to reject the null hypothesis of no cointegration at the 2.5% significance level. The no-cointegration null hypothesis is never rejected when the critical values for $\mu = 0$ in the DGP are used, even at the 5% significance level. The evidence strongly supports no cointegration and appears temporally stable.

Overall, the Johansen cointegration tests applied to daily spot and forward rates for the seven major currencies overwhelmingly indicate the absence of any long-run equilibrium relationship. This evidence appears to be temporally stable, contradicting that of SL to a great extent.

II.B. Monthly Data

In this section we report the results of cointegration tests on monthly foreign exchange spot rates originally used by Hakkio and Rush (1989). Our data set consists
of end-of-month U. S. dollar noon-time bid rates in the New York foreign exchange market for the seven major currencies and cover the period 1973:1 to 1993:10. Following SL we first used a four exchange rate model which included the CD, DM, BP, and JY. We estimated the trace test statistic for sample periods that begin in July 1975 and span from January 1979 to October 1993 (SL’s study covered the period until December 1988). Figures 3a to 3d present the cointegration test results. We fail to reject the no-cointegration null hypothesis in all cases except for the subsamples consisting of 57, 58, and 65 observations at which we obtain evidence of four cointegrating vectors (all spot rates are stationary). This evidence strongly favors absence of cointegration in the four-currency system since it is extremely weak and occurs at relatively short subsamples. When the tests are conducted at the 2.5% significance level there is evidence of four cointegrating vectors only at the subsample of 58 observations indicating that the evidence against no cointegration is extremely weak or non-existent. If the critical values corresponding to the case of \( \mu = 0 \) in the DGP are used the inference is no cointegration across all subsamples at the 5% significance level.

To ensure robustness of our results we repeated the experiment for the whole period 1973:1-1993:10 in the following way. We performed the test on an initial sample of the first 48 observations (1973:1-1976:12) and then repeated it on subsamples obtained by adding one additional observation until 1993:10. As Figures 4a to 4d indicate, we fail to reject the null hypothesis of no cointegration in all but fifteen cases (out of a total of 205 cases) corresponding to the relatively small subsamples of 78 to 80, 85 to 87, and 90 to 97 observations where we find evidence for four cointegrating vectors. This indicates that the four currency rates are stationary contrary to substantial empirical evidence of a unit root in these series. As Figure 4d indicates, the trace test statistic for the null hypothesis \( r \leq 3 \) clearly converges to a value well below the 5% asymptotic critical value as the sample size
increases. At the 2.5% significance level there is evidence for four cointegrating vectors only in five cases corresponding to subsamples consisting of 89 to 92 and 95 observations. If the critical values for the case of $\mu = 0$ in the DGP are used the inference is always no cointegration at the 5% significance level. These results are indicative of temporally stable evidence in favor of no cointegration in the system of four currency rates in question.

We repeated the same experiment as above for the full system, that is, for the seven exchange rate model over the 1973:1-1993:10 period. The no-cointegration null hypothesis is never rejected for any subsample providing strong evidence in support of temporally stable inference based on the Johansen cointegration method (detailed results and figures are available upon request).

Finally, we repeated SL’s experiment for the bivariate case of the DM and the BP. As Figures 5a and 5b indicate there is evidence of temporal instability in the bivariate case at the 5% level. On several occasions the inference is that one cointegrating vector exists but in most cases the inference is that two cointegrating vectors exist, that is, both series are stationary, contrary to unit root test results. The inference clearly converges towards no cointegration with sample size. The evidence is weak against the no-cointegration null hypothesis since, at the 2.5% level, it is concentrated only on certain small subsamples. One issue that must be kept in mind in this bivariate system is the omitted-variable problem. Clearly the bivariate case is a restricted framework to analyze long-run relationships in the foreign exchange markets and one should not exclude relevant variables in cointegration analysis. To obtain additional insight, we resort to a longer sample. We utilize all data available (1973:1-1993:10) and, similarly as before, we perform the trace test on an initial sample of 48 observations and on subsequent samples generated by sequentially adding one observation until all data points are exhausted. The evidence, summarized in Figures 6a and 6b, overwhelmingly favors
no cointegration except in a small number of cases where the evidence marginally supports the presence of two cointegrating vectors at the 5% significance level (never is the no-cointegration null hypothesis rejected at the 2.5% significance level). The longer sample strongly points toward the direction of temporally stable evidence in favor of no cointegration even in the bivariate case.

Overall, the cointegration test results obtained here provide temporally stable evidence which supports absence of cointegration among foreign exchange rates. Even when we obtain occasional evidence supporting cointegration, one would expect about five per cent rejections with a large number of independent test statistics, if the null hypothesis of no cointegration is true. The plots in the figures indicate much fewer rejections. However, these test statistics are correlated as they are estimated from overlapping subsamples. The occasional rejections of the no cointegration null hypothesis is consistently restricted to small subsamples. Also, the longer the sample period the more stable the evidence in favor of no cointegration, which provides with additional assurance to the validity of our inference.

Among the two improvements we suggested, namely, inclusion of a drift term in the VECM and correction of the trace statistic for finite-sample bias, the former appears to be primarily responsible for providing stable inference and overturning SL’s results. The presence of the drift term in the VECM induces significant relations between the variances in (3), affecting the value of the trace test statistic and therefore cointegration inference. When the Johansen cointegration tests are performed on different subsamples, the effect of excluding the drift term from the VECM on the variances in (3) may vary across subsamples, which could account for the fragile evidence obtained by SL. Inference is affected by the presence of the drift term and, without a strong prior that it is absent, a drift term should be included in the VECM specification. Baillie and Bollerslev (1989) did not include a
drift term in the VECM and found a single cointegrating relationship in a system of seven spot and forward currency rates. When Diebold, Gardeazabal, and Yilmaz (1994) allowed for a drift term in the VECM the cointegrating relationship disappeared. Hakkio and Rush (1989) found no evidence of a long-run relationship between the DM and BP spot rates using the Engle-Granger method with a constant term in the cointegrating regression. SL were able to overturn the findings of Hakkio and Rush (1989) using the Johansen method, but they did not include a constant term in the VECM specification. The discrepancy in their results can reasonably be attributed to the treatment of the constant term in the estimation process, and with the inclusion of the constant term the evidence is much more robust in support of absence of cointegration.

III. Conclusions

We improve upon SL’s analysis of the stability properties of the Johansen cointegration tests of foreign exchange market efficiency by i) estimating a more appropriate VECM in the Johansen procedure, and ii) correcting for finite-sample bias in the Johansen test statistic. With these methodological improvements, the instability from the cointegration tests found in SL generally disappears.

The analysis was based on data of various frequencies (daily, monthly) as well as systems of various orders (multivariate, bivariate). One interesting avenue for future research is to investigate the performance of the Johansen method when going from lower- to higher-frequency data as well as from higher- to lower-order systems. With respect to the former, research has been done for some of the residual-based tests for cointegration (Hakkio and Rush (1991)). Regarding the latter, the evidence here indicates a tendency to find cointegration more often in lower- as opposed to higher-order systems. In addition, the possibility of non-linear
cointegration which might plausibly exist in the foreign exchange markets should also be addressed. Sephton (1994) made the first step in that direction.\textsuperscript{8}
References


Notes

1. The presumed relationship between market efficiency and cointegration is not, however, unanimously accepted. Bossaerts (1988) using a Lucas-type, multiple-period general equilibrium model showed that in economies that do not exhibit separation, but that move close to separation, cointegration among asset prices emerges. Using canonical correlation analysis he found substantial evidence of cointegration among the prices of five size-based as well as industry-based portfolios of common stock. The presumed theoretical relationship between market efficiency and cointegration has also been challenged by Dwyer and Wallace (1992) who argued that there is no general equivalence between them, and that the existence or lack of cointegration among a set of asset prices is a function of the relevant model. Dwyer and Wallace illustrated their argument that cointegration may be consistent with market efficiency in the cases of spot exchange rates, interest rates, spot and forward exchange rates, and asset prices, with or without reinvestment of income flows.

The fact that cointegration among asset prices is consistent with conditions of equilibrium in competitive markets populated by rational agents does not exclude the possibility that the cointegration relationship(s) can help predict future asset returns, thus leading to profitable arbitrage opportunities. Cerchi and Havenner (1988) using the system theoretic time series procedure due to Aoki (1987a,b) found that the stock prices of five major department stores possess one common non-stationary trend. Using the in-sample parameter estimates of the trend and cycle models they established satisfactory out-of-sample performance on the basis of both conventional forecasting measures and a non-parametric direction of change test. More importantly, Cerchi and Havenner constructed a new asset as a linear combination of the existing assets and, based on the price forecasts of the new asset,
they showed statistically significant positive profits. Cerchi and Havenner concluded that "...the forecasts are informative, certainly large enough for the forecasts to have practical value and larger than might be expected for such difficult series" (p. 345). This evidence clearly suggests that the presence of cointegrating relationships among asset prices can lead to significant forecasting improvements.

2. During the period(s) examined several events occurred in the foreign exchange markets, for example, central bank interventions and the formation of the European Monetary System (EMS) which might have impacted the long-term behavior of foreign exchange rates in certain ways. A systematic investigation of the long-term effects of such events on foreign exchange rates is beyond the scope of this paper.

3. The unit-root and cointegration tests were performed on the logs of the original series.

4. Using Monte Carlo simulations, Cheung and Lai (1993) showed that for autoregressive processes standard selection criteria, like the Schwartz Information Criterion (SIC) and Akaike Information Criterion (AIC), can be useful for selecting the correct lag structure for the Johansen tests. They found that the SIC performs slightly better than the AIC.

5. Alternative lag structures were used but the results were insensitive to the particular choice of lag length. More importantly, at this lag order the residual vectors from the system equations are serially uncorrelated, which is a critical assumption in the Johansen methodology.

6. A concern regarding this multi-stage testing process is the effective levels of type I errors. Operating in a cointegration framework and given the dependence of the sequential tests, this is a very difficult issue to resolve. We thank an anonymous referee for pointing out this important issue.
7. The finite-sample correction is more important in small samples as with large samples the modified and unmodified trace test statistics are very close.

Table 1: Asymptotic Critical Values for the Johansen Trace Statistic

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>( \mu \neq 0 ) in the DGP</th>
<th>( \mu = 0 ) in the DGP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r \leq 6 )</td>
<td>3.76 4.95 6.65</td>
<td>8.18 9.72 11.65</td>
</tr>
<tr>
<td>( r \leq 5 )</td>
<td>15.41 17.52 20.04</td>
<td>17.95 20.08 23.52</td>
</tr>
<tr>
<td>( r \leq 4 )</td>
<td>29.68 32.56 35.65</td>
<td>31.52 34.48 37.22</td>
</tr>
<tr>
<td>( r \leq 3 )</td>
<td>47.21 50.35 54.46</td>
<td>48.28 51.54 55.43</td>
</tr>
<tr>
<td>( r \leq 2 )</td>
<td>68.52 71.80 76.07</td>
<td>70.60 74.04 78.87</td>
</tr>
<tr>
<td>( r \leq 1 )</td>
<td>94.15 98.33 103.18</td>
<td>85.18 99.32 104.20</td>
</tr>
<tr>
<td>( r = 0 )</td>
<td>124.24 128.45 133.57</td>
<td>124.25 129.75 136.06</td>
</tr>
</tbody>
</table>

The alternative hypothesis is that the impact matrix \( \Pi = \alpha \beta' \) is full rank, that is, \( r = 7 \). The critical values were obtained from Tables 1 and 1.1* in Osterwald-Lenum (1992).
Figure 1: Johansen Cointegration Test Results for the Seven Currency Model with Daily Spot Data (01/01/1980-10/31/1993)

Figure 2: Johansen Cointegration Test Results for the Seven Currency Model with Daily Forward Data (01/01/1980-10/31/1993)
Figure 5: Johansen Cointegration Test Results for the Bivariate (DM vs BP) Currency Model with Monthly Spot Data (1975:7-1993:10)

(a) Null Hypothesis: 0 Cointegrating Vectors

(b) Null Hypothesis: At Most 1 Cointegrating Vectors
Figure 6: Johansen Cointegration Test Results for the Bivariate (DM vs BP) Currency Model with Monthly Spot Data for the Longer Sample (1973:1-1993:10)

(a) Null Hypothesis: 0 Cointegrating Vectors

(b) Null Hypothesis: At Most 1 Cointegrating Vectors
Note: for technical reasons Figures 3 and 4 have not been included in this downloadable copy. Please request them from the authors at baum@bc.edu.