Stability of the S&P 500 Futures Market Efficiency Conditions

William J. Crowder\textsuperscript{1}
Associate Professor
Department of Economics
University of Texas at Arlington
Arlington, TX 76019
crowder@uta.edu

January 12, 2004

\textsuperscript{1} Chanwit Phengpis provided valuable research assistance
Abstract

Brenner and Kroner (1995) laid out the necessary conditions for futures market efficiency when the asset price data are characterized by stochastic trends. Specifically, a no arbitrage profit condition implies that spot, futures and cost-of-carry will be cointegrated, unless the cost-of-carry is stationary, in which case spot and futures will be cointegrated. We examine this for the S&P 500 futures market. Our results are intriguing since we initially find evidence that spot and futures are themselves cointegrated. But a deeper analysis demonstrates that this cointegrating relationship is not stable. However, including the three-month Treasury bill rate as a proxy for the cost-of-carry yields one stable cointegrating (or equilibrium) relationship. This suggests that the evidence of cointegration between spot and futures alone is spurious and that researchers need to be careful about conclusions drawn from cointegration analysis of market efficiency conditions.
1 Introduction

Asset markets, like commodity, equity and foreign exchange markets, represent an important link between theory and the real world in financial economics. Because asset markets generally represent markets that are as close to perfectly competitive as real world markets can be, the efficiency of asset markets is critical to the efficacy of rational financial economic theory. Fama (1970) delineates the conditions that must hold for varying degrees of informational efficiency. At the most basic level, financial theory tells us that rational economic agents should not ignore risk-free profit opportunities. As such, the no-arbitrage-profit condition represents the easiest hurdle one might expect theory to encounter in the real world. A slightly more stringent test of market efficiency is the unbiasedness hypothesis. This is a weak form efficiency hypothesis where asset prices should not be predictable based on their own past history.

There is a rather large literature that tests the weak form efficiency of futures and forward markets in commodities, foreign exchange and securities. The use of unit root and cointegration techniques is ubiquitous in this literature due to the fact that asset price data are characterized by stochastic trends. Brenner and Kroner (1995) detail the relevant issues involved when using cointegration techniques to test the efficiency of futures markets. An important consideration is the appropriate specification of the cointegrating relationship. The simple efficiency or unbiasedness hypothesis assumes that the futures price is an unbiased predictor of the future spot price, implying that spot and futures prices share a one-to-one long-run equilibrium. In markets like forward foreign exchange where the no-arbitrage profit condition entails an interest rate differential, via the covered interest parity condition, this simple efficiency model may be appropriate if the differential is stationary. But for other asset markets, like commodity futures, which entails a single variable measuring the cost-of-carry, the appropriate efficiency hypothesis is based on a no-arbitrage-profit condition. As Brenner
and Kroner demonstrate, spot and futures prices will only be cointegrated if the cost-of-carry is stationary. If this condition is violated, i.e. the cost-of-carry is non-stationary, then the no-arbitrage-profit condition links the futures rate, spot rate and cots-of-carry measure in a cointegrating relationship.

Brenner and Kroner (1995) review the literature on futures markets and find very little evidence favoring simple market efficiency. They explain this by demonstrating that these studies estimate misspecified models when the cost-of-carry measure is non-stationary. Brenner and Kroner make no distinction in their analysis between storable and non-storable commodities. However, Yang et al. (2001) claim that non-storable commodities will not entail cost-of-carry expenses and the appropriate specification for such markets is based on the simple unbiasedness hypothesis. They do not examine this distinction in their own empirical analysis.

In contrast to Yang et al. (2001), Crowder and Hamed (1993) find that oil futures markets can be characterized by the unbiasedness hypothesis and that storage costs or convenience yields are not important in the forward pricing role of this futures market. More recently papers by Beck (1994), Pizzi et al. (1998) and Chu et al. (1999) all find evidence of the simple efficiency hypothesis for various storable and non-storable commodity futures markets. Pizzi et al. and Chu et al. both find evidence that the S&P 500 futures price is an unbiased predictor of the future cash price with no role played by cost-of-carry measures. However, Ackert and Racine (1999) find evidence consistent with both the unbiasedness hypothesis and the no-arbitrage-profit condition for futures market efficiency as laid out by Brenner and Kroner (1995).\footnote{Although not mentioned in their paper, Ackert and Racine (1999) find that S&P 500 spot and futures contracts are cointegrated at the 10\% level (they use a 1\% level test) with a cointegrating vector insignificantly different from \([1, -1]\) which supports the unbiasedness hypothesis.} Since both specifications cannot be correct, there is a discrepancy that has yet to be investigated in the literature.
In this study we examine the S&P 500 futures market efficiency first by examining the simple efficiency or unbiasedness specification. We find that initial results imply this specification is consistent with the data. Specifically we find one cointegrating vector between spot and futures prices and the estimated relationship is very close to \([1, -1]'\) as suggested by the unbiasedness hypothesis. But an analysis of the stability of this relationship yields evidence of substantial instability. We then add a proxy for the cost-of-carry, the three-moth Treasury bill rate, and find evidence consistent with only one cointegrating vector between the three variables, consistent with the cost-of-carry model, and this relationship has been stable over the entire sample period under examination. Our results demonstrate two points; 1) the futures premium appears to have a small but important non-stationary component that has analogs in the forward exchange rate literature, and 2) standard cointegration and unit root tests will generally not pick up on this important fact.

The rest of the paper is organized as follows; section 2 presents the no-arbitrage-profit conditions for futures market efficiency and the conditions under which the unbiasedness hypothesis may hold. Section 3 discusses the econometric procedures used in the empirical analysis. Section 4 presents the empirical results and section 5 concludes.

\section{Cointegration and Futures Market Efficiency}

Brenner and Kroner (1995) delineate the conditions under which asset prices which are generally integrated or non-stationary processes will be cointegrated. Assume that the cash or spot price of an asset, \(S_t\), evolves according to the \(n\)-factor geometric Brownian motion,

\[ dS_t = \mu S_t dt + \sum_{i=1}^{n} \gamma_i S_t dW_{i,t}, \]  \hspace{1cm} (1)
where $W_{i,t}, i = 1, \ldots, n$ are $n$ independent Brownian motions that represent the $n$ factors or sources of uncertainty, $\mu$ is the expected return and the $\gamma_i$ are diffusion coefficients. Equation (1) implies that the natural log of the spot price evolves as a random walk. Under suitable conditions, a closed-form solution for the futures price will be given by

$$F_{t,t-k} = S_{t-k} \cdot D_{t,t-k} \cdot \exp(Q_{t,t-k}),$$

(2)

where $D_{t,t-k}$ is the cost-of-carry over the life of the contract and $Q_{t,t-k}$ is an adjustment for the mark-to-market pricing that occurs in the futures market. This adjustment term vanishes as $k$ gets larger.

Taking natural logs of (2) yields a model of the log futures price, $f_{t,t-k}$ as,

$$f_{t,t-k} = s_{t-k} + d_t + \omega_t$$

(3)

where $s_{t-k}$ is the log spot price observed in period $t-k$, $f_{t,t-k}$ is the log futures price in $t-k$ for a contract that matures in period $t$, $d_t$ is the log cost-of-carry or stochastic convenience yield (or a combination of the two) and $\omega_t$ is the (white noise) error. If the cost-of-carry is stationary, then the no-arbitrage model implies that the log futures price is cointegrated with the log spot price with a cointegrating vector of $[1, -1]'$. If the cost-of-carry is integrated, however, (3) implies a cointegrating vector of $[1, -1, -1]'$ between the three variables $f_{t,t-k}, s_{t-k}$ and $d_t$.

One issue arising in this model and not discussed by Brenner and Kroner is the proper measure of $d_t$. Yang et al. (2001) make the point that $d_t$ should be interpreted as the interest cost ($r_{T-t}$), not the interest rate ($r_t$). The difference between the two measures, interest cost vs. interest rate, will imply a cointegrating relationship among the three variables $f_{t,t-k}, s_{t-k}$ and $r_t$ of $[1, -1, -\beta_t]'$, with $r_t$ serving as a proxy for the cost-of-carry.
3 Empirical Methodology

The potential for trends in financial time series data has important implications for the choice of appropriate estimators. Specifically, if the data are characterized by stochastic trends, i.e. non-stationarity due to a unit root in the characteristic polynomial of the data generating process, then the use of Ordinary Least Squares (OLS) or a variant of OLS, i.e. GLS, 2SLS, etc., is inappropriate because of the non-standard distribution of the parameter estimates and their standard errors. We discuss in some detail the tests used to determine if the data are integrated and the statistical procedures appropriate for estimating parameters when the data are I(1).

3.1 Univariate Methodology

Several procedures exist to test for the presence of stochastic trends in time series data. The most commonly applied is the augmented Dickey-Fuller (1979) (ADF) test based on equation (4).

\[
\Delta X_t = \mu + \delta t + \rho X_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta X_{t-j} + \varepsilon_t
\]  

(4)

The null hypothesis, \( \rho = 0 \), can be tested using a t-statistic. The lag length of the augmentation terms, \( j \), is chosen as the minimum necessary to reduce the residuals to white noise. The distribution of this test is, however, not standard. Critical values for this Dickey-Fuller (DF) test have been calculated using Monte Carlo methods, most accurately by MacKinnon, et al. (1991).

Perron (1989) has demonstrated that such standard tests are subject to misspecification bias and size distortion when the series involved have undergone structural shifts. We address this issue by using the procedure advocated by Zivot and Andrews (1992). Zivot and Andrews
(1991) developed a recursive test for stochastic trends based upon equation (5).

\[
\Delta X_t = \mu + \delta t + \theta D_t + \rho X_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta X_{t-j} + \varepsilon_t
\]  \tag{5}

The variable \( D_t \) is a dummy variable that takes a value of one when \( t \) is greater than \( \lambda T \), where \( T \) is the sample size and \( \lambda \in [0.1, 0.9] \), and a value of zero otherwise.

The ZA test for a unit root in the series \( X_t \) is the minimum value of the sequence of t-statistics testing the hypothesis that \( \hat{\rho} = 0 \) over all such calculated t-statistics. Nunes et al. (1997) suggest that an alternative with good small sample properties is to choose the t-statistic testing the null that \( \hat{\rho} = 0 \) where the SBC statistic is minimized instead of where the t-statistic sequence is minimized. The SBC statistic is calculated as,

\[
SBC(T_B) = \ln \hat{\sigma}^2(T_B) + (k + c) \ln(T)/T
\]  \tag{6}

where \( \hat{\sigma}^2(T_B) \) is the residual variance from the regression in (5) with \( k \) augmentations and \( c = 5 \), the additional regressors in (5). In the empirical section we present results for both the ADF and Zivot-Andrews test procedures.

### 3.2 Multivariate Methodology

The multivariate unit root analysis is carried out using the techniques advocated by Johansen (1988, 1991). The Johansen procedure is based upon a vector autoregressive (VAR) process for \( X_t \) as in (7),

\[
X_t = \Phi_1 X_{t-1} + \ldots + \Phi_k X_{t-k} + \mu + \delta t + \varepsilon_t
\]  \tag{7}

where \( X_t \) is \( p \)-dimensional vector of variables integrated of order one or less, i.e., I(2) and higher orders are ruled out, \( \Phi_j \) are \((p \times p)\) coefficient matrices, \( \mu \) is a \((p \times 1)\) vector of
constants, $\delta$ is a $(p \times 1)$ vector of coefficients on linear trend terms and $\varepsilon_t$ is a white noise error vector with non-diagonal covariance matrix $\Omega$.

The Johansen tests are calculated from the transformed version of (7) into its error-correction model (VECM) form as in (8).

$$
\Delta X_t = \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi X_{t-1} + \mu + \delta t + \varepsilon_t
$$

(8)

The long-run multiplier matrix $\Pi = \Phi(1) - I$ can be decomposed into two $(p \times r)$ matrices such that $\alpha \beta' = \Pi$. The $(p \times r)$ matrix $\beta$ represents the cointegrating vectors or the long-run equilibria of the system of equations. The $(p \times r)$ matrix $\alpha$ is the matrix of error-correction coefficients which measure the rate each variable adjusts to the long-run equilibrium.

Maximum likelihood estimation of (8) can be carried out by applying reduced rank regression. Johansen (1988, 1991) suggests first concentrating out the short-run dynamics by regressing $\Delta X_t$ and $X_{t-1}$ on $\Delta X_{t-1}, \Delta X_{t-2}, \ldots, \Delta X_{t-k+1}$ and $t$, and saving the residuals as $R_{0t}$ and $R_{1t}$, respectively. Next calculate the product moment matrices $S_{ij} = T^{-1} \sum_{t=1}^{T} R_{it} R_{jt}'$ and solve the eigenvalue problem $|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0$. Then order the estimated eigenvalues from largest to smallest ($\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_p$). The test for cointegration is a test for the number of non-zero eigenvalues and the estimated cointegration space is the space spanned by the eigenvectors associated with these non-zero eigenvalues. The likelihood ratio statistic testing the rank of $\Pi$, or equivalently the number of non-zero eigenvalues, is given by $-T \sum_{i=r+1}^{p} \ln(1 - \hat{\lambda}_i)$ and is called the trace statistic by Johansen (1988, 1991). The distribution of this statistic is non-standard and depends upon, among other things, the specification of the deterministic components $\mu$ and $\delta$ in the VAR in (7).

In order to see the implications for the specification of constants, $\mu$, and linear trend, $\delta$, terms in (7) it is instructive to examine the moving average representation (MAR) for (8).
Equation (9) displays the MAR found by inverting (8),

$$X_t = C \sum_{i=1}^t \varepsilon_i + C \mu t + \frac{1}{2} C \delta t^2 + C^*(L)(\varepsilon_t + \mu + \delta t)$$  \hspace{1cm} (9)

where

$$C = \beta_\perp (\alpha'_\perp (I - \sum_{i=1}^{k-1} \Gamma_i) \beta_\perp)^{-1} \alpha'_\perp$$  \hspace{1cm} (10)

and $C^*(L)$ is a matrix polynomial with all characteristic roots strictly outside the unit circle and $\alpha_\perp$ and $\beta_\perp$ are the orthogonal complements to the error-correction and cointegration spaces, respectively. Equation (9) decomposes $X_t$ into permanent and transitory components. Note from (9) that inclusion of a constant in the VAR, i.e. $\mu \neq 0$ in (7) means that the underlying data generating process will include a linear trend in $X_t$. Similarly including a time trend, i.e. $\delta \neq 0$ in (7) implies that the process for $X_t$ will contain a quadratic trend. Hansen and Juselius (1995) stress that it is important to understand the role played by the constant and trend terms in the VECM because the asymptotic distribution of the tests for cointegration depend on the deterministic specification.

There are five possible specifications on the deterministic components in (8) and three that are of particular interest in this study. To more clearly see the restrictions involved with different deterministic specifications, decompose $\delta$ and $\mu$ into,

$$\delta = \alpha_1 \delta_1 + \alpha_\perp \delta_2,$$  \hspace{1cm} (11)

$$\mu = \alpha_1 \mu_1 + \alpha_\perp \mu_2,$$

where $\delta_2 = (\alpha'_\perp \alpha_\perp)^{-1} \alpha'_\perp \delta$ is a $(p - r)$-dimensional vector of quadratic trends coefficients in the data, $\delta_1 = (\alpha' \alpha)^{-1} \alpha' \delta$ is an $r$-dimensional vector of linear trend coefficients in the cointegrating relations or stationary variables, $\mu_2 = (\alpha'_\perp \alpha_\perp)^{-1} \alpha'_\perp \mu$ is a $(p - r)$-dimensional
vector of linear trend slopes in the data and $\mu_1 = (\alpha' \alpha)^{-1} \alpha' \mu$ is an $r$-dimensional vector of intercepts in the cointegrating relations or means of the stationary variables. Restrictions on the deterministic components in (8) translate into restrictions on the parameters $\delta_1, \delta_2, \mu_1$ and $\mu_2$.

- Case 1. No restrictions on $\delta_1, \delta_2, \mu_1$ and $\mu_2$. This is consistent with a quadratic trend in the levels of the data that is eliminated by the cointegrating relations and linear trends in the differenced data. Hypothesis tests for cointegration rank under this specification are denoted $H_0(r)$.

- Case 2. $\delta_2 = 0, \delta_1, \mu_1$ and $\mu_2$ unrestricted. This specification allows for a linear trend in the cointegration space, thus the cointegrating relations are trend stationary, but excludes quadratic trends from the data. This is also called "stochastic" cointegration because it eliminates only the stochastic trends in the data. Hypothesis tests for cointegration rank under this specification are denoted $H^*_0(r)$.

- Case 3. $\delta_1 = \delta_2 = 0, \mu_1$ and $\mu_2$ unrestricted. This case allows for trends in the levels of the data or non-zero means in the differences, but these trends are eliminated by the cointegrating relations. This is also called "deterministic" cointegration because it eliminates both the stochastic trends and deterministic trends in the data.

- Case 4. $\delta_1 = \delta_2 = \mu_2 = 0$ and $\mu_1$ unrestricted. This case allows for a non-zero constant in the cointegration space but excludes all deterministic trends in the levels data. Hypothesis tests for cointegration rank under this specification are denoted $H^*_1(r)$.

- Case 5. $\delta_1 = \delta_2 = \mu_1 = \mu_2 = 0$. This specification is the most restrictive in that it does not allow for any deterministic terms in the VECM. Hypothesis tests for cointegration rank under this specification are denoted $H_2(r)$.
Since it is rare that the econometrician knows the true deterministic specification a priori, Johansen (1994) has developed a likelihood ratio (LR) procedure for testing which deterministic component specification is appropriate. The form of the LR test depends on whether one is testing $H_i(r)$ versus $H_i^*(r)$ or $H_i^*(r)$ versus $H_{i+1}(r)$. Equation (12) presents the form of the LR test when testing $H_{i+1}(r)$ versus $H_i^*(r)$ using the $r$-largest eigenvalues.

$$-2 \ln \mathcal{L} \{H_{i+1}(r) \mid H_i^*(r)\} = T \sum_{j=1}^{r} \ln \left[ \frac{1 - \hat{\lambda}_{j,i}^*}{1 - \hat{\lambda}_{j,i+1}} \right]$$ (12)

The form of the alternative test is given in equation (13) which is based upon the $(p-r)$-smallest eigenvalues from estimation of (8).

$$-2 \ln \mathcal{L} \{H_i^*(r) \mid H_i(r)\} = -T \sum_{j=r+1}^{p} \ln \left[ \frac{1 - \hat{\lambda}_{j,i}^*}{1 - \hat{\lambda}_{j,i}} \right]$$ (13)

Johansen (1994) derives the distribution of the two tests in (12) and (13) and shows that they are distributed $\chi^2(r)$ and $\chi^2(p-r)$, respectively.

For example, if one wished to test a model that allowed only stochastic cointegration under a null of $H_0^*(r)$ consistent with case 2 above versus a model that included deterministic cointegration under the null $H_1(r)$ consistent with case 3 above, the appropriate form of the test would follow (12). On the other hand, if one wanted to test whether the data contained linear trends under $H_1(r)$ consistent with case 4 above versus a model that allowed only a constant in the cointegration space under the null of $H_1^*(r)$ consistent with case 3 above, one would use the LR test in (13).

The estimate of $\beta$, $\hat{\beta}$, is given by the $r$-largest eigenvectors associated with the eigenvalues $\hat{\lambda}$. Hypothesis tests on $\hat{\beta}$ can be conducted using likelihood ratio (LR) tests with standard $\chi^2$ inference. Let the form of the linear restrictions on $\beta$ be given by $\beta = H\varphi$ where $H$ is a $p \times s$ matrix of restrictions and $\varphi$ is a $s \times r$ matrix of unknown parameters. The LR test
statistic is given by,

\[ T \sum_{i=1}^{r} \ln \left( \frac{1 - \tilde{\lambda}_i}{1 - \hat{\lambda}_i} \right) \sim \chi^2_{(p-s)} \]  

(14)

where \( \tilde{\lambda}_i \) are the eigenvalues from the restricted MLE.

The estimated cointegration parameters are assumed to be constant throughout the sample. This may not be the case if the variables are subject changing economic environments, policy regimes, etc. Hansen and Johansen (1993) have developed a recursive procedure to test the stability of the cointegration parameters. They suggest basing the \( \hat{\beta} \) stability tests on the model with the short-run dynamics set to their full-sample values. This simply means basing the \( \hat{\beta} \) stability tests on the equation \( R_{0t} = \alpha \beta' R_{1t} + \epsilon_t \), where \( R_{0t} \) and \( R_{1t} \) have been defined earlier. The hypothesis of interest is that \( \beta_t = \beta \) \( \forall \ t = T_0, \ldots, T \) so that \( H = \beta \) and \( \varphi \) is the \( r \times r \) identity matrix. For each period \( t \) the LR test is calculated as,

\[ t \sum_{i=1}^{r} \ln \left( \frac{1 - \tilde{\lambda}_i}{1 - \hat{\lambda}_i} \right), t = T_0, \ldots, T, \]  

(15)

which is distributed \( \chi^2_{(p-r)} \) at each iteration.

Below we examine the appropriate deterministic specification in the VECM and then calculate the tests for cointegration based on this specification. We then examine the stability of the estimated cointegrating relations using the Hansen and Johansen technique.

4 Empirical Results

4.1 Data

We retrieved the S&P 500 spot and futures closing prices for each trading day from 21 April 1982 to 5 June 2003 from the DRI data base for a total of 5,512 observations. The futures prices are those from the nearest contract. Contracts are rolled over to the next
nearby contract one day prior to expiration. The T-bill data were obtained from the St. Louis Federal Reserve Bank. The equity index data have been transformed into natural logarithmic form and all three series are plotted in figure 1.

4.2 Univariate Unit Root Results

Implementing the ADF test from equation (4) yields test statistics of -1.39, -1.37 and -2.82 for the S&P 500 cash price and the futures price and 3-month T-bill rate, respectively.\(^2\) Comparing these calculated test statistics to the 5% critical value of -3.41 results in a non-rejection of the unit root null hypothesis.

As discussed earlier, these tests have low power when the underlying time series have undergone structural changes such as intercept changes or slope changes. To investigate the possibility that the ADF results are driven by such breaks we implement the Zivot-Andrews procedure from (5). Nunes et al. (1997) and Lee and Strazicich (2001) demonstrate that the critical values for the ZA test depend on the size of the break as measured by \(\theta\) the coefficient on \(D_t\). In figure 2 we display the ZA test statistics for both the spot and futures price series. Also included is the calculated SBC statistics and the estimated coefficient on \(D_t\), \(\hat{\theta}\). The ZA test statistics have been normalized by the appropriate 5% critical value so that values greater than one are statistically significant. Examining figure 2 reveals no evidence of rejecting the unit root null hypothesis for any series.

4.3 Multivariate Unit Root Results

The first step in our multivariate analysis is to determine the appropriate deterministic specification for the bivariate VECM of (8) spot and futures prices. Table 1 displays the

\(^2\)The ADF regressions included a constant and linear trend. The lag truncation was set at 50 for each series. The results were not sensitive to this choice.
results of the various hypothesis tests for the bivariate system. From the results in table 1 we concluded that a specification that allows for a deterministic trend in the cointegration space was most appropriate so that the estimated equilibrium is trend stationary.3

The calculated trace test statistic for the null hypothesis of zero cointegrating relations is 130.09 which can be compared to the 5% critical value from Hansen and Juselius (1995) of 25.47 and we can easily reject this hypothesis. Examining the trace statistic for the null of nor more than one cointegrating of 4.09 which is much less than the 5% critical value of 12.39. Thus we cannot reject this hypothesis and conclude that there is one cointegrating vector.

The estimated cointegrating vector is \[1.000, -1.001\]′. This is very close to the theoretically implied value. But when we test this restriction using the LR test from (14) the calculated statistic is 7.60. This statistic has an asymptotic \(\chi^2\) distribution with one degree of freedom so that this hypothesis is rejected at very high levels of significance.

Figure 3 displays the test for the stability of the equilibrium relationship given in (15) where the test statistic has been normalized by the appropriate 5% critical value such that values greater than unity imply statistical significance. Examining this graph makes it clear that the relationship between the S&P 500 spot index and the futures contract has not been stable over the entire sample. So while there appears to be strong evidence in favor of the unbiasedness hypothesis, the instability of the relationship casts doubt on its validity.

Adding the three-month Treasury bill interest rate to the model alter the results. Table displays the analysis of deterministic component specification. One again the evidence suggests that the long-run equilibrium is trend stationary. The calculated trace statistic for the null of zero cointegrating vectors is 265.36 which is much larger than the 5% critical value of 42.20 given in Hansen and Juselius (1995). The trace statistic for the null that there are

3This is consistent with the manner in which the futures price data were collected. Specifically, the time to maturity is not constant but declining over contract length.
no more than one cointegrating vector is 17.21. This can be compared to the 5% critical value of 25.47 and we can conclude that there is only one cointegrating relationship among the three variables.

This result seems at odds with those discussed earlier for the bivariate system. One possibility is that the Treasury bill rate is not statistically significant in the model. The estimated cointegrating vector is \([1.000, -1.000, 0.002]\)' so that the coefficient on the T-bill rate is very small. But the LR test excluding it from the model is 118.48 which is distributed as a \(\chi^2(1)\) variate. This null is clearly rejected and we cannot exclude the T-bill rate from the equilibrium relationship. Testing the exclusion of the spot and futures price yields test statistics of 233.88 for each test implying that these series also cannot be excluded.

Testing the symmetry between spot and futures prices, i.e. that the cointegrating vector has the form \([1, -1, *, *]'\), yields a statistic of 0.27 which is distributed \(\chi^2(1)\) and thus not significant.

The stability of this no-arbitrage-profit equilibrium can be examined by referring to figure 4. This displays the Hansen and Johansen (1993) stability tests analogous to those in figure 3. The equilibrium is now stable over the entire sample. This provides very strong evidence that the cost-of-carry model is the appropriate one for the S&P 500 index futures market.

4.4 Weak Exogeneity and Price Discovery

In this section we briefly highlight another result from our analysis. Since our preferred model has three variables and one cointegrating vector, there must exist two common trends underlying the three series.\(^4\) From the estimated VECM we can determine which of the variables are the sources of the common trends. To see this note that the model in (9) implies that the common trend components are loaded into the series through the \(C\) matrix.

in (10). Johansen (1991) interprets $\alpha'_\perp$ as the common trends and $\beta_\perp(\alpha'_\perp(I - \sum_{i=1}^{k-1}\Gamma_i)\beta_\perp)^{-1}$ as the factor loadings. This is a natural interpretation since $\alpha'_\perp X_t$ is weakly exogenous to the long-run parameters and does not respond to any disequilibrium in the system. Thus weakly exogenous variables are the sources of the common trends and they do not adjust when the system is not in equilibrium but serve as the attractors for the endogenous variables, i.e. those that do adjust to disequilibria.

In the context of the futures market, weakly exogenous variables represent those that are causally prior to the others. Price discovery, i.e. which market conveys new information first, can be determined by examining the weak exogeneity of the variables in the system. This is operationalized by testing the adjustment coefficients, $\alpha$, in the VECM. Specifically, for variables that are weakly exogenous the error correction coefficients are equal to zero. Testing this hypothesis is similar to testing restrictions on $\beta$ as in (14).

Testing the weak exogeneity of the S&P 500 cash price index yields a $\chi^2(1)$ distributed statistic equal to 1.36. This null cannot be rejected and we can conclude that the spot index is weakly exogenous. The analogous test for the futures contract is 31.47 which is significant at very high levels. Testing the weak exogeneity of the T-bill rate yields a test statistic of 2.77 so that the interest rate is also weakly exogenous. The implication of these results is that new information arrives in the spot market prior to the futures market so that the spot market serves the important role of price discovery.

5 Conclusion

This study was an investigation of the appropriate specification for the S&P 500 futures market efficiency. We examined both the simple efficiency or unbiasedness hypothesis and the no-arbitrage-profit or cost-of-carry model. While the bivariate specification of the unbiasedness hypothesis yielded a significant cointegrating relationship between spot and futures
prices, it was found to unstable over the sample. This suggested model misspecification. The trivariate cost-of-carry model also yielded one significant cointegrating relationship but this vector was found to be stable implying that this specification is more appropriate.

This apparent discrepancy seems to suggest that the spot-future differential contains a small but important permanent component. This is consistent with findings in the forward foreign exchange market, e.g. Newbold et al. (1999). It may be appropriate to model the futures premium as a non-linear process, i.e. threshold autoregression.
References


Table 1: Test for Deterministic Components in Bivariate VECM

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Test Statistic</th>
<th>5% Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0^*(r)$ vs. $H_0(r)$</td>
<td>2.51</td>
<td>3.84</td>
</tr>
<tr>
<td>$H_1^<em>(r)$ vs. $H_0^</em>(r)$</td>
<td>6.95</td>
<td>3.84</td>
</tr>
<tr>
<td>$H_1^*(r)$ vs. $H_1(r)$</td>
<td>8.50</td>
<td>3.84</td>
</tr>
<tr>
<td>$H_2(r)$ vs. $H_1^*(r)$</td>
<td>805.38</td>
<td>3.84</td>
</tr>
</tbody>
</table>

Note: Entries represent the calculated LR test statistics from (12) or (13).

Table 2: Test for Deterministic Components in Trivariate VECM

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Test Statistic</th>
<th>5% Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0^*(r)$ vs. $H_0(r)$</td>
<td>5.11</td>
<td>5.99</td>
</tr>
<tr>
<td>$H_1^<em>(r)$ vs. $H_0^</em>(r)$</td>
<td>15.87</td>
<td>3.84</td>
</tr>
<tr>
<td>$H_1^*(r)$ vs. $H_1(r)$</td>
<td>13.24</td>
<td>5.99</td>
</tr>
<tr>
<td>$H_2(r)$ vs. $H_1^*(r)$</td>
<td>973.29</td>
<td>3.84</td>
</tr>
</tbody>
</table>

Note: See notes to table 1.