

Análisis de la eficiencia del mercado de futuros: El caso del mercado de café

Óscar Hamilton García Xitumul

**Escuela Agrícola Panamericana, Zamorano
Honduras**

Noviembre, 2016

ZAMORANO
CARRERA DE ADMINISTRACIÓN DE AGRONEGOCIOS

Análisis de la eficiencia del mercado de futuros: El caso del mercado de café

Proyecto especial de graduación presentado como requisito parcial para optar
al título de Ingeniero en Administración de Agronegocios en el
Grado Académico de Licenciatura

Presentado por:

Óscar Hamilton García Xitumul

Zamorano, Honduras
Noviembre, 2016

Análisis de la eficiencia del mercado de futuros: el caso del mercado de café

Oscar Hamilton García Xitumul

Resumen: La bolsa de futuros se originó de la necesidad de fijar los precios e intercambiar diferentes productos agrícolas. La eficiencia de mercado es reflejada en la información incorporada en los precios. Asumiendo equilibrio competitivo, no es posible obtener la arbitrariedad, ganar sobre activos sobre o subvalorados. De los tres niveles de eficiencia de mercado, nuestro estudio analizó la forma débil: los precios reflejan solo información histórica. Se evaluó el mercado de café en USA durante el periodo enero 2007 - abril 2016 con el objetivo de determinar el grado de eficiencia débil tanto en precios futuros de café como en precio contado (spot). Se utilizó modelos multivariantes autoregresivos (VAR) Models. Se determinó la dinámica de los precios y la interrelación entre ellas con el tiempo. Se infiere que debido a la evidencia de cointegración entre ambas series de tiempo, existe evidencia de eficiencia de mercado a largo plazo.

Palabras clave: Eficiencia débil, modelos VAR, series de tiempo.

Abstract: Futures market was originated from the need of fixing prices and exchanging different agricultural products. Market efficiency reflects the information incorporated in the prices. Assuming perfect competition, the theory indicates that is not possible to have arbitrary opportunities: gains due mispriced assets. Among the three levels of market efficiency, we focused on the weak form which indicates that prices reflect all historical information. We evaluated the US market for the period January 2007 - April 2016. The main purpose of this article is to determine level of weak efficiency form of coffee futures and spot prices by using multivariate time series analysis (VAR models) procedure. The results show that prices are influenced at least, by lagged variables of both, futures and spot prices. Evidence of cointegration between both time series, shows that there is evidence of long-run market efficiency.

Keywords: Time series, VAR models, weak efficiency.

CONTENIDO

Portadilla	i
Página de firmas	ii
Resumen	iii
Contenido	iv
Índice de cuadros, figuras y anexos.....	v
1. INTRODUCCIÓN.....	1
2. METODOLOGÍA.....	2
3. RESULTADOS Y DISCUSIÓN.....	9
4. CONCLUSIONES	11
5. RECOMENDACIONES	14
6. LITERATURA CITADA.....	15
7. ANEXOS	17

ÍNDICE DE CUADROS, FIGURAS Y ANEXOS

Cuadros	Página
1. Resumen VAR Model	9
2. Coeficientes y nivel de significancia estadística.	9
3. Resumen VEC Model.....	11
 Figuras	 Página
1. Ejemplo de una serie de tiempo estacional (n=1000 observaciones).	4
2. Test Johansen de cointegración.	11
 Anexos	 Página
1. Contrato de futuros de café de CME GROUP.....	17
2. Sintaxis de STATA 14 para análisis.....	18
3. Comportamiento de precios periodo 2007-2016.	18
4. Resumen VAR model.....	18
5. Resultado VAR models precios Futuros vs precios Spot.	19
6. Pre estimación de número de rezagos.	19
7. Resumen VEC model.	20
8. Resultados VEC Model.	20
9. Ecuaciones de cointegración.	21
10. Restricción de normalización de Johansen impuesta.	21
11. Test causalidad de Granger.	21

1. INTRODUCCIÓN

El mercado de futuros de Chicago se especializa en la compra y venta de contratos de commodities (Fisanotti, 2014). Los precios de café de la bolsa reaccionan rápidamente a la oferta y la demanda, clima en la zona de producción y políticas gubernamentales entre otras. Mientras que los precios del mercado físico, debido a que el precio se descubre al momento de la compra-venta, reaccionan de manera lenta a estos cambios (Vázquez, 2002).

Estados Unidos es el segundo importador de café en el mundo, detrás de la Unión Europea. El total de las importaciones en 2015 fueron de USD 6,000 millones (ITC, 2016). El café arábica es el más comercializado con más de 155.7 millones de sacos 60kg, representando el 60% de la producción global, seguido por el café robusta (40% del mercado) (USDA, 2016). Los mayores exportadores de café a EE.UU son Brasil (con 454 mil toneladas métricas [MTM]), Colombia (318 MTM), Vietnam (171 MTM), Indonesia (72 MTM) y Guatemala (71 MTM) (ITC, 2016).

El café tipo Arábica es cosechado principalmente en Centro y Sur América. Es de alta calidad reflejándolo en su precio. Brasil posee alrededor del 29% de la cuota internacional de mercado (Cenessa, 2014). En Centroamérica el café representa una fuente importante de ingresos por exportaciones del producto (ANACAFE, 2016).

Los productores utilizan los precios futuros de commodities de la bolsa para la toma de decisiones, por ello sus precios deben ser eficientes y reflejar toda la información posible. Eficiencia de mercado se define cuando los precios en cualquier momento reflejan toda la información disponible (Fama, 1970).

Para este análisis se tomaron precios históricos de futuros de café de NYMEX, mercado de materias primas que pertenece a CME Group, y precios históricos de mercado físico (spot prices) en Estados Unidos de enero 2007 a marzo 2016, ambos ajustados a precios reales. Se comparó el grado de ajuste de futuros de café con “spot” prices. La teoría afirma que los precios futuros predicen los precios spot de mercado, haciendo que el grado de eficiencia sea alto. El análisis realizado para este estudio se basa en modelos de series temporales, cuyo objetivo es realizar un pronóstico de valores futuros (Villavicencio, s. f.).

El estudio determina el grado de eficiencia débil de mercado en futuros de café. Para esto, se requirió el uso de análisis de multi-series temporales utilizando modelos Vectores Autoregresivos (“VAR models”). Con ello se busca cuantificar el grado de ajuste de los precios futuros con los precios físicos de café. Establecer un posible modelo estadístico que relacione los precios de Futuros y Spot.

2. METODOLOGIA

1. Concepto de series de tiempo.

Se define como series de tiempo a una serie de eventos espaciados, ocurridos de manera cronológica. Este hecho genera dependencia entre los datos. Asimismo, poseen elementos estocásticos (cuyo resultado proviene de procesos aleatorios) y determinísticos. El principal objetivo de su análisis es comprender la dinámica de los datos y predecir valores futuros. (Villavicencio, s.f).

En las series temporales se identifican tres componentes: Tendencia, estacionalidad y componente aleatorio.

- (i) La tendencia hace referencia al cambio que se produce en el promedio de la serie en el largo plazo.
 - (ii) La estacionalidad es la periodicidad cíclica en una determinada fracción de tiempo.
 - (iii) Y el componente aleatorio es el resultado de valores imprevistos.
- (Villavicencio, s.f)

Los primeros dos componentes son determinísticos, y el tercero (como su mismo nombre menciona) es aleatorio. Por lo que se puede denotar las series de tiempo X_t como:

$$X_t = T_t + E_t + I_t \quad [1]$$

(Villavicencio, s.f)

Dónde: T_t es la tendencia, E_t el componente estacional e I_t el componente aleatorio.

Se denota x_t como una observación de la serie de tiempo (variable aleatoria) X_t la cual es denominada realización del proceso aleatorio. El comportamiento del proceso aleatorio X_t es descrito por la distribución de funciones. Para comprender este concepto, imaginemos que la serie de tiempo X_t tenga solo una sola observación x_t , entonces la distribución de la observación (probabilidad que sea menor que un dato fijo X_{t1}) es:

$$F_{t1}(x_1) = P(X_{t1} < x_1), \quad [2]$$

(Villavicencio, s.f)

Si en caso la serie dos observaciones (en tiempos t_1 y t_2), la distribución de las observaciones x_1 y x_2 están brindadas como:

$$F_{t1,t2}(x_1, x_2) = P(X_{t1} < x_1, X_{t2} < x_2), \quad [3]$$

(Villavicencio, s.f)

Por tanto, una serie de tiempo X_t está definida por la distribución conjunta de todas sus observaciones en cada tiempo (desde t_1 a t_n) de la siguiente forma general:

$$F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = P(X_{t_1} < x_1, X_{t_2} < x_2, \dots, X_{t_n} < x_n) \quad [4]$$

(Villavicencio, s.f)

2. Propiedades de las series de tiempo.

Estacionalidad estricta. Una serie de tiempo X_t es estrictamente estacional si las propiedades de la distribución de una sección son similares a los otros segmentos. En otras palabras si $\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$ y $\{X_{t_1+\tau}, X_{t_2+\tau}, \dots, X_{t_n+\tau}\}$ tienen las mismas distribuciones para cualquier entero $n \geq 1$ y cualquier entero τ (Adkins, 2011).

Series de ruido blanco IID. El modelo más simple para de una serie de tiempo es el ruido “IID” (todas las observaciones son independientes e idénticamente distribuidas). En notación científica, quiere decir que la distribución de cada observación no depende de la anterior:

$$F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = P(X_{t_1} < x_1, X_{t_2} < x_2, \dots, X_{t_n} < x_n) \quad [5]$$

$$= P(X_{t_1} < x_1) P(X_{t_2} < x_2) \dots P(X_{t_n} < x_n) = F_{t_1}(x_1) F_{t_2}(x_2) \dots F_{t_n}(x_n) \quad [6]$$

$$= F(x_1) F(x_2) \dots F(x_n) \quad [7]$$

En series de tiempo lo que se busca es que el residuo de nuestro modelo sea IID. (Adkins, 2011)

Promedio y autocovarianza. En economía, la distribución es discreta: Los intervalos de tiempos son definidos y limitados. Por lo que definimos a $p(x_t)$ como la probabilidad que ocurra x_t . Así, el promedio (ponderado) de $\{X_t\}$ es definido como el valor esperado en cada momento en el tiempo,

$$\mu(t) = E\{X_t\} = \sum_t x_t p(x_t) \quad [8]$$

La función de autocovarianza de $\{X_t\}$ es definido como la covarianza de dos observaciones con respecto a su promedio:

$$\gamma(t_1, t_2) = \text{Cov}(X_{t_1}, X_{t_2}) = \sum_t [x_{t_1} - \mu(t_1)][x_{t_2} - \mu(t_2)] p(x_{t_1}, p(x_{t_2})) \quad [9]$$

Esta mide la dependencia lineal entre dos puntos en la misma serie de tiempo observada en diferentes momentos (Adkins, 2011).

La función de Varianza, es la diferencia cuadrada de cualquier observación con respecto al promedio, es un caso especial de la autovarianza cuando $t_1 = t_2 = t$.

$$\sigma^2(t) = \sum_t [x_{t_1} - \mu(t_1)]^2 p(x_{t_1}) \quad [10]$$

(Adkins, 2011)

Propiedades de las series estacionarias. Para series de tiempo estacionarias cambiar el tiempo de origen no tiene efecto en la distribución conjunta. Esto implica que las diferencias de X_t es la misma para todo t ,

$$f_t(x) = f_{t+\tau}(x) = f_{t+(-t)}(x) = f_0(x) \Rightarrow \mu(t) = \mu(t + \tau) = \mu(t + (-t)) = \mu(0) \quad [11]$$

Por tanto, los primeros dos momentos (promedio y varianza) son finitos y constantes,

$$\mu(t) = \mu, \quad \sigma^2(t) = \sigma^2 \quad [12]$$

Entonces la función de autocorrelación depende únicamente de $t_2 - t_1 = h$, esto es llamado rezago. Y se denota como:

$$\gamma(h) = \text{Cov}(X_t, X_{t+h}) = E\{[X_t - \mu][X_{t+h} - \mu]\}; \quad (t_1 = t, t_2 = t + h) \quad [13]$$

(Adkins, 2011)

Serie de tiempo (o proceso) débil estacionario. Un procesos es estacionario en segundo orden (o débilmente estacionario) si su media es constante y la función de autocovarianza depende únicamente del rezago,

$$E(X_t) = \mu \quad \text{y} \quad \text{Cov}(X_t, X_t + h) = \gamma(h) \quad [14]$$

Note (dejando $h = 0$) la varianza al igual que la media son constantes.
Gráficamente:

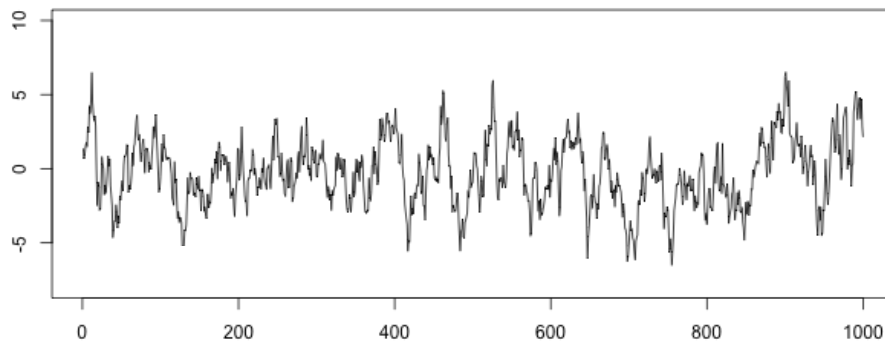


Figura 1. Ejemplo de una serie de tiempo estacional ($n=1000$ observaciones).

Podemos observar en la figura 1 la estacionalidad de la serie, pues su promedio es fijo en 0, mientras que mantiene una varianza similar a lo largo de las 1000 observaciones (Adkins, 2011).

3. Series de tiempo univariantes.

Es un modelo estocástico en función de una sola variable. Existen dos razones para utilizar modelos univariantes: (i) por motivos pedagógicos (es más sencillo comprender los conceptos) y (ii) para generar importante información como características importantes de la serie y predecir su comportamiento futuro (FES Acatlán, s. f.).

Modelos ARMA(p,q). Son los más comunes en análisis de series de tiempo **estacionario**. La variable en este modelo está en función de su propia evolución histórica. Este proceso tiene dos partes: una parte autoregresiva (AR) y una parte de medias móviles (MA).

Proceso autoregresivo (AR(p)). Un proceso autoregresivo de primer orden AR(1) se define como:

$$y_t = \mu + \lambda y_{t-1} + \varepsilon_t \quad [15]$$

Donde: μ es una constante (promedio), ε_t es ruido blanco IID. y_t es un proceso autoregresivo donde se asume que sus valor actual depende de un suceso anterior (FES Acatlán, s. f.).

Modelo de promedio móvil (MA(q)). En este modelo y_t es expresado en términos de una constante y un proceso de media móvil. Un proceso de media móvil de primer grado MA(1) es el siguiente:

$$y_t = \mu + \varepsilon_t - \theta \varepsilon_{t-1} \quad [16]$$

Donde: y_t es función de una constante μ y del elemento aleatorio de la observación anterior ε_{t-1} (FES Acatlán, s. f.).

Modelo Generalizado Autoregresivo Condicional Heterocedástico (GARCH). En ocasiones, la varianza puede ser heterocedástica (no constante a través del tiempo), lo cual hace que la serie de tiempo sea no estacionaria. Para este caso es necesario modelar también el componente de varianza utilizando modelos condicionales. El más popular es GARCH(P,Q). En el cual se modela la función ARMA dejando un residuo ε_t , (no distribuido idénticamente). Por tanto, su varianza σ_t^2 se modela con respecto a los residuos pasados (previos P) y a su varianza histórica (previos Q). Para el caso de GARCH(1,1) se denota así por ser modelada por su residuo y varianza previas (uno solo periodo atrás). En notación científica:

$$\varepsilon_t = \sigma_t A_t \quad \text{Con} \quad \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \delta_1 \sigma_{t-1}^2 \quad [17]$$

Donde: $\alpha_0 > 0$, $\alpha_1 > 0$, $\delta_1 < 1$, $\alpha_1 + \delta_1 < 1$ y (A_t) es independiente de (σ_t).

Un modelo GARCH como el anterior implica que los componentes del proceso (ε_t) no están autocorrelacionados aunque no son estadísticamente independientes debido a que la varianza condicionada de ε_t está en función de ε_{t-1} (Mauricio, 2007).

4. Series de tiempo multivariantes.

Consiste en un conjunto de series, por ejemplo:

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix}$$

Comúnmente existe una variable dependiente, o endógena cuyo valor se explica a través de un modelo, y una serie de variables explicativas o exógenas, determinadas fuera del modelo. Por lo que se puede postular un sistema de ecuaciones:

$$x_t = \gamma_0 + \gamma_1 x_{t-1} + \gamma_2 y_{t-1} + \gamma_3 z_{t-1} + u_t \quad [18]$$

$$y_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 y_{t-1} + \beta_3 z_{t-1} + \varepsilon_t \quad [19]$$

$$z_t = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \alpha_3 z_{t-1} + v_t \quad [20]$$

(Castillo, Varela, s. f.)

Modelos de Vectores Autoregresivos (VAR). Son los más utilizados para el análisis de series de tiempo multivariantes. Este procedimiento se usa para describir el comportamiento dinámico de series de tiempo en economía, finanzas y pronósticos. A diferencia de las univariantes, se pueden establecer relaciones de dependencias entre diferentes series de tiempo. Por tanto, es basado en ecuaciones simultáneas para generar mejores pronósticos (Wang, J., Zivot, E., 2006).

Un modelo bivariado VAR₂(1) presenta la siguiente forma matricial.

$$\begin{pmatrix} y_{t1} \\ y_{t2} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \beta_{11}^1 & \beta_{12}^1 \\ \beta_{21}^1 & \beta_{22}^1 \end{pmatrix} \begin{pmatrix} y_{t1-1} \\ y_{t2-1} \end{pmatrix} + \begin{pmatrix} \beta_{11}^2 & \beta_{12}^2 \\ \beta_{21}^2 & \beta_{22}^2 \end{pmatrix} \begin{pmatrix} y_{t1-2} \\ y_{t2-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{pmatrix} \quad [21]$$

Donde: β es la representación de los parámetros, $cov(\varepsilon_{t1}, \varepsilon_{t2}) = \sigma_{12}$, para $t = s$; 0 en otro caso. Nótese que cada ecuación tiene la misma cantidad de regresores retardados para y_{t1-2} y y_{t2-2} . Este es el caso más simple del modelo de VAR(p) (donde p representa el periodo de tiempo pasado) está formado solamente por regresiones aparentemente no relacionadas con variables retardadas y términos determinísticos como regresores comunes (Novales, 2014). Existen casos, sin embargo donde puede existir relación en la matriz de covarianza, en las denominadas VAR estructurales.

VARMA. Comúnmente utilizado para modelar series correlacionadas en tiempo. Un proceso multivariante estacionario (Y_t) vectorial ARMA Vectorial [VARMA (p, q)] se define como:

$$Y_t = \mu + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \dots - \theta_q A_q \quad [22]$$

Para todo $t=0, \pm 1, \pm 2, \dots$, (A_t) y μ como matriz de constantes, $\beta_1, \beta_2, \dots, \beta_p$ representan los parámetros de la autocorrelacion y $\theta_1, \theta_2, \dots, \theta_q$ los parámetros de promedios móviles.

Asimismo es estacionario si cumple la siguiente condición del determinante (Muricio, 2007).

$$|I - \beta_1 x - \beta_2 x^2 - \dots - \delta_p x^p| = 0 \quad [23]$$

Las raíces de la determinante deben estar fuera del círculo unitario.

Modelo de Error de Corrección del Vector (VEC). El modelo de Error de Corrección del Vector (VEC) es un caso especial de VAR para variables que son estacionarias en sus diferencias. Por tanto, VEC incluye la relación de cointegración (que las series de tiempo poseen una tendencia estocástica en común) entre las variables dependientes.

Tomando de ejemplo que dos series de tiempo con variables Y_t y X_t están interrelacionadas, se tiene un sistema de ecuaciones VAR cointegrado orden 1 (VAR(1) con I(1)):

$$\begin{aligned} \Delta y_t &= \beta_{10} + \beta_{11} y_{t-1} + \beta_{12} x_{t-1} + v_t^y \\ \Delta x_t &= \beta_{20} + \beta_{21} y_{t-1} + \beta_{22} x_{t-1} + v_t^x \end{aligned} \quad [24]$$

Cada variable dependiente es tomada la diferencia del periodo anterior (Δ representa diferencia) y está en función de sus propios rezagos y valores históricos de las otras variables en el sistema (Adkins, 2011).

5. Justificación Modelo VAR.

Con el objetivo de determinar la eficiencia débil de mercado de futuros de café el presente estudio aplica el procedimiento estadístico de Vector Autoregressive models. Este modelo es utilizado para describir la dinámica del comportamiento económico y financiero de series de tiempo y de previsión de precios. Genera mejores previsiones que los modelos univariados de series de tiempo estableciendo una teoría basada en modelos de ecuaciones simultáneas (Wang, J., Zivot, E., 2006).

El modelo VAR fue introducido como una alternativa a los modelos macroeconómicos debido a que es común que series de precios agrícolas sean afectados por índices económicos. Asimismo, la literatura sostiene comunes relaciones entre series de precios. En nuestro proyecto se ha buscado que evaluar precios de mercado futuro y presentes, por lo que su estimación es justificable. Por otro lado, la modelación VARMA no es utilizada en el campo de la economía agrícola, pues su resolución son órdenes que no han sido desarrolladas para esa aplicación. Además, VAR provee aproximaciones más sistemáticas a las restricciones impuestas y permite identificar irregularidades que permanecerían ocultas si se utilizan dos modelos univariados (Bjørnland, 2000).

Un modelo VAR se utiliza cuando se desea predecir varias series de tiempo de manera simultánea debido a que tienen a estar relacionadas entre sí. Tal es el caso de los precios de futuros y spot de café en USA, donde los precios café están en función de sus valores históricos.

Mercados Eficientes. Es aquel mercado cuyos precios reflejan cierto nivel de información. Así, las empresas pueden tomar decisiones de inversión o producción (Fama, E. 1970).

La información del mercado está basada en la teoría del paseo aleatorio: no es posible predecir los cambios en precios de acciones. Por tanto, no existe arbitrariedad (no es posible obtener ganancias sobre activos sobre o subvalorados). Así, enunciar que los mercados reflejan información eficientemente implica (1) que los cambios en precios ocurren de manera aleatoria y (2) que estos cambios se distribuyen de manera uniforme.

La teoría del paseo aleatorio afirma que Existen tres niveles de eficiencia:

- (i) Eficiencia débil: los precios de mercado reflejan toda la información pasada.
 - (ii) Eficiencia Semi fuerte: La información pasada y también pública están reflejados en los precios.
 - (iii) La eficiencia fuerte: La información pasada, pública y privada ya ha sido asimiladas por el mercado y por tanto se refleja en el precio de los activos.
- (Fama, E. 1970).

Datos. Se analizaron dos series de tiempo, los precios de futuro y los spot prices de café de Estados Unidos. La serie de datos contiene 2338 observaciones diarias durante el periodo 11 de enero 2007 al 22 abril 2016. Los datos futuros son precios de cierre reportados por NYMEX y la serie de spot prices son reportados por MacroTrends. Ambas series se ajustaron a precios reales usando como base el precio de futuros de enero 2016.

3. RESULTADOS Y DISCUSIÓN

Para realizar el procedimiento estadístico, se definieron tres variables rezagadas o Lags, esto indica que tanto los precios futuros como spot estarán en función de tres periodos de tiempo anteriores. Dando como resultado seis variables para cada ecuación, tres coeficientes en función de los precios futuros y tres en función de los precios spot.

Cuadro 1. Resumen VAR Model

Ecuación	R ²	X ²	p > X ²
Futuros	0.9911	260,227	0.0000
Spot	0.9923	302,255	0.0000

Los valores de R² para los modelos de Ft (Futuros) y St (spot) reflejan un buen ajuste de modelo es de 0.99 para ambas. Asimismo, la prueba chi² por ser menor a 0.01 sugiere que es poco probable ambas variables sean independientes entre sí. De esta manera afirmamos que existe dependencia entre las series, las variables presentan dependencia, y que existe causalidad de las variables en el modelo (ver Cuadro 13).

Cuadro 2. Coeficientes y nivel de significancia estadística.

Ecuación	Coeficiente	p> Z
Futuros	f_{t-1}	0.56857
	f_{t-2}	0.07078
	f_{t-3}	0.18524
	s_{t-1}	0.57251
	s_{t-2}	-0.19783
	s_{t-3}	-0.19938
	Constante	0.01256
		0.000
Spot	f_{t-1}	0.10864
	f_{t-2}	-0.01386
	f_{t-3}	-0.04136
	s_{t-1}	0.75207
	s_{t-2}	0.11207
	s_{t-3}	0.07894
	Constante	0.00165
		0.600

Como puede observarse el Cuadro 2 muestra los resultados del procedimiento estadístico del modelo VAR₂(3) model. Las salidas del procedimiento muestran los coeficientes, las desviaciones estándar, valores Z, niveles de significancia y los intervalos de confianza. Aquí, podemos ver efectivamente, que se necesitan tres rezagos para la predicción de la serie multivariable.

Los resultados muestran que existen variables que no presentan significancia estadística como el caso de la constante y el segundo rezago en precios futuros de la segunda regresión (ver Ec. 25), es posible excluir variables en función a su significancia estadística, sin embargo hay razones para no hacerlo (Novales, 2014). “Primero, si se mantiene el mismo número de variables explicativas en todas las ecuaciones, facilita la estimación por mínimos cuadrados ordinarios (MCO, por sus siglas en inglés) ecuación por ecuación. Segundo, la presencia de rezagos como variables exógenas hacen la colinealidad entre variables, importante, por lo que se pierde precisión en la estimación de un modelo. Finalmente no es recomendable excluir aquellas variables cuyos coeficientes no presenten significancia estadística, debido a la colinealidad inherente del modelo” (Novales, 2014).

Los resultados del procedimiento estadístico generaron las siguientes regresiones:

$$F_t = 0.01256 + 0.56857 * f_{t-1} + 0.07078 * f_{t-2} + 0.18524 * f_{t-3} + 0.57251 * s_{t-1} - 0.19783 * s_{t-2} - 0.19938 * s_{t-3} \quad [25]$$

$$S_t = 0.00165 + 0.10864 * f_{t-1} - 0.01386 * f_{t-2} - 0.04136 * f_{t-3} + 0.75207 * s_{t-1} + 0.11207 * s_{t-2} + 0.07894 * s_{t-3} \quad [26]$$

Donde:

F_t es el precio futuro en el periodo t y S_t es el precio spot del periodo t .

f_{t-i} es el precio futuro rezagado i días.

s_{t-i} es el precio spot rezagado i días.

Las variables están expresadas en USD por libra de café.

Cada $\hat{\beta}_i$ es la derivada parcial del valor esperado de F_t y S_t (Y) dado los rezagos (X) $f_{t-1}, f_{t-2}, f_{t-3}, s_{t-1}, s_{t-2}, s_{t-3}$ en función de X_i ,

$$\hat{\beta}_i = \frac{\partial E(Y|X)}{\partial X_i} \quad [27]$$

La interpretación es de la siguiente manera: Para la ecuación (25), f_{t-1} Indica que por cada dólar adicional en el precio de futuros del día anterior, el precio de futuros que se espera recibir se incrementa en USD 0.56857. Por la misma lógica. f_{t-2} muestra que por cada dólar adicional en el precio de futuros de dos días anteriores el precio de futuros esperado a obtener aumenta en USD 0.07078, y así sucesivamente. Esta lógica también se aplica para la interpretación de los demás coeficientes en la ecuación 25 y 26.

Johansen tests for cointegration					
Trend: constant			Number of obs =		2335
Sample: 4 - 2338			Lags =		3
maximum				trace	5%
rank	parms	LL	eigenvalue	statistic	critical
0	10	8313.5441	.	153.0644	15.41
1	13	8388.7005	0.06235	2.7515*	3.76
2	14	8390.0763	0.00118		
maximum				max	5%
rank	parms	LL	eigenvalue	statistic	critical
0	10	8313.5441	.	150.3129	14.07
1	13	8388.7005	0.06235	2.7515	3.76
2	14	8390.0763	0.00118		

Figura 2. Test Johansen de cointegración.

Debido a que el estadístico supera al valor crítico del cinco por ciento, se rechaza H_0 de no cointegración (relación entre las variables a largo plazo). La hipótesis alterna (H_1)= afirma que por lo menos existe una ecuación con cointegración. Podemos observar, específicamente, que existe al menos integración de orden 1 ($2.75 < 3.76$), por tanto también se propone como alternativa el uso de vectores de corrección para errores (VEC models).

Cuadro 3. Resumen VEC Model.

Ecuación	R ²	X ²	p > X ²
Futuros	0.2513	781	0.0000
Spot	0.0537	132	0.0000

Cuadro 5. Coeficientes y nivel de significancia estadística.

Ecuación	Coefficiente	p> Z
Futuros	f_{t-1}	-0.25636
	f_{t-2}	-0.18566
	s_{t-1}	0.39648
	s_{t-2}	0.19880
	Cointegración	-0.17483
	Constante	0.00002
Spot	f_{t-1}	0.05470
	f_{t-2}	0.04074
	s_{t-1}	-0.19208
	s_{t-2}	-0.07978
	Cointegración	0.05428
	Constante	0.00008

Los resultados del procedimiento estadístico generaron las siguientes regresiones:

$$\Delta F_t = -0.17483(0.00002 + 0.25636f_{t-1} + 0.18565f_{t-2}) + 0.39647s_{t-1} + 0.19880s_{t-2} \quad [28]$$

$$\Delta S_t = 0.05428(-0.00008 - 0.05470f_{t-1} - 0.04074f_{t-2}) - 0.19208s_{t-1} - 0.07978s_{t-2} \quad [29]$$

Donde:

F_t es el precio futuro en el periodo t y S_t es el precio spot del periodo t .

f_{t-i} es el precio futuro rezagado i días.

s_{t-i} es el precio spot rezagado i días.

Las variables están expresadas en USD por libra de café.

La interpretación es de la siguiente manera: Para la ecuación (28), f_{t-1} Indica que por cada dólar adicional en el precio de futuros del día anterior, el cambio de precios de un día para otro, se reduce en USD 0.25636. Por la misma lógica. f_{t-2} muestra que por cada dólar adicional en el precio de futuros de dos días anteriores el cambio de precios futuros esperado a obtener se reduce en USD 0.18565, y así sucesivamente. Esta lógica también se aplica para la interpretación de los demás coeficientes en la ecuación 28 y 29.

4. CONCLUSIONES

- Ambas series de tiempo presentan un nivel de significancia alto ($p > 0.0000$). Esto asume que los precios están en función de la dinámica entre las variables rezagadas de precios futuros y spot.
- Se proponen dos alternativas el uso de vectores autoregresivos (VAR) y de vectores de corrección del error (VEC Model) por la posibilidad de cointegración. Esto evidencia relación entre las series de tiempo a largo plazo.

5. RECOMENDACIONES

- Ampliar el estudio a más cultivos que representen importancia para las economías de Estados Unidos y Centro América.
- Analizar otros niveles de eficiencia de mercado de futuros.
- Evaluar la factibilidad de utilizar los contratos de futuros como fondos de cobertura para productores de materias primas.

6. LITERATURA CITADA

Adkins, L. (2011). Vector Error Correction Models. Octubre 10, 2016, de LearnEconometrics Sitio web:

<http://www.learneconometrics.com/class/5263/notes/Vector%20Error%20Correction%20Models.pdf>

Bjørnland, H. (2000). VAR Models in Macroeconomics Research. Octubre 6, 2016, de Statistisk sentralbyrå Sitio web: https://www.ssb.no/a/histstat/doc/doc_200014.pdf

Castillo, R., Varela, R. (s. f.). ECONOMETRÍA PRÁCTICA: Fundamentos de series de tiempo. Octubre 6, 2016, de Cal State LA Sitio web: <http://web.calstatela.edu/faculty/rcastil/Personal/LibroEconometriaInternet.pdf>

Cenessa, R. (2014). Futuros sobre café. Octubre 6, 2016, de Técnicas de Trading Sitio web: <http://www.tecnicasdetrading.com/2014/01/futuros-sobre-cafe.html>

CME Group. (2013). CME Group Overview. Septiembre 6, 2016, de CME Group Sitio web: <http://www.cmegroup.com/company/files/cme-group-overview.pdf>

CME Group. (s. f.). CME Group Coffee Futures Contract Specs. Septiembre 6, 2016, de CME Group Sitio web: http://www.cmegroup.com/trading/agricultural/softs/coffee_contract_specifications.html

Fama, E. (1970, mayo). Efficient capital markets: a review of theory and empirical work. The Journal of Finance, Vol. 25, No 2, 383-417. 2016, septiembre 7, De Jstore.org Base de datos.

Fisanotti, L. (2014). Antecedentes históricos de los mercados de futuros y opciones: cobertura y especulación. Septiembre 6, 2016, de Universidad Invenio Sitio web: <https://dialnet.unirioja.es/descarga/articulo/4839237.pdf>

FES Acatlán. (s. f.). Capítulo 2. Modelos de series de tiempo estacionarios univariados. Octubre 2, 2016, de UNAM Sitio web: <https://es.scribd.com/doc/307478687/Modelos-de-Series-de-Tiempo-Estacionarios-Univariados>

ITC. (2016). List of supplying markets for the product imported by United States of America in 2015 Metadata Product: 0901 Coffee, whether or not roasted or decaffeinated; coffee husks and skins; coffee substitutes containing coffee in any proportion. Octubre 6, 2016, de International Trade Centre Sitio web: http://trademap.org/Country_SelProductCountry.aspx

Mauricio, J. (2007). Análisis de Series Temporales. Octubre 10, 2016, de Universidad Complutense de Madrid Sitio web: <https://www.ucm.es/data/cont/docs/518-2013-11-11-JAM-IASST-Libro.pdf>

Novales, A. (2014). Modelos Vectoriales Autoregresivos (VAR). Septiembre 8, 2016, de Universidad Complutense Sitio web: <https://www.ucm.es/data/cont/media/www/pag-41459/VAR.pdf>

USDA. (2016). Coffee: World Market and Trades. Octubre 6, 2016, de Foreign Agricultural Services. Sitio web: <http://www.fas.usda.gov/data/coffee-world-markets-and-trade>

Vázquez, M. (2002). El café y bolsa de Nueva York. Septiembre 6, 2016, de Fundación Internacional para el Desafío Económico Global Sitio web: <http://www.elobservadoreconomico.com/articulo/259>

Villavicencio, J. (s.f). Introducción a series de tiempo. Septiembre 7, 2016, de Instituto de estadísticas de Puerto Rico Sitio web: http://www.estadisticas.gobierno.pr/iepr/LinkClick.aspx?fileticket=4_BxecUaZmg%3D

Wang, J., Zivot, E. (2006). Vector Autoregressive Models for Multivariate Time Series. En Modeling Financial Time Series with S-Plus®(385-428). United States of America: Springer.

7. ANEXOS

Anexo 1. Contrato de futuros de café de CME GROUP.

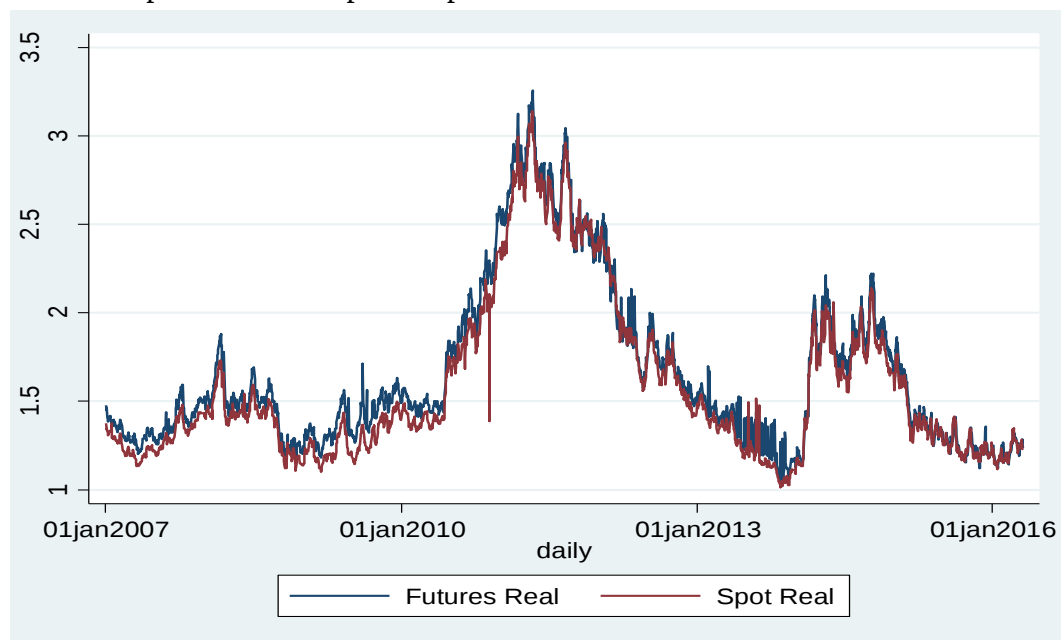
Contract Unit	37,500 pounds	
Price Quotation	U.S. Dollars per pound	
Trading Hours	CME Globex:	Sunday - Friday 6:00 p.m. - 5:00 p.m. (5:00 p.m. - 4:00 p.m. Chicago Time/CT) with a 60-minute break each day beginning at 5:00 p.m. (4:00 p.m. CT)
	CME ClearPort:	Sunday - Friday 6:00 p.m. - 5:00 p.m. (5:00 p.m. - 4:00 p.m. Chicago Time/CT) with a 60-minute break each day beginning at 5:00 p.m. (4:00 p.m. CT)
Minimum Price Fluctuation	\$0.0005 per pound	
Product Code	CME Globex: KT CME ClearPort: KT Clearing: KT	
Listed Contracts	Trading is conducted in the March, May, July, September, and December cycle for the next 23 months.	
Settlement Method	Financially Settled	
Termination Of Trading	Trading terminates on the day immediately preceding the first notice day of the corresponding trading month of Coffee "C" futures at ICE Futures U.S.	
Settlement Procedures	Nymex Coffee Settlement Procedures	
Position Limits	NYMEX Position Limits	
Exchange Rulebook	NYMEX 931	
Block Minimum	Block Minimum Thresholds	
Price Limit Or Circuit	Price Limits	
Vendor Codes	Quote Vendor Symbols Listing	

Fuente: CME Group

Anexo 2. Sintaxis de STATA 14 para análisis.

```
. tsset daily, daily
      time variable: daily, 03jan2007 to 22apr2016, but with gaps
      delta: 1 day
. var futuresreal spotreal, lags(1/3) lutstats nocnsreport vsquish
```

Anexo 3. Comportamiento de precios periodo 2007-2016.



Anexo 4. Resumen VAR model.

```
. var futuresreal spotreal, lags(1/3) lutstats
```

Vector autoregression

```
Sample: 4 - 2338
Log likelihood = 8390.076      (lutstats) AIC      = -12.85184
FPE              = 2.63e-06    HQIC      = -12.84106
Det(Sigma_ml)   = 2.59e-06    SBIC      = -12.82226
```

Equation	Parms	RMSE	R-sq	chi2	P>chi2
futuresreal	7	.043565	0.9911	260227.9	0.0000
spotreal	7	.039879	0.9923	302255.9	0.0000

Anexo 5. Resultado VAR models precios Futuros vs precios Spot.

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
futuresreal						
futuresreal						
L1.	.5685738	.0215507	26.38	0.000	.5263353	.6108124
L2.	.0707779	.0242824	2.91	0.004	.0231852	.1183706
L3.	.1852369	.0198867	9.31	0.000	.1462596	.2242142
spotreal						
L1.	.5725116	.0240175	23.84	0.000	.5254382	.619585
L2.	-.1978298	.0295449	-6.70	0.000	-.2557367	-.1399229
L3.	-.1993785	.0260668	-7.65	0.000	-.2504685	-.1482886
_cons	.0125647	.0034383	3.65	0.000	.0058258	.0193037
spotreal						
futuresreal						
L1.	.1086444	.0197273	5.51	0.000	.0699797	.1473092
L2.	-.0138552	.0222279	-0.62	0.533	-.0574211	.0297107
L3.	-.041356	.0182041	-2.27	0.023	-.0770354	-.0056766
spotreal						
L1.	.7520725	.0219854	34.21	0.000	.708982	.795163
L2.	.1120742	.0270451	4.14	0.000	.0590668	.1650816
L3.	.078944	.0238613	3.31	0.001	.0321767	.1257112
_cons	.00165	.0031474	0.52	0.600	-.0045188	.0078187

Anexo 6. Pre estimación de número de rezagos.

```
. varsoc futuresreal spotreal, maxlag(3)
```

Selection-order criteria

Sample: 4 - 2338

Number of obs = 2335

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	1712.09				.000792	-1.46475	-1.46295	-1.45982
1	8174.93	12926	4	0.000	3.1e-06	-6.99694	-6.99155	-6.98215
2	8308.52	267.17	4	0.000	2.8e-06	-7.10794	-7.09896	-7.08329
3	8390.08	163.12*	4	0.000	2.6e-06*	-7.17437*	-7.1618*	-7.13986*

Endogenous: futuresreal spotreal

Exogenous: _cons

Anexo 7. Resumen VEC model.

```
. vec futuresreal spotreal, trend(constant) lags(3)
```

Vector error-correction model

Sample: 4 - 2338	No. of obs	=	2335
	AIC	=	-7.174048
Log likelihood = 8388.701	HQIC	=	-7.162373
Det(Sigma_ml) = 2.60e-06	SBIC	=	-7.142003

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_futuresreal	6	.043565	0.2513	781.617	0.0000
D_spotreal	6	.039892	0.0537	132.0447	0.0000

Anexo 8. Resultados VEC Model.

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
D_futuresreal						
_cel						
L1.	-.1748283	.0176513	-9.90	0.000	-.2094241	-.1402324
futuresreal						
LD.	-.2563623	.0230479	-11.12	0.000	-.3015352	-.2111893
L2D.	-.1856578	.0199122	-9.32	0.000	-.224685	-.1466306
spotreal						
LD.	.3964779	.0269509	14.71	0.000	.3436551	.4493007
L2D.	.1988026	.0260996	7.62	0.000	.1476482	.2499569
_cons	-.0000247	.0009016	-0.03	0.978	-.0017917	.0017424
D_spotreal						
_cel						
L1.	.0542825	.0161631	3.36	0.001	.0226033	.0859617
futuresreal						
LD.	.0547051	.0211048	2.59	0.010	.0133405	.0960697
L2D.	.040743	.0182335	2.23	0.025	.005006	.0764799
spotreal						
LD.	-.1920821	.0246787	-7.78	0.000	-.2404515	-.1437126
L2D.	-.0797828	.0238993	-3.34	0.001	-.1266245	-.0329411
_cons	-.0000794	.0008256	-0.10	0.923	-.0016975	.0015387

Anexo 9. Ecuaciones de cointegración.

Cointegrating equations

Equation	Parms	chi2	P>chi2
_cel	1	12494.3	0.0000

Identification: beta is exactly identified

Anexo 10. Restricción de normalización de Johansen impuesta.

Johansen normalization restriction imposed

beta	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_cel						
futuresreal	1	.	.	.	.	.
spotreal	-1.010743	.0090424	-111.78	0.000	-1.028466	-.9930199
_cons	-.0537576	.	.	.	.	.

Anexo 11. Test causalidad de Granger.

. vargranger

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
futuresreal	spotreal	345.56	3	0.000
futuresreal	ALL	345.56	3	0.000
spotreal	futuresreal	8.7541	3	0.033
spotreal	ALL	8.7541	3	0.033