



Universidad de  
**San Andrés**

**Universidad de San Andrés**

Departamento de Finanzas

Trabajo de Graduación de la Licenciatura en Finanzas

¿Cuál es la dinámica entre los precios *spot* y  
futuro del Bitcoin?

Análisis de causalidad en sentido de Granger  
durante la corrección de 2018.

Autor: Nicolas Goldrossen

Numero de Legajo: 28150

Mentor de Tesis: Gabriel Basaluzzo

Buenos Aires, Argentina – Diciembre del 2021.

# Índice

1. <b>Introducción</b> .....	4
2. <b>Revisión literaria</b> .....	6
3. <b>Datos</b> .....	7
4 <b>Metodología</b> .....	8
4.1 Cointegración .....	8
4.1.1 Cointegración por Granger.....	9
4.1.2 El enfoque de Johansen – Juselius.....	10
4.2 Modelo de Corrección de Errores .....	11
5. <b>Análisis y resultados</b> .....	13
5.1 Cointegración vía Engle y Granger .....	13
5.2 Cointegración vía Johansen .....	15
5.3 Modelo de Corrección de Errores .....	17
6 <b>Conclusión</b> .....	24
7. <b>Bibliografía</b> .....	25
8. <b>Anexo</b> .....	26

## Resumen

A fines de 2017 el precio del Bitcoin tuvo una caída de aproximadamente 70% de su valor en dos meses. Esta caída coincidió con la introducción del mercado de futuros. La entrada de este derivado pudo ser un condicionante del desplome ya que a partir de este instrumento los inversores pueden tomar una posición corta con más facilidad vendiendo futuros. En esta tesis vamos a realizar un análisis empírico sobre la relación de causalidad entre los futuros y el precio del activo subyacente mediante modelos econométricos. A partir de este análisis podremos realizar una interpretación para ver la relación entre mercados.

**Palabras clave:** Price Discovery, Bitcoin, Futuros



## 1. Introducción

Bitcoin es la más común y conocida de las criptomonedas en el mundo, y su popularidad ha ido incrementando día a día. Debido a que no tiene un valor intrínseco o fundamental como una empresa, su precio depende mayoritariamente de la oferta y demanda.

Los distintos contextos cambiantes en el mundo han llamado mucho la atención sobre las criptomonedas, tanto que para el 18 de diciembre de 2021 el capital de mercado de Bitcoin ronda los 888 mil millones de dólares.

Las criptomonedas son conocidas por ser uno de los activos más especulativos y volátiles en la historia. En los 13 años de vida desde su creación se han observado grandes fluctuaciones, una de las más fuertes ocurrió a finales de 2017, cuando su valor cayó 70% en dos meses.

Los futuros son un mercado regulado en donde los participantes acuerdan comprar o vender un activo en una específica fecha futura a un precio establecido hoy. El contrato futuro es comúnmente utilizado para tomar una posición “corta” sobre un activo, es decir, obtener retorno ante una caída en el precio del subyacente. Este instrumento también se utiliza para obtener la tasa de interés implícita del activo. A partir de estos fundamentos sospechamos que la fuerte caída en el precio de Bitcoin ante la creación de futuros no parece ser una simple coincidencia sino una consecuencia del comportamiento de los tenedores. Los contratos futuros sobre Bitcoin se introdujeron el 13 de diciembre de 2017, por la CME <sup>1</sup> y el CBOE <sup>2</sup>.

El objetivo del presente trabajo consiste en evaluar la interacción entre los mercados *spot* y futuro de Bitcoin en el marco de la corrección que comenzó a fines de 2017 y finalizó en noviembre 2018. Analizaremos la causalidad y la dinámica de la relación, comprobaremos si el precio *spot* está

---

<sup>1</sup> *Chicago Mercantile Exchange* es un mercado estadounidense de derivados financieros y de productos básicos.

<sup>2</sup> El *Chicago Board Options Exchange* es el mayor intercambio de opciones de EE. UU.

liderado por los precios futuros o no y, en paralelo, realizaremos interpretaciones financieras para enriquecer el análisis econométrico. En otras palabras, a lo largo de la tesis intentaremos responder a las siguientes preguntas: “¿Qué mercado se movió primero?” y “¿Qué mercado provoca un efecto mayor en el otro?”.

Lo que motivó la elección de este objetivo es comprender la importancia de la introducción del mercado de derivados sobre un activo financiero tan volátil debido a que se estarán implementando estos en otras criptomonedas, tal es así que el 8 de febrero de 2021 se inauguró el mercado de futuros sobre Ether (\$ETH).

El trabajo está estructurado de la siguiente manera: la sección 2 contiene una revisión de literatura donde se consignan los distintos trabajos que han contribuido para realizar los modelos econométricos; en la sección 3 se describen los distintos modelos que inspiraron la metodología desarrollada para el estudio de la relación futuro-*spot*. Finalmente, la sección 4 contiene una conclusión con los resultados.

## 2. Revisión Literaria

Existen decenas de modelos (tanto paramétricos como no paramétricos) que establecen una relación entre dos activos a partir de una muestra de datos, lineales como no lineales.

Lo primero que se debe definir por medio de estos modelos es la exogeneidad/endogeneidad de ambos mercados entre las cuales veremos varias posturas. Por un lado, Karkkainen (2018) establece que el precio *spot* sucede al futuro. Por el otro lado, Baur y Dimpfl (2019) establecen que el mercado de futuros maneja el precio *spot*. Otros autores que argumentan que ambas se mueven simultáneamente (Giudici (2019)). La teoría marca que el mercado de futuros es más exógeno que el *spot*, y uno de los objetivos del trabajo será definir y verificar el comportamiento en el mercado de criptomonedas.

Estos autores previamente mencionados estudian los precios del Bitcoin mediante dos métodos, *Vector Error Correction Models* (VECM) y Vectores Autorregresivos (VAR). Baur y Dimpfl (2019) utilizan VECM, Karkkainen (2018) aplica la metodología VAR para examinar el modelo de precios mediante la descomposición de Cholesky para luego realizar un chequeo robusto.

En cuanto a los modelos no paramétricos, Pagnottoni (2019) desarrolla un modelo para analizar la relación entre el mercado de futuros y *spot* a través de redes neuronales.

Los dos autores que más contribuyeron al trabajo observan distintos métodos y resultados: Karkkainen, T. (2018) utiliza una base de datos de mayor frecuencia (1 minuto, 5 minutos, 15 minutos, 30 minutos, 1 hora y 1 día) en un periodo menor de tiempo (13 de diciembre 2017 a 16 de

abril 2018), y llega a la conclusión de que a mayor frecuencia el nivel de significatividad es cada vez más determinante en favor a los futuros estadísticamente causando los precios *spot*. En su estudio realiza un modelo de corrección de errores y un modelo lineal en donde llega a la misma conclusión: en términos intradiarios los precios futuros preceden al *spot* y en términos diarios los precios *spot* precede a los futuros.

Por otro lado, Hattori, T. y Ishida, R (2020) realizan un análisis financiero donde estudian la relación del mercado de futuros y el spot del Bitcoin a través de la arbitrariedad. Ellos concluyen que no hay evidencia de una conexión entre los dos mercados durante el periodo analizado. Utilizan los precios diarios del Bitcoin entre el 13 de diciembre de 2017 y el 13 de diciembre de 2018.



### 3. Datos

Para crear la base de datos se utilizaron precios en dólares y volúmenes diarios de Bitcoin *spot* descargados de Investing.com <sup>3</sup> y los precios futuros diarios por parte de la CME <sup>4</sup>. La serie de tiempo obtenida es desde el 13 de diciembre de 2017 hasta el 16 de noviembre de 2018, lo que serían los primeros 11 meses desde la inauguración del mercado de futuros. Los precios *spot* utilizados provienen de Bitfinex <sup>5</sup>, y son los precios al cierre del día. La CME utiliza precios de otra plataforma por lo que son ligeramente distintos. Los contratos futuros de la CME contienen el *ticker* BMC. Se calculan a partir de un promedio ponderado de Bitstamp, Coinbase, itBit, and Kraken. Al construir la serie de precios tuvimos únicamente en cuenta los contratos a un mes de duración debido a que eran los que predominaban en cuanto a volumen. Utilizamos los datos de la CME en vez de la CBOE debido a que había mejor disponibilidad de datos. En las situaciones donde no había volumen o el mercado se encontraba cerrado, las observaciones de ambas series fueron omitidas<sup>6</sup>. Los precios de Bitcoin en Bitfinex están abiertos todos los días de la semana.

Para detectar si la introducción del mercado de futuros generó la caída en el precio de la criptomoneda se examina la relación entre los retornos de ambos activos. La Tabla 1 y la Tabla 2 exhiben un resumen estadístico de ambas series tanto en retornos como en nivel. Se puede observar que en la media el retorno es negativo tanto para el precio *spot* como para el precio en el mercado de futuros. La media de los retornos diarios en el precio *spot* es -0,522%, con un desvío estándar de

---

<sup>3</sup> <https://www.investing.com/crypto/bitcoin/btc-usd-historical-data> - Precio diario del BTC desde el 13 de diciembre de 2017 hasta 11 de noviembre de 2018. Se omitieron los días en donde el mercado de futuros estaba cerrado (sábados, domingos y feriados)

<sup>4</sup> <https://www.investing.com/crypto/bitcoin/bitcoin-futures> - Precio diario del BTC futuro desde el 13 de diciembre de 2017 hasta 11 de noviembre de 2018.

<sup>5</sup> <https://www.bitfinex.com/>

<sup>6</sup> CME Bitcoin futures horarios de trading son 8:30a.m. – 3:30 p.m. CT los lunes y 8:30 a.m. a 3:15 p.m. los martes a viernes.

5,88%. Los resultados son parecidos a los del mercado de futuros, que tiene una media de -0,596% y un desvío de 5,68%. El exceso de curtosis para ambos mercados indica que las distribuciones son leptocúrticas, y esto lo podemos ver a través de los rangos, por un lado, los retornos diarios llegan de -20,38% hasta 14,06% y por el otro de -18,96% hasta 13,72%. Al calcular la correlación de ambas series en nivel se obtiene un valor de 0,98. Llamativamente al calcular la correlación de ambas series en retornos se obtiene un valor de 0,53.



## 4. Metodología

El principal enfoque del presente trabajo se basa en los conceptos de cointegración, causalidad, modelos VECM y análisis de estudio de eventos a través de modelos de retornos normales y anormales.

### 4.1 Cointegración

Antes de identificar cualquier relación de sucesión entre las series se debe estudiar si estas cointegran. La medida de cointegración muestra si hay una relación de equilibrio a largo plazo entre las dos series para estimar un precio, a pesar de su desvío a corto plazo. En el estudio de cointegración hay distintas técnicas, de las cuales usaremos el modelo de Johansen (1991) y el de Granger (1987).

Antes de que se estableciera la cointegración en la econometría para poder realizar una regresión entre dos series era estadísticamente necesario que ambas fuesen estacionarias. En el caso de que no lo sean, se diferenciaban las series y se realizaba un modelo con las diferencias. Esto permite ver si hay un efecto significativo entre las variables, pero no permite trabajar con el modelo en niveles. La diferenciación resuelve el problema desde el punto de vista estadístico (permite correr un modelo entre las dos variables), pero desde el punto de vista económico no es posible modelar la eventual relación en niveles por la no estacionariedad. A partir del estudio de la cointegración es posible estimar una regresión en niveles entre variables de mismo orden de integración  $I(1)$  si los errores son  $I(0)$ .

Para verificar la no estacionariedad de las series se utilizará el test de Dickey-Fuller aumentado (ADF), que se puede describir de la siguiente forma:

$$\Delta Y_t = \alpha + \beta Y_{t-1} + \delta \Delta Y_{t-1} + \delta_2 \Delta Y_{t-2} + \dots + \delta_n \Delta Y_{t-n} \varepsilon_t \quad (1)$$

donde,  $\Delta Y$  es el cambio en el logaritmo tanto de la serie de Bitcoin (*spot* y futuros),  $\alpha$  es una constante y  $\beta$  es el coeficiente de la tendencia a largo plazo. Si el  $\beta = 0$ , la serie no es estacionaria. Cuando se determinen los rezagos en este estudio se utilizará el modelo MCE identificados por los criterios de información Aikaike (AIC), Hannan-Quinn o Bayesian (BIC). Siempre se utilizará el nivel 1% de valor crítico.

#### 4.1.1 Cointegración vía Engle - Granger

En el caso de la técnica de Granger, en un sistema con  $n$  variables diremos que están cointegradas si cada una de ellas es integrada de orden superior a 0 (no son estacionarias), ambas son del mismo orden y existe un vector de parámetros  $\beta$  tal que la combinación lineal de estas variables resulta ser integrada de orden inferior. Es decir, le reducimos un orden de integración. En la Tabla 3 del Apéndice puede observarse que en este caso tanto el *spot* como el futuro son de orden 1 (I(1)).

Una vez verificada la cointegración mediante la técnica de Granger podremos analizar la relación. Para realizar esto utilizaremos dos ejercicios de pronóstico, llamados *Future Prices Model (FPM)* y *Augmented Future Price Model (AFPM)*.

El ejercicio de *Future Prices Model* consiste básicamente en estimar un modelo en el cual el *spot* depende del futuro y viceversa. Lo que se hace es estimar la constante  $\alpha$  y la pendiente  $\sigma$ . Si  $\alpha = 0$  quiere decir que hay un supuesto de insesgadez, si  $\sigma = 1$  hay un supuesto de proporcionalidad.

$$S_{t+h/T} = \alpha + \sigma F_t^h \quad (2)$$

La técnica *Augmented Future Prices Model* es una versión aumentada del FPM. En este caso se descuenta la tasa implícita que supone el mercado de futuros por la tasa de interés de referencia, esto de

alguna manera incorpora el efecto que puede tener la tasa de interés.

$$S_{t+h/T} = \alpha + \gamma \frac{F_t^h}{(1+i^h)_t} \quad (3)$$

#### 4.1.2 El enfoque de Johansen - Juselius

El procedimiento de Johansen es una generalización multivariada del test de Dickey-Fuller.

Como se puede ver en la ecuación (1), en una prueba de Dickey — Fuller univariado, si el  $\beta = 0$  la serie tiene una raíz unitaria. En el caso multivariado se considera una generalización a  $n$  variables:

$$\Delta X_t = \phi_1 X_{t-1} - X_{t-1} + \varepsilon_t \quad (4)$$

$$\Delta X_t = (\phi_1 - I) X_{t-1} + \varepsilon_t \quad (5)$$

$$\Delta X_t = \pi X_{t-1} + \varepsilon_t \quad (6)$$

donde  $X_t$  y  $\varepsilon_t$  son vectores ( $n \times 1$ ),  $\phi$  es una matriz ( $n \times n$ ) de parámetros.  $I$  es la matriz ( $n \times n$ ) identidad, y  $\pi$  se define como  $(\phi - I)$ . Para evaluar la cointegración, Johansen (1991) propone analizar el rango de la matriz  $\pi$ . Si el rango de  $\pi$  es menor a  $n$  y mayor a 0 las variables cointegran debido a que hay alguna combinación lineal de las variables no estacionarias del sistema que resulta estacionaria. En el caso del BIVAR (sistema con dos variables) las posibilidades son  $r(\pi) = 0$ ,  $r(\pi) = 1$  (cointegración) y  $r(\pi) = 2$ .

## 4.2 Modelo de Corrección de Errores

Una vez confirmada la cointegración entre dos o más variables, se puede desarrollar un modelo de corrección de errores. La ventaja de estimar este modelo es que permitirá evaluar los efectos de corto y largo plazo que puede tener una variable sobre la otra. Este modelo de corrección de errores va a incorporar la dinámica en el modelo. Permite explicar los retornos logarítmicos de  $\Delta Y_t$  en función del movimiento en  $\Delta X_t$  y el desequilibrio pasado entre los niveles de  $Y_t$  y  $X_t$ .

$$\Delta Y_t = \beta_1 \Delta X_{t-1} + \beta_2 \Delta Y_{t-1} - (1 - \beta_3) \left[ Y_{t-1} - \frac{\beta_1 + \beta_2}{1 - \beta_3} X_{t-1} \right] + \varepsilon_t \quad (7)$$

$$\frac{\beta_1 + \beta_2}{1 - \beta_3} = K \quad (8)$$

$$(1 - \beta_3) = \alpha \quad (9)$$

donde  $(1 - \beta_3)$  mide la velocidad a la que va a ajustar la variable dependiente cada vez que se produzca un desequilibrio para retornar al equilibrio de largo plazo y  $K$  representa las consecuencias de largo plazo que tiene la variable dependiente en  $Y$ . El coeficiente  $\beta_1$  va a ser el efecto de corto plazo que tenga  $X$  en  $Y$ . El coeficiente  $\beta_2$  representa el ajuste a largo plazo que afecta directamente a  $\beta_1$ .

Por esto el MCE implica una modelación conjunta de la dinámica de corto y de largo plazo con variables en diferencia y en nivel.

## 5 Análisis y resultados

En esta sección se presentan algunos resultados para analizar la eficiencia de los modelos utilizados. Desarrollamos los algoritmos en lenguaje R utilizando las librerías correspondientes. Primero se realizó la prueba de cointegración que resultó positiva tanto vía Johansen como vía Engle y Granger.

### 5.1 Cointegración vía Engle y Granger

Para evaluar la dirección de la causalidad y la robustez comenzamos aplicando el método de Engle y Granger. Como el modelo requiere estacionariedad, las series temporales bivariadas las tuvimos que transformar en series de retornos. Se utilizó la prueba de Dickey – Fuller para evaluar la estacionariedad de las series. En todos los niveles los resultados muestran estacionariedad estadísticamente significativa al nivel de valor crítico de 1%. Aquí identificamos si los datos rezagados de la variable Y proporcionan alguna información sobre la variable X teniendo en cuenta también los datos anteriores de X. La librería seleccionada para realizar estos procedimientos se llama *urca*. Sin embargo, todavía no hay evidencia para concluir que estas series están cointegradas, ya que debemos probar estacionariedad en los residuos de la muestra. En la *tabla 5.1* podemos observar que hay evidencia para rechazarla hipótesis nula a partir de los distintos *tests* de raíz unitaria sobre los residuos y, por lo tanto, podemos asumir cointegración.

**Tabla 5.1**

Cointegración vía Engle & Granger – Serie en retornos

<b>Engle &amp; Granger Cointegration</b>	<b>P-value</b>
Augmented Dickey Fuller (ADF)	0.00374 ***
Phillips – Perron (PP)	0.0001 ***
Pantula, Gonzales - Farias and Fuller (PGFF)	0.0001 ***
Elliot, Rothenberg, and Stock (ERSD)	0.16636
Johansen Trace Test (JOT)	0.0001 ***
Schmidt and Phillips Rho (SPR)	0.0038 ***

Nota: Los asteriscos denotan el nivel de significatividad: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

El único método que no muestra evidencia para rechazar la hipótesis nula es el de Elliot, Rothenberg y Stock, los otros 5 concluyen que estamos en condiciones de rechazar a nivel 1%.

## 5.2 Cointegración vía Johansen

Si bien el procedimiento de Engle y Granger es efectivo de implementar tiene la desventaja de que nos obliga a decidir cual variable será la dependiente y cual la explicativa. Esto no es lo conveniente debido a que se podrían determinar de forma simultánea por lo que convendría estimarlos en forma conjunta.

Por esta razón se realizaron las pruebas de cointegración vía Johansen. A continuación, se mostrarán los resultados en la Tabla 5.2. Se utilizaron los datos en logaritmo para evitar problemas de normalidad en los errores. Para realizarlo se utilizó la librería *urca* de R *Studio* que genera el modelo. Se realiza esta prueba para tener una metodología alternativa de cointegración en donde se puede extender la idea uniecuacional del modelo de Engel y Granger a un sistema conjunto.

El nivel de rezagos utilizado fue 1 siguiendo los criterios *Aikaike* y *Bayes*. A partir de este nivel rechazo la hipótesis de ausencia de autocorrelación y podría asumir que se controla toda la dinámica de esta serie según el *serial.test*. El *normality.test* concluye que hay normalidad entre los errores, no hay problemas de asimetría ni de curtosis tal como se puede observar en la Tabla 9 en el Apéndice.

**Tabla 5.2**

Cointegración vía Johansen – Serie en logaritmo

<b>Result</b>	<b>Daily Co-integrated</b>
<b>Trace Test</b>	
$r = 0$	63.11 ***
<i>Critical Value 0.05 = 18.96, Critical Value 0.01 = 23.65</i>	
$r = > 1$	5.5
<i>Critical Value 0.05 = 12.25, Critical Value 0.01 = 16.26</i>	
<b>Eigen Test</b>	
$r = 0$	57.61 ***
<i>Critical Value 0.05 = 18.96, Critical Value 0.01 = 23.65</i>	
$r = > 1$	5.5
<i>Critical Value 0.05 = 12.25, Critical Value 0.01 = 16.26</i>	
<i>Lags</i>	2
<i>Observations</i>	128

Nota: Los asteriscos denotan el nivel de significatividad: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Tanto la prueba de traza como de autovalor concluyen que ambas series cointegran a un nivel de 1% de valor crítico. Esto significa que las series bivariadas son de orden de integración 1 (I (1)).

Logro rechazar la hipótesis nula de que el rango es igual a 0 (ausencia de cointegración) a nivel 1% con un estadístico de 57,61 y no rechazo la hipótesis nula de que el rango sea menor o igual a uno a nivel 1% que define que hay única relación de cointegración entre el *spot* y el contrato futuro.

### 5.3 Modelo de Corrección de Errores

El modelo fue aplicado a las series en logaritmo. De aquí vamos a obtener el término de corrección de errores, y luego la exogeneidad de las variables. Este análisis se realizó nuevamente mediante la librería *urca* y *dynlm* en *R Studio*. Los resultados en la Tabla 8 del Apéndice muestran que el término de corrección de errores no es significativo a ningún nivel de p-valor. Esto denota que no hay evidencia de un impacto de los precios *spot* sobre la serie de precios futuros. Sin embargo, realizando el análisis del impacto de la serie de futuros a los precios *spot* que se observan en la Tabla 9 del Apéndice, el término de corrección de errores es significativo a nivel 1%, con un coeficiente  $\alpha$  de 0,48 que indica la velocidad que van a tomar los precios futuros para retornar al equilibrio de largo plazo, es decir, cada vez que haya un desequilibrio de largo plazo, en un periodo (un día en nuestro modelo) el precio futuro va a ajustar en un 48% para retornar al equilibrio de largo plazo.

Pruebo la ausencia de autocorrelación en los errores para lograr el modelo que mejor ajusta en el precio de los futuros, y llego a la conclusión a partir de la Tabla 9 y Tabla 10 que figuran en el apéndice que el mejor modelo se da a partir del término de corrección de errores sumado al primer *lag* del spot que logra un nivel de significatividad de 10%. Con este modelo logro no rechazar la prueba de Breusch – Godfrey por lo que el modelo captura completamente la dinámica.

**Tabla 5.3**

Modelo de Corrección de Errores- Precios en logaritmo

<b>Vector Error Correction Model</b>	Daily
<b>Error Correction Term</b>	
Spot	-0.042 (0.953)
Futures	0.4877 *** (2.1e-09)
Correlation of Residuals	0.2836
<b>Residual Standard Error</b>	
Spot	0.2397
Futures	0.06955
<b>Adjusted R-squared</b>	
Spot	-0.01
Futures	0.148
<i>Lags</i>	<i>1</i>
<i>Observations</i>	<i>280</i>

Nota: Los asteriscos denotan el nivel de significatividad: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Los resultados resumidos en la tabla 5.3 muestran lo que remarcamos anteriormente: el coeficiente de corrección de errores denota el impacto de la serie futuros sobre los precios *spot* y no hay evidencia de que haya un impacto del mercado *spot* hacia el mercado de futuros a ningún nivel de significatividad.

El hecho de que el término de corrección de errores de una serie sea significativo y de la otra serie no Engle, Hendry y Richard (1983) llaman exogeneidad débil. Si una de las variables no ajusta cada vez que se produce un desequilibrio de largo plazo (el coeficiente del término de corrección de errores no es significativo) y la otra sí, es débilmente exógena. Esto quiere decir que la única ecuación válida es la que tiene como variable dependiente la serie *spot*, porque es la que responde a los cambios de las otras variables y no a la inversa.

En función de justificar los resultados otorgados por la estimación del MCE realizamos distintas pruebas que comprueban los resultados anteriores obtenidos:

1. Causalidad en sentido de Granger: Las conclusiones que obtenemos a partir de este *test* son que tenemos evidencia que la serie futuros precede temporalmente a la serie *spot* a un nivel de significatividad de 1%. Contrario a esto, tal como vimos en el análisis previo, los resultados arrojan que no tenemos evidencia de que la serie de *spot* tenga causalidad en sentido de Granger sobre la serie futuros.

**Table 5.7**  
Causalidad en sentido de Granger – Serie en retornos

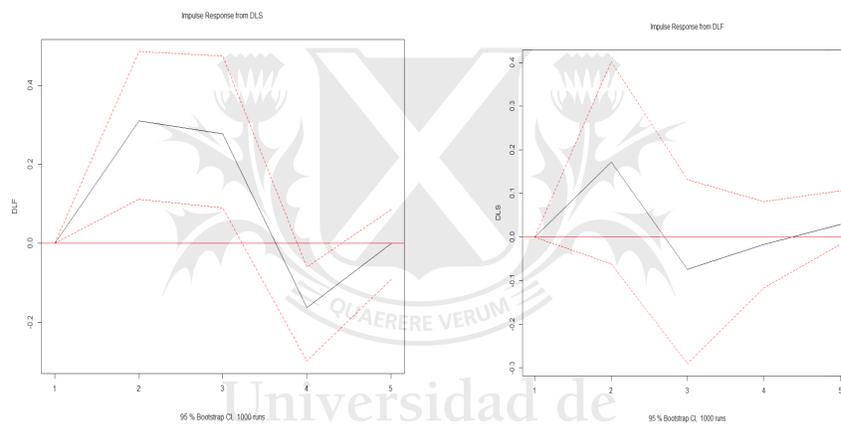
<b>Granger Causality</b>	Daily
Spot does Granger cause Futures	0.5875
Futures does not Granger cause Spot	0.0002 ***
<i>Lags</i>	2
<i>Observations</i>	280

Note: Los asteriscos denotan el nivel de significatividad: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

2. Función Impulso – Respuesta: En esta simulación realizamos un impulso o *shock* en una de las dos variables y vemos cómo reacciona la otra ante este shock. En los gráficos veremos por un lado como un shock en la serie futuros afecta en mayor dimensión a la serie *spot* y tarda más días en converger hacia el cero.

**Tabla 5.8**

Función Impulso – Respuesta – Serie en logaritmos.



Universidad de  
San Andrés

## 6 Conclusión

Ambos modelos de cointegración llegan a la conclusión de que las series están cointegradas tal como se prevé que suceda en una serie de futuros y *spot* de un activo.

El análisis de la relación entre el mercado *spot*-futuro del Bitcoin posterior al análisis de cointegración sugieren que los precios futuros suceden el precio *spot* de la criptomoneda, causándola en sentido de Granger. Mediante la función impulso - respuesta y a través del estudio de causalidad en sentido de Granger obtenemos estos resultados con una gran contundencia.

Debido a que la series no son estacionarias y habría que trabajar con el modelo en retornos no hicimos un análisis de vectores autorregresivos ya que no causaría el mismo efecto enriquecedor para el trabajo. Es mejor utilizar los modelos de cointegración de Johansen y el modelo de corrección de errores para sacar conclusiones. Estos también siguen la misma línea que los modelos nombrados anteriormente, con la misma contundencia. En el modelo de corrección de errores podemos ver significatividad muy fuerte utilizando los precios futuros como variable dependiente, y todo lo contrario al utilizar a la serie *spot* como variable dependiente. Con un coeficiente  $\alpha$  de 0,48 se indica la velocidad que van a tomar los precios *spot* para retornar al equilibrio de largo plazo, es decir, cada vez que el promedio de los precios futuros sea alto, los precios *spot* ajustaran en un periodo (un día en nuestro caso) un 48% hacia el nivel de precios de la serie de futuros.

Comparando con los métodos de los autores analizados podemos apoyarnos en Karkkainen, T. (2018) que concluye que a mayor frecuencia el nivel de significatividad es cada vez más determinante a favor de los futuros causando los precios *spot*. Sin embargo, disentimos en los resultados obtenidos utilizando una serie de precios diaria. A través del análisis empírico realizado por Hattori, T. y Ishida, R. (2020), ellos concluyen que no hay evidencia de una relación entre los dos mercados durante el

periodo analizado debido a que el único momento en donde hay una oportunidad de arbitraje es cuando se realizó el caída más vertiginosa. Luego a lo largo de toda la muestra se estabiliza. Esto se puede observar en la figura 2.

Lo que comprobamos a lo largo de este trabajo es que efectivamente hay una relación entre los dos activos. Por otro lado, observamos que la relación es asimétrica: el precio *spot* se ve más influenciado por los futuros que los precios futuros por el *spot*. Estos resultados me permitirían plantear una conjetura para un trabajo posterior con otra metodología sobre la posibilidad si la introducción de futuros podría ser un detonante de la corrección.



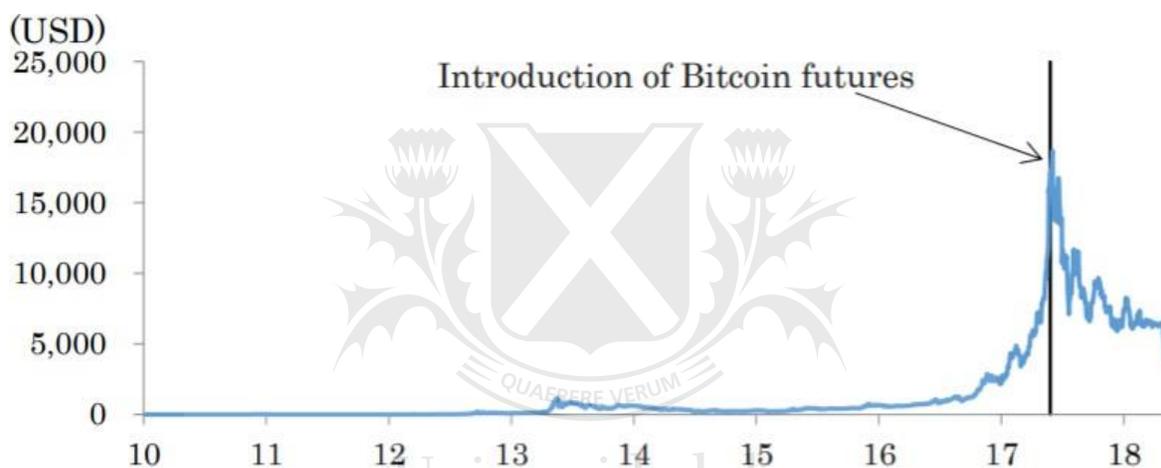
## 7. Bibliografía

1. Ahumada, H., Cornejo, M., 2015. Explaining commodity prices by a cointegrated time series-cross section model. *Empir. Econ.* 48, 1667–1690.
2. Baur, G., Dimpfl, T. 2018. Price Discovery in Bitcoin *Spot* or Futures?
3. Clements, M.P., Hendry, D.F., 1998. *Forecasting Economic Time Series*. Cambridge University Press, Cambridge.
4. Cox, J.C., Ingersoll, J.E., Ross, S.A., 1981. The relation between forward prices and futures prices. *J. Financ. Econ.* 9, 321–346.
5. Engle, R.F., Hendry, D.F., Richard, J.-F., 1983. Exogeneity. *Econometrica* 277–304.
6. Giudici, P., Pagnotoni, P., 2019a. Vector error correction models to measure connectedness of Bitcoin exchange markets.
7. Hattori, T., Ishida, R., 2020. The relationship between arbitrage in futures and *Spot*
8. Karkkainen, T., 2018. Price discovery in the Bitcoin futures and cash markets markets and Bitcoin price movements: Evidence from the Bitcoin markets.
9. Pagnotoni, P., 2019. Neural network models for bitcoin option pricing. *Front. Artif. Intel.* 2(5), 1–9.
10. Reichsfeld, D.A., Roache, S.K., 2011. Do Commodity Futures Help Forecast Spot Prices? International Monetary Fund.
11. Yang, J., Bessler, D.A., Leatham, D.J., 2001. Asset storability and price discovery in commodity futures markets: A new look. *J. Futures Markets* 21, 279–300.

## 8. Anexo

**FIGURA 1: Grafico de precios históricos.**

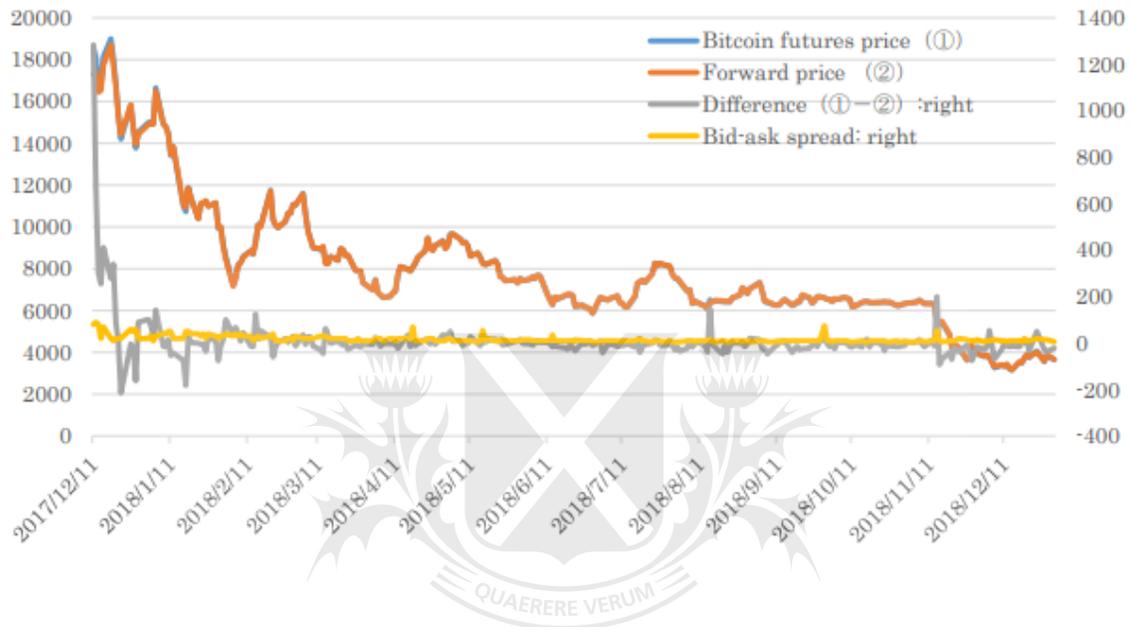
Gráfico de precios históricos de Bitcoin señalando el momento de entrada del mercado de futuros.



Elaboración de *Bloomberg* en base a datos de *Investing.com*

## FIGURA 2: Arbitrariedad en los mercados.

Gráfico utilizado por Hatori, T. y Ishida, R. (2020) para mostrar los momentos de arbitraje entre ambos mercados.



Note: This figure describes the Bitcoin futures, forward, differential, and bid-ask prices. The sample period is from December 11, 2017 to December 28, 2018. The timing of futures and forward prices is 4 PM ET and the data frequency is daily.

San Andrés

**FIGURA 3: Serie de precios *spot* y futuros.**



Elaboración propia en base a datos obtenidos de *investing.com*

Universidad de  
**San Andrés**

**Tabla 1: Resumen estadístico**

Precios diarios *spot* de Bitcoin y de los futuros por el CBOE.

	Bitcoin <i>spot</i>	Bitcoin Futuros
Media	10425	10429
Mediana	9427	9445
Desvío Estándar	2912	2972
Max.	18971	19595
25%	8398	8505
75%	11414	11429
Min	6610	6630
Curtosis	0.3	0.55
Simetría	1.06	1.14
Observaciones	280	280

**Tabla 2: Resumen estadístico**

Retornos diarios de precio *spot* de Bitcoin y de los futuros por el CBOE.

	Retorno Bitcoin <i>spot</i>	Retorno Bitcoin Futuros
Media	-0.522%	-0.596%
Mediana	-0.308%	-0.484%
Desvío Estándar	6%	6%
Max.	14.056%	13.729%
25%	-3.942%	-0.456%
75%	3.35%	3.519%
Min	-20.381%	-18.961%
Curtosis	0.66	-0.15
Simetría	-0.38	-0.15
Observaciones	279	279

Elaboración propia en base a datos obtenidos de *investing.com*

### Tabla 3: Cointegración vía Engle y Granger

Output de R Studio al realizar el análisis de cointegración vía Engle y Granger.

$$Y[i] = 0.9986 X[i] + 0.0203 + R[i], R[i] = 0.3884 R[i-1] + \text{eps}[i], \text{eps} \sim N(0, 0.0470^2)$$

(0.0161)                      (0.1505)                      (0.0826)

$$R[128] = 0.0415 \quad (t = 0.819)$$

#### Unit Root Tests of Residuals

	Statistic	p-value
Augmented Dickey Fuller (ADF)	-4.668	0.00374
Phillips-Perron (PP)	-81.622	0.00010
Pantula, Gonzales-Farias and Fuller (PGFF)	0.362	0.00010
Elliott, Rothenberg and Stock DF-GLS (ERSD)	-1.999	0.16636
Johansen's Trace Test (JOT)	-44.448	0.00010
Schmidt and Phillips Rho (SPR)	-43.797	0.00383

#### Variances

SD(diff(x))	=	0.077995
SD(diff(y))	=	0.081662
SD(diff(residuals))	=	0.057194
SD(residuals)	=	0.050664
SD(innovations)	=	0.047039

Half life	=	0.732842
R[last]	=	0.041500 (t=0.82)



Universidad de  
**San Andrés**

## Tabla 4: Cointegración vía Johansen

Output de R Studio al realizar el análisis de cointegración vía Johansen.

Test de Autovalor.

```
#####  
# Johansen-Procedure #  
#####  
  
Test type: maximal eigenvalue statistic (lambda max) , with linear trend in cointegration  
  
Eigenvalues (lambda):  
[1] 3.669537e-01 4.268542e-02 2.992398e-17  
  
Values of teststatistic and critical values of test:  
  
      test 10pct 5pct 1pct  
r <= 1 | 5.50 10.49 12.25 16.26  
r = 0  | 57.61 16.85 18.96 23.65  
  
Eigenvectors, normalised to first column:  
(These are the cointegration relations)  
  
      LS.l1      LF.l1      trend.l1  
LS.l1  1.0000000000 1.000000000 1.000000000  
LF.l1  -0.9961754674 0.641643377 -3.47433447  
trend.l1 0.0001192593 0.006614515 -0.04695185  
  
weights w:  
(This is the loading matrix)  
  
      LS.l1      LF.l1      trend.l1  
LS.d 0.002730001 -0.04436953 -4.303524e-17  
LF.d 0.723590675 -0.02396820 -7.284363e-18
```

## Tabla 5: Cointegración vía Johansen

Test de Traza.

```
#####  
# Johansen-Procedure #  
#####  
  
Test type: trace statistic , with linear trend in cointegration  
  
Eigenvalues (lambda):  
[1] 3.669537e-01 4.268542e-02 2.992398e-17  
  
Values of teststatistic and critical values of test:  
  
      test 10pct 5pct 1pct  
r <= 1 | 5.50 10.49 12.25 16.26  
r = 0  | 63.11 22.76 25.32 30.45  
  
Eigenvectors, normalised to first column:  
(These are the cointegration relations)  
  
      LS.l1      LF.l1      trend.l1  
LS.l1  1.0000000000 1.000000000 1.000000000  
LF.l1  -0.9961754674 0.641643377 -3.47433447  
trend.l1 0.0001192593 0.006614515 -0.04695185  
  
weights w:  
(This is the loading matrix)  
  
      LS.l1      LF.l1      trend.l1  
LS.d 0.002730001 -0.04436953 -4.303524e-17  
LF.d 0.723590675 -0.02396820 -7.284363e-18
```

## Tabla 6: Causalidad en sentido de Granger

Análisis de causalidad en sentido de Granger.

Utilizando Spot como variable dependiente.

```
Granger causality test
test ([D_ls]LD.lf=0) ([D_ls]L1._ce1=0)

( 1) [D_ls]LD.lf = 0
( 2) [D_ls]L._ce1 = 0

      chi2( 2) =    16.89
Prob > chi2 =    0.0002
```

Utilizando los futuros como variable dependiente.

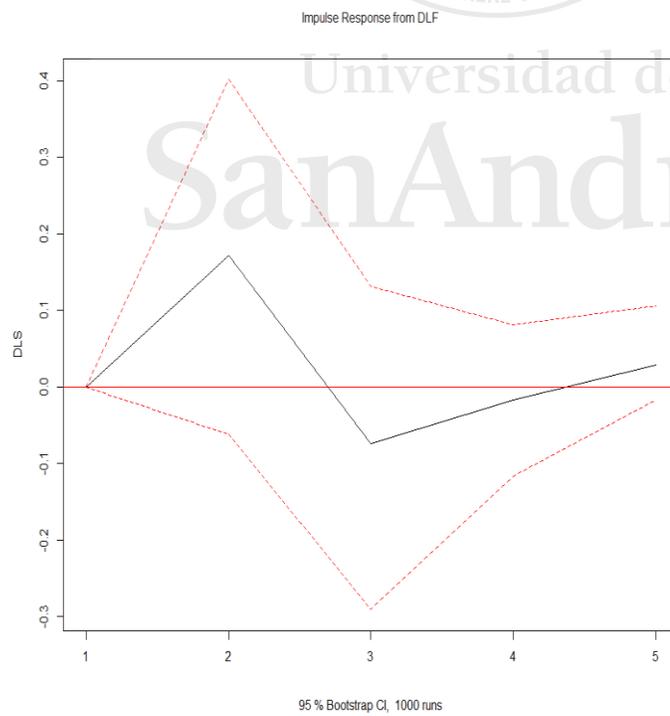
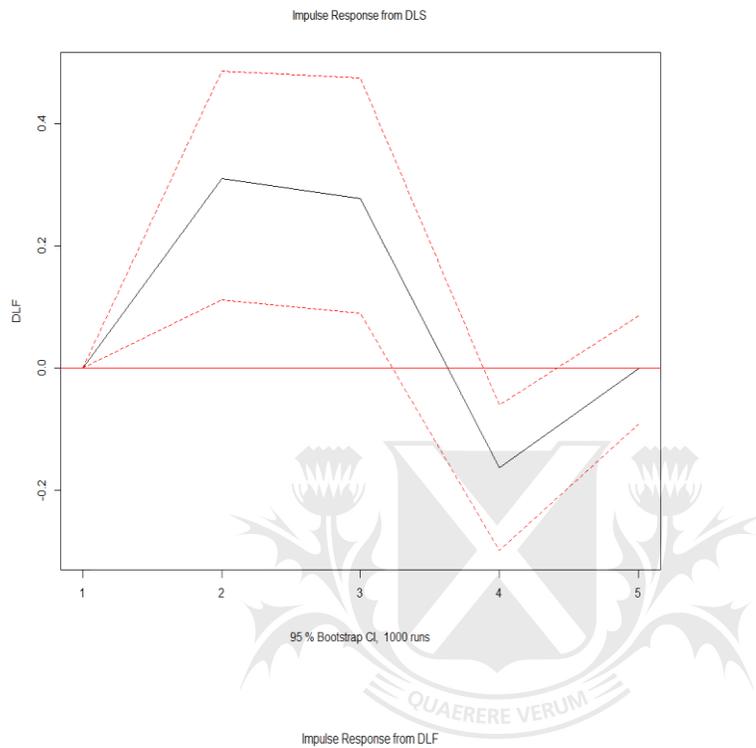
```
Granger causality test
. test ([D_lf]LD.ls=0) ([D_lf]L1._ce1=0)
> n 2 lags

( 1) [D_lf]LD.ls = 0
( 2) [D_lf]L._ce1 = 0

      chi2( 2) =    1.06
Prob > chi2 =    0.5875
```

## Tabla 7: Función Impulso - Respuesta

Función Impulso – Respuesta de las series en retornos a partir del análisis VAR.



## Tabla 8: Modelo de Corrección de Errores

Output de R Studio para el análisis de Modelo de Corrección de Errores (MCE).

Utilizando precios futuros como variable dependiente.

```
Time series regression with "ts" data:
Start = 2017(14), End = 2018(11)

Call:
dynlm(formula = d(LF) ~ L(tce) + L(d(LS)) + L(d(LF), 1), data = yyy)

Residuals:
    Min       1Q   Median       3Q      Max
-0.19224 -0.02430 -0.00293  0.01988  1.00627

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0002691  0.0045568   0.059   0.953
L(tce)       -0.0437891  0.0806945  -0.543   0.588
L(d(LS))     0.0323991  0.0826945   0.392   0.696
L(d(LF), 1) -0.0126688  0.0902532  -0.140   0.888

Residual standard error: 0.07597 on 274 degrees of freedom
Multiple R-squared:  0.002617, Adjusted R-squared:  -0.008303
F-statistic: 0.2397 on 3 and 274 DF, p-value: 0.8686
```

## Tabla 9: Modelo de Corrección de Errores

Modelo más explicativo utilizando un rezago.

```
Time series regression with "ts" data:
Start = 2017(14), End = 2018(11)

Call:
dynlm(formula = d(LS) ~ L(tce) + L(d(LF)) + L(d(LS), 1), data = yyy)

Residuals:
    Min       1Q   Median       3Q      Max
-0.21989 -0.01999  0.00161  0.01626  0.92921

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.0001768  0.0041713  -0.042   0.9662
L(tce)       0.4877351  0.0738680   6.603 2.09e-10 ***
L(d(LF))     -0.1090918  0.0826180  -1.320   0.1878
L(d(LS), 1)  0.1333644  0.0756988   1.762   0.0792 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06955 on 274 degrees of freedom
Multiple R-squared:  0.1579, Adjusted R-squared:  0.1487
F-statistic: 17.13 on 3 and 274 DF, p-value: 3.185e-10
```

## Tabla 10: *Test de Breusch - Godfrey*

Modelo para comprobar la ausencia de autocorrelación en los residuos.

Test de Breusch – Godfrey de la serie sin rezagos.

```
Breusch-Godfrey test for serial correlation of order up to 2  
data: mce  
LM test = 8.4325, df = 2, p-value = 0.01475
```

Test de Breusch – Godfrey de la serie con un rezago.

```
Breusch-Godfrey test for serial correlation of order up to 1  
data: mce  
LM test = 0.60854, df = 1, p-value = 0.4353
```

