

An analysis of price discovery between Bitcoin futures and spot markets

Burcu Kapar

Jose Olmo*

American University in Dubai

University of Southampton

Abstract

This paper analyzes the Bitcoin price discovery process. We collect data on futures and spot prices for the period December 2017 to May 2018 and compute Hasbrouck's information share and Gonzalo and Granger's common factor component to quantify the contribution of each market to the price discovery process. Both measures coincide in suggesting that the Bitcoin futures market dominates the price discovery process. We also find that both prices are driven by a common factor that is given by a weighted combination of the futures and spot market. Finally, we observe that deviations from the equilibrium condition equating the futures and spot log-price have predictive ability for the return on the Bitcoin spot price but not on the futures price.

Keywords: Bitcoin, information share, price discovery, cointegration, permanent-transitory decomposition

*Corresponding author: Department of Economics, University of Southampton, Highfield Lane, Southampton, SO17 1BH. United Kingdom. E-mail: j.b.olmo@soton.ac.uk

1 Introduction

In an informationally efficient market asset prices reflect information available to market participants. Price discovery is the process by which markets incorporate this information to arrive at equilibrium asset prices. If a single financial asset or multiple highly related financial assets are traded on more than one market, each market may be involved in the price discovery process, but the one that provides a combination of the greatest liquidity, lowest execution costs and greatest leverage opportunities should dominate. In short, price discovery studies attempt to answer the following questions: "Which market moves first?" and "Which market product moves closer to the intrinsic value?" The answer to this question is of great interest to market participants and regulators.

Whether the spot or the futures market is the center of price discovery for different types of asset classes has been discussed for a long time in the literature. Theoretically, futures markets are better positioned as leaders in the price discovery process. The reason for this is the inherent leverage, low transaction costs, and the absence of any short-selling constraints. Stein (1961) showed that futures and spot prices for a given commodity are determined simultaneously. Garbade and Silber (1983) develop a model of simultaneous price dynamics in which they establish that price discovery takes place in the market with the highest number of participants. More recently, the price discovery research has focused on microstructure models and on methods to measure it. This line of literature applies two methodologies: the Information Share of Hasbrouck (1995) and the Permanent-Transitory decomposition of Gonzalo and Granger (1995). The prevailing view in this literature is that across different assets future markets dominate the price discovery process, e.g., markets for equities and stock indices (Stoll and Whaley, 1990), markets for carbon emissions (Mizrach and Otsubo, 2014), markets for commodities (Figuerola-Ferretti and Gonzalo, 2010), markets for VIX indices (Chen and Tsai, 2017) and markets for currency futures (Tse et al., 2006), among a few others. However some of the papers have advocated opposite results for commodity markets (Chiou-Wei et al., 2013) or bidirectional feedback between both markets for commodities (Kaufmann and Ullman, 2009).

The purpose of this article is to reconsider the market leadership question for the recently created class of assets given by cryptocurrencies, and in particular, we focus on studying the price discovery process for Bitcoin. Bitcoin future contracts have been traded in regulated exchanges since December 2017. There are two related articles that study the price discovery process for cryptocurrencies. Brauneis and Mestel (2018) check the efficiency of 73 crypto-currencies with a market capitalization over USD 1 million between 31 August 2015 and 30 November

2017. These authors find that Bitcoin is the most efficient cryptocurrency and hence the least predictable. Their study only focuses on spot markets and, hence, the price discovery analysis involves comparison of different cryptocurrencies and not of the same cryptocurrency on different (spot and future) markets as we do here. Similarly, Brandvold et al. (2015) investigate the price discovery process for cryptocurrencies across different exchanges but not between the spot and future markets. These authors conclude that those exchanges with highest traded volume are the prominent price leaders in terms of information share. Smaller exchanges provide less information and usually follow the market with a lag.

An important exception is Baur and Dimpfl (2018). These authors analyze price discovery between Bitcoin spot and futures prices with high frequency data between 12 December 2017 and 26 March 2018 using both CME and CBOE futures contracts. Using an econometric methodology similar to us, their study concludes that price discovery takes place in the spot market, not in the futures market. This result is at odds with the price discovery literature on financial markets and also with the findings of our current paper. In line with the literature, we find that the futures market drives the price discovery process. Our results are robust across price discovery measures such as Gonzalo and Granger (1995) and Hasbrouck (1995) information share. The discrepancy between our results and those of Baur and Dimpfl (2018) may be due to the choice of data frequency. Whereas we analyze daily data on Bitcoin futures and spot prices these authors entertain intradaily data at a five-minute frequency.

2 Price discovery measures

If a single security is traded on parallel markets the corresponding prices share a common component, the fundamental value or efficient price. Relevant information can potentially enter the price of the asset traded in different markets at a different pace. Market microstructure research tries to identify where exactly price discovery takes place and which market incorporates relevant information first and in so doing takes informational leadership. Research in this field also tries to clarify the relative share a single market has in the process of price discovery.

Let s_t denote the log of the spot price at time t and $\Delta s_t = s_t - s_{t-1}$ the corresponding log return. Similarly, f_t is the log of the future price and $\Delta f_t = f_t - f_{t-1}$ is the corresponding log-return. Let $y_t = (f_t, s_t)'$ and $\Delta y_t = (\Delta f_t, \Delta s_t)'$. The VECM is given by

$$\Delta y_t = \alpha z_{t-1} + \sum_{i=1}^p \Gamma_i \Delta y_{t-i} + \epsilon_t. \quad (1)$$

The equilibrium condition is characterized by the error correction variable $z_t = \beta' y_t$ with β the cointegrating vector; α is the vector of loadings that measures the speed of adjustment of the departures from equilibrium z_{t-1} to the dynamics of the vector Δy_t . The matrices Γ_i capture the short-run Granger-casual relationships between the variables in first differences. The vector $\varepsilon_t = (\varepsilon_t^f, \varepsilon_t^s)'$ is a zero-mean vector of serially uncorrelated disturbances with covariance matrix denoted by Ω with variance terms σ_f^2 and σ_s^2 , respectively.

2.1 The Common Factor Component Model

Both futures and spot markets are disturbed by the arrival of new information such that the relationship between both prices moves away from long-term equilibrium. The CFC model uncovers a common factor that drives the long-term dynamics of both prices and that is interpreted as the efficient price that contains the fundamental value of the asset. Using Gonzalo and Granger (1995), we define the common factor as $g_t = \alpha'_\perp y_t$, with α_\perp a vector satisfying $\alpha'_\perp \alpha = 0$. The corresponding permanent-transitory decomposition of the process y_t is

$$y_t = A_1 g_t + A_2 z_t, \quad (2)$$

with $A_1 = \beta_\perp (\alpha'_\perp \beta_\perp)^{-1}$ and $A_2 = \alpha (\beta' \alpha)^{-1}$, where β_\perp is a vector satisfying $\beta'_\perp \beta = 0$.

The presence of cointegration between the spot and future prices also allows us to measure the contribution of each of the two markets to the common factor. These contributions are determined by the speed of adjustment parameter $\alpha = (\alpha_f, \alpha_s)$ in (1), such that

$$CFC_f = \frac{\alpha_s}{\alpha_s - \alpha_f}; CFC_s = \frac{-\alpha_f}{\alpha_s - \alpha_f} = 1 - CFC_f. \quad (3)$$

The market which dominates the process of price discovery will have a small α and thus a large CFC, while on the other hand, the market which predominantly follows in the process of price discovery will have a large α and thus a small CFC.

2.2 Hasbrouck Information Share Model

The information share associated with a particular market is defined as the proportional contribution of that market's innovations to the innovation in the common efficient price. The assumption

of covariance stationarity of the vector Δy_t implies the following vector moving average (VMA):

$$\Delta y_t = \Psi(L)\varepsilon_t. \quad (4)$$

$\Psi(1)$ is the sum of the moving average coefficients of the VMA representation. All of the rows of $\Psi(1)$ are identical. Intuitively, $\Psi(1)\varepsilon_t$ constitutes the long-run impact of a disturbance on each of the prices. That the rows are identical implies that this long-run impact is the same for all prices. Let $\psi = (\psi_f, \psi_s)$ denote the common row vector of $\Psi(1)$.

The increment $\psi\varepsilon_t$ is the component of the price change that is permanently impounded into the security price and is presumably due to new information. The variance of this term is $\psi\Omega\psi'$. If Ω is diagonal, then $\psi\Omega\psi'$ will consist of two terms, each of which represents the contribution to the random-walk innovation from the futures or spot market. The proportion of this for the futures market relative to the total variance of $\psi\varepsilon_t$ is defined by Hasbrouck (1995) as the futures market information share IS^f . Using the decomposition of $\Psi(1)$ in Johansen (1991) and defining $\alpha_\perp = (\gamma_f, \gamma_s)'$, we note that

$$IS^f = \frac{\gamma_f^2 \sigma_f^2}{\gamma_f^2 \sigma_f^2 + \gamma_s^2 \sigma_s^2}, \text{ and } IS^s = 1 - IS^f. \quad (5)$$

3 Dataset and Empirical Results

3.1 Bitcoin Futures Contracts

Bitcoin futures contract represents an agreement to buy or sell Bitcoin at a given point in time in the future. Currently, there are two publicly regulated exchanges where Bitcoin futures contracts are traded such as Chicago Board Options Exchange (CBOE) since December 10, 2017 and Chicago Mercantile Exchange & Chicago Board of Trade (CME) since December 18, 2017. The underlying asset for CBOE futures contracts is Gemini Exchange auction price for Bitcoin in U.S. dollar and for CME futures contracts is Bitcoin reference rate derived from an aggregate of major exchanges in U.S. dollar. Both contracts are cash-settled. Margin requirements for the contracts are high as Bitcoin prices are highly volatile and the market is immature yet. CBOE Bitcoin futures contracts have 40% margin requirement and CME contracts have 35% margin requirement.

Our dataset of Bitcoin spot and future prices range from 12 December 2017 to 16 May 2018 (124 daily observations). For Bitcoin spot price, we use Coindesk Bitcoin USD Price Index, a

simple average of global Bitcoin/USD exchange prices. It is expressed as the midpoint of the bid/ask spread across a number of global exchanges meeting certain minimum criteria with regard to minimum trade size, trading volume and others. For Bitcoin Future Price, we use Chicago Merchantile Exchange & Chicago Board of Trade (CME) contracts. The reason for preferring CME contracts over CBOE contracts is due to data availability. Further, by using Coindesk Bitcoin USD Price Index as a post price, we use a series that is very close to CME contracts reference rate. Both spot and future price datasets are publicly available.

3.2 Empirical Results

Standard Dickey-Fuller tests show that the series of Bitcoin spot and future prices are both unit roots. Model (1) is fitted to our sample of daily log prices from December 2017 to May 2018. We have implemented different specifications of the VECM in (1). Information criteria such as AIC, BIC and HQC and misspecification tests such as lagrange multiplier tests are favourable to the model specification given by five lags. For consistency across the paper we retain this model for the subsequent analyses.

The estimation of model (1) confirms empirically the equilibrium condition given by the vector $\beta = (1, -1)'$ and defining the stationary variable $z_t = f_t - s_t$. In equilibrium this variable is zero, however, in practice, both markets do not react simultaneously to market news, hence, the variable z_t can be different from zero. The estimates of the vector α are (2.02, 3.17) with standard errors (1.39, 1.40), respectively. This vector contains the long-term causality from the vector z_t to the returns on Bitcoin future and spot markets. The lack of statistical significance of α_1 suggests that departures from equilibrium between the futures and spot markets have predictive ability on the returns on the Bitcoin spot market but not on the futures market. This is evidence of long-term exogeneity of the futures market and endogeneity of the Bitcoin spot market. Furthermore, the positive and significant value of α_2 suggests that positive departures from the equilibrium condition $f_t - s_t = 0$ are reflected in increases in the spot price next period. That is, when the market is in contango we can expect increases in the spot price in the next period. In contrast, when the market is in backwardation, $f_t < s_t$, the VECM suggests a fall in spot prices to correct departures from equilibrium. These empirical findings are in line with most of the empirical literature on price discovery between futures and spot prices for equity and commodity markets that highlight the leading role of the futures market.

The above model analyzing the long-run relationship between futures and spot prices can be further exploited to disentangle the Bitcoin price discovery process. For our sample, we obtain

$\alpha_{\perp} = (-0.84, 0.53)'$ and $A_1 = (-3.28, -3.28)$. These estimates characterize the permanent and transitory component of both futures and spot markets. In particular, the permanent component driving both markets is given by a common factor $g_t = -0.84f_t + 0.53s_t$. Furthermore, the Bitcoin future price can be decomposed as $f_t = -3.28g_t + u_t^f$ and the spot price as $s_t = -3.28g_t + u_t^s$; u_t^f and u_t^s are stationary variables capturing the transitory components. The common permanent component given by $-3.28g_t$ suggests that the fundamental value of Bitcoin futures and spot prices is the same.

To conclude the analysis we also compute the information share of Hasbrouck (1995) and the CFC of Gonzalo and Granger (1995). For Hasbrouck (1995)'s approach, we find that Bitcoin futures prices have 88.7% information share and spot prices have 11.3% information share. Similarly, according to Gonzalo and Granger (1995)'s approach, the information share of futures market is 2.75 and of the spot market is -1.75 . Both sets of estimates clearly reflect the leadership of the Bitcoin futures markets with respect to the spot market.

References

- [1] Baur, Dirk G. and T. Dimpfl (2018) Price Discovery in Bitcoin Spot or Futures? Available at SSRN: <https://ssrn.com/abstract=3171464>.
- [2] Brandvold, M., P. Molnr, K. Vagstad and O.C.A. Valstad (2015) Price Discovery on Bitcoin Exchanges. *Journal of International Financial Markets, Institutions and Money* 36, pp. 18-35.
- [3] Brauneis, A. and R. Mestel (2018) Price Discovery of Cryptocurrencies: Bitcoin and Beyond. *Economics Letters* 165, pp. 58-61.
- [4] Chen, Y. and W. Tsai (2017) Determinants of Price Discovery in the VIX Futures Market. *Journal of Empirical Finance* 43, pp. 59-73.
- [5] Chiou-Wei, S.Z., S.C. Linn and Z. Zhu (2013) The Response of US Natural Gas Futures and Spot Prices to Storage Change Surprises: Fundamental Information and the Effect of Escalating Physical Gas Production. *Journal of International Money and Finance*, 42, pp. 156-173.
- [6] Figuerola-Ferretti, I. and J. Gonzalo (2010) Modelling and Measuring Price Discovery in Commodity Markets. *Journal of Econometrics* 158(1), pp. 95-107.

- [7] Garbade, K.D., and W. L. Silber (1983) Price Movements and Price Discovery in Futures and Cash Markets. *Review of Economics and Statistics* 65, pp. 289-297.
- [8] Gonzalo, J., and C. W. J. Granger (1995) Estimation of Common Long-Memory Components in Cointegrated Systems. *Journal of Business and Economic Statistics* 13, pp. 27-36.
- [9] Hasbrouck, J. (1995) One Security, Many Markets: Determining the Contributions to Price Discovery. *Journal of Finance* 50, pp.1175-1199.
- [10] Johansen, S. (1991) Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59, 1551-1580.
- [11] Kaufmann, R. and B. Ullman (2009) Oil Prices, Speculation, and Fundamentals: Interpreting Causal Relations among Spot and Futures Prices. *Energy Economics* 31(4), pp. 550-558.
- [12] Mizrach, B., and Y. Otsubo (2014) The Market Microstructure of the European Climate Exchange. *Journal of Banking and Finance* 39 (1), pp. 107-116.
- [13] Stein, J.L. (1961) The Simultaneous Determination of Spot and Future Prices. *American Economic Review* 51, pp. 1012-1025.
- [14] Stoll, H. R., and R. E. Whaley (1990) The Dynamics of Stock Index and Stock Index Futures Returns. *Journal of Financial and Quantitative Analysis* 25, pp. 441-468.
- [15] Tse, Y., J. Xiang and J. Fung (2006) Price Discovery in the Foreign Exchange Futures Market. *Journal of Futures Markets* 26 (11), pp. 1131 - 1143.