Detecting Information Flows in Markets*

Matthew Anderson† Jeff L. McMullin‡

November 30, 2018

Abstract

Using transfer entropy (TE), a measure from information theory, we detect sub-minute, directed information flows across prices. Examining data simulated with known levels of information transfer, we first document the properties of a TE estimator. Next, to determine if TE estimates capture information flows between fundamentally related markets, we estimate TE between equity options and the underlying equity. We find TE is generally strongest for at-the-money options, lower for in-the-money options, and near zero for out-of-the-money options. We detect at what timescale these information flows are the strongest and find TE peaks at timescales between 8 and 14 seconds. We find peak TE is higher when more news articles are published about the firm and when institutional investor attention to the firm is higher. Overall, our evidence supports using TE to measure information flows between prices, and develops a method for future research into price discovery and market information networks.

Keywords: price discovery, information transfer, information theory, transfer entropy
JEL classification: G12, G14, M40, M21

*This project has benefited from discussions with Glenn Lawyer and Ghazale Haratinezhad Torbati. The figures in the paper are best viewed in color.

†School of Informatics, Computing, and Engineering, Indiana University, andersmw@indiana.edu
‡Kelley School of Business, Indiana University, jemcmull@indiana.edu.

1 Introduction

Prices contain information and change as new information arrives. This fundamental assumption pervades modern economic thought. Thus, understanding the price discovery process is paramount in economic inquiry and driven by our ability to measure the process. An important part of this process is how information flows from one price to another price. In an efficient market, this information transfer is posited to occur rapidly, if not instantaneously. However, price changes at small time intervals appear to be noisy, martingale processes. Thus, measuring directed information flow requires a method that detects the transfer amid the noise. Building on research in the information theory field, we propose a measure of sub-minute information transfers across security markets and empirically examine the information flows between equity options and the underlying equity to assess the validity of this measure.

Although we assume prices reflect information, rarely if ever do we know the entire set of information prices reflect. Moreover, ascribing price changes to a specific piece of information is difficult. In contrast to theories that markets efficiently impound information into price within seconds of the information arrival, commonly used methods to examine information's impact on price typically examines price changes over wide windows. For firms experiencing common information events, long-window event studies examine the price response over many days or a few months (Kothari and Warner 1997; Barber and Lyon 1997), and short-window event studies examine the change in price over a few days. Prior studies use this short-window approach to determine how a firm's disclosures contemporaneously impact peers' security prices (e.g., Foster 1981 and Baginski 1987). Other studies examine lead-lag connections with weekly returns (Hou 2007). Examining price changes from the start to end of the window may capture only a fraction of the information flow as they ignore the price movements between the beginning and end of the window. In other words, these methods may be missing much of the price adjustment process. We propose a method that may better capture this price adjustment process.

The field of information theory offers an alternative approach for understanding information in price. This field models the quantification and communication of information and demonstrates how information can be passed through a noisy channel, without respect to the meaning of the information (MacKay 2003). These models form the foundation of modern communication technologies (e.g., modems, cell phones, data storage, etc.). An important aspect of this field is detecting the flow of information between two systems. We base our measure of information flow in securities markets on a recently developed measure of information transfer, transfer entropy (Schreiber 2000; Paluš, Komárek, Hrnčíř, and Štěrbová 2001).[1]

Transfer entropy (TE) captures the directed exchange of information between two systems, even when the information channel between the two systems is noisy (Schreiber 2000). In the context of market prices, the TE from security B to security A is the reduction in uncertainty about the future price of security A by knowing the past price(s) of security B, conditional on knowing the past price(s) of security A. This example may help with the intuition. No one can be certain what IBM's stock price will be tomorrow. Although knowing that IBM's stock price today is $120 reduces this uncertainty, some amount of uncertainty remains. If knowing HP's stock price today is $23 reduces this remaining uncertainty, transfer entropy is positive indicating an information transfer from HP to IBM. Although this example is over two-periods to aid intuition, transfer entropy is estimated across a longer time series.

Prior research develops a few approaches for estimating transfer entropy. In all approaches, researchers must specifying the time between periods (time scale), the length of the time series (time series size), the number of past prices that inform future price (lookback or embedding dimension), and the direction of information transfer.[2] The correct estimation approach depends on, at least in part, the underlying data. Estimation of TE across

[1] See Bossomaier, Barnett, Harré, and Lizier (2016) for an introduction to transfer entropy.
[2] Figure 1 depicts these concepts.

discrete data (e.g., binary) is simple and computation is trivial. This method amounts to calculating various conditional probabilities by counting. Unfortunately, using this approach on continuous data requires discretizing the data and the resulting estimates of TE are quite sensitive to choices made in discretizing. Addressing this problem, prior research proposes a number of TE estimation approaches when data are continuous (e.g., Moon, Rajagopalan, and Lall 1995; Kraskov, Stögbauer, and Grassberger 2004; Vicente, Wibral, Lindner, and Pipa 2011). One of the most widely used approaches for estimating TE is based on Kraskov, Stögbauer, and Grassberger (2004). This approach estimates TE using a k-nearest neighbor algorithm. While this method slightly underestimates TE (downward bias), it is consistent and relatively insensitive to choices of free parameters (i.e., the number of nearest neighbors). Computation of this approach is non-trivial.

We begin our analysis by examining the properties of the (Kraskov, Stögbauer, and Grassberger 2004) TE estimator. Using data simulated with varying levels of information transfer, we examine the conditions under which the TE estimator captures information transfer. In addition to varying the strength of the connection, we vary the time scale, the time series size, the lookback, and the direction. We find that as the strength of the information transfer increases the downward bias in TE grows, but that the estimator remains consistent. We find TEs estimated using one, two, or three nearest neighbors converge as the time series size increases. We also find that modeling a lookback that is higher the correct number lookback periods biases TE estimates downward. Finally, we hold constant the time series size and examine the effect of downsampling; modeling a longer timescale than the correct timescale. Our results indicate that downsampling biases TE estimates downward. Collectively, these results indicate the Kraskov, Stögbauer, and Grassberger (2004) TE estimator identifies information transfers with some downward bias, and that selecting the incorrect lookback, or time scale increases this bias, but that with a sufficiently long time series the number of nearest neighbors does not influence estimates. Given the direction of the bias, we see minimal concern of false positives, and given the estimator is

consistent (low variance), we proceed with using the Kraskov, Stögbauer, and Grassberger (2004) estimator of TE to measure information flow in securities markets.

Next, we analyze TE between equity and option markets. We select this setting because equity price and option prices should be influenced by the same set of information. Although textbooks explain that option prices are determined by equity prices (Hull 2012), indicating TE should be non-zero from equity to options and zero from options to equity, prior research documents trading in option markets can impact equity prices (Roll, Schwartz, and Subrahmanyam 2010; Johnson and So 2012; Ge, Lin, and Pearson 2016). One such case is pinning, where the price of the underlying is pulled toward the strike price of an expiring option (Avellaneda and Lipkin 2003; Golez and Jackwerth 2012). Since information transfers may occur in either direction, we examine TE from a firm's equity to all options on the equity and from these options to the equity. We first compute TE between SPY, an exchange traded fund (ETF) which tracks the S&P 500, and options on this ETF $\text{TE}_{SPY \to O}$.

The results of analyzing TE between SPY and its options indicate strong information flows between the two markets. Using a time scale of two seconds, we document a strong transfer entropy from SPY to in-the-money options $\text{TE}_{SPY \to O}$. The strength of $\text{TE}_{SPY \to O}$ grows the closer the in-the-money options are to being at-the-money. In fact, the strongest $\text{TE}_{SPY \to O}$ occur to options that are at-the-money. Surprisingly we find $\text{TE}_{SPY \to O}$ sharply drops to zero for out-of-the-money options. Examining information flows in the other direction ($\text{TE}_{O \to SPY}$) reveals a similar, but weaker, pattern which is consistent with conventional practice that the option price is determined by the underlying, while also supporting prior findings that the option market impacts the equity market.

As a confirmation that the TE pattern we document across the range of strike prices captures the pattern of information flow, we compute the correlation of the absolute value of returns between SPY and all options on SPY. Prior research uses absolute value of returns to measure information arrival. As prices in these two markets should be driven by the same underlying information, we posit prices react simultaneously to some information and

the strength of these reactions will be correlated with the delayed flow of other information across markets. Given this, we expect and find a similar pattern in the correlation of contemporaneous unsigned returns.

Modeling TE with longer lookbacks allows us to assess whether information flows from *multiple* past periods to the current period. If information efficiently transfers from the immediately prior period, we would expect to find minimal information transfers from earlier periods. If this is the case, we expect that modeling TE using a lookback greater than one will reduce estimated TEs. Consistent with this expectation, we find lower TEs for SPY when we select a lookback of two, indicating there is minimal information in prices two periods prior and supporting the notion information is efficiently transferred from one period to the next.

The time scale in the TE estimation captures the speed of the estimated information flows. By searching for the time scale where TE peaks, we identify the speed where information transfers generally occur. We compute $\text{TE}_{SPY \to O}$ across range of time scales (between 2 seconds and 60 seconds), and find a time scale of 12 seconds maximizes $\text{TE}_{SPY \to O}$. $\text{TE}_{SPY \to O}$ at this time scale is approximately 50% higher than $\text{TE}_{SPY \to O}$ with a time scale of two seconds. Examining the other direction $\text{TE}_{O \to SPY}$, we find the time scale that maximizes TE from option to SPY is 14 seconds, indicating a slightly slower information transfer from options to SPY. Comparing $\text{TE}_{SPY \to O}$ with a time scale of 12 seconds to $\text{TE}_{O \to SPY}$ with a time scale of 14 seconds reveals that, at the time scales where each TE peaks, the information flow from SPY to options is about twice as strong as the information flow from options to SPY.

In the remaining analysis, we examine transfer entropy from the underlying equity to options on the equity $\text{TE}_{U \to O}$ and from the options to the underlying $\text{TE}_{O \to U}$ for 78 tickers.[3] We first perform a search for the time scale where the TE between the underlying and their respective at-the-money options peaks. Plotting the peak $\text{TE}_{U \to O}$ and peak $\text{TE}_{O \to U}$ reveals

[3] As detailed in Section 4.1, data collection limitations only permit the analysis of 78 tickers.

these two TEs are noisily correlated. Interestingly, it appears there are two types of firms. One group of firms experiences similar levels of $TE_{U \to O}$ and $TE_{O \to U}$. The other group of firms experience much higher $TE_{O \to U}$ than $TE_{U \to O}$, indicating a much stronger information flow from options to equity. In this sample, we do not find firms that experience a much stronger information flow from equity to options. By plotting TEs between the underlying and all options at the identified peak time scale, we find that many tickers have a pattern of TEs similar to the patterns we document for SPY. There are some notable exceptions. For example, the drop in TE for Amazon's out of the money options is more gradual as options are more out-of-the-money. This analysis also allows us to compare information transfer time scales. While the majority of firms experience the strongest information transfers between 12 and 14 seconds, there are some with a large disparity between time scales. For example, Netflix has the strongest $TE_{O \to U}$ at a time scale of 8 seconds and the strongest $TE_{U \to O}$ at 12 seconds.

In our final analysis, we compare our measure of information transfer with other measures of information arrival. We examine the number of news articles published about a firm, the intensity of readership from Bloomberg terminals (i.e., institutional investor attention), and the number of EDGAR filings. These results indicate that TE is stronger for firms that have more news stories and higher readership, supporting the validity our TE measure captures information flow. We find no evidence that more EDGAR filings results in higher information flow. This may be either evidence that EDGAR filings contain stale information, or it may be due to our sample period covering a time period where these firms were not filing key filings (e.g., 10-Ks, 10-Qs, 8-Ks).

Our study is first to apply information theory to detect information flow at sub-minute levels. While other studies have examined TE in market settings, these tend to examine TE across market indexes, at very long time scales, and using TE estimators that are sensitive researcher choices (Marschinski and Kantz 2002; Kiyono, Struzik, and Yamamoto 2006; Harré and Bossomaier 2009; Reddy and Sebastin 2009; Walker, Cisneros, and Davies 2012;

Sensoy, Sobaci, Sensoy, and Alali 2014). In contrast, modeling TE at time scales closer to the transfer rates predicted by efficient market theories and using a consistent TE estimator, we examine securities which theoretically should be linked (equities and options on the equities). Our results contribute to the accounting, finance, and economics literature by demonstrating how TE can be used to capture directed information flows from one price to another, determine information transfer speed, and examine the efficiency of the price discovery process. Future research may find these aspects of TE useful in understanding the price discovery in any market, mapping market information networks, and examining how specific types of information (e.g., earnings announcements, management forecasts, analysts forecasts, material disclosures, etc.) is transferred. Additionally, our study provides insights into the literature that examines the connections between the option and equity markets. While prior studies demonstrate these markets influence each other (Golez and Jackwerth 2012; Johnson and So 2012; Ge, Lin, and Pearson 2016), we document the strength, timing, and variability of these connections across securities and firms.

We make three caveats. First, we acknowledge that the dyadic approach to measuring the price adjustment process may be overly simplified since a price may be simultaneously influenced by multiple prices (James, Barnett, and Crutchfield 2016). Nonetheless, we see examining information flow across dyads as a necessary step in developing a comprehensive understanding of the price adjustment process. Second, inherent in our approach is the assumption that the information flow between two securities is stable over the measurement window. Finally, computing TE across continuous processes takes a nontrivial amount of processing time, especially when computing many TEs. Computing TE between two 100,000 period time series takes a standard processor approximately twenty seconds. Estimation thus becomes non-trivial when examining many connections, performing searches for the time scale where TE peaks, or testing statistical significance of TE estimates.[4]

[4] For example, computing a single TE between every pair of 3,000 equities with a time series size of 100,000 would require approximately 9 million comparisons and take about 5.71 years (179.94 million seconds/60/60/24/365) of processing time. Of course, actual time can be reduced by estimating TEs in

2 Related Literature

Examining security prices is common place in accounting, finance, and economic research. Studies in the 1960s established the standard research design we still used today to capture how price changes to capture information. Ball and Brown (1968) and Fama, Fisher, Jensen, and Roll (1969) employ the event study design in which common information events for firms are aligned in time and price changes are calculated over a commonly sized window. These windows may be a few days to a few months. The assumption behind this design is that price changes result from market participants interpreting new information and taking actions that impound the information into price. Examining stock price changes with this design quickly became the standard way of using price to make inferences about information and continues to be extensively used by researchers.

Further research examines price connections across firms. Focusing on earnings announcements (Foster 1981; Pownall and Waymire 1989; Han and Wild 1990; Freeman and Tse 1992; Ramnath 2002) or on management earnings forecasts (Baginski 1987; Han, Wild, and Ramesh 1989; Pyo and Lustgarten 1990; Kim, Lacina, and Park 2008; Koo, Wu, and Yeung 2017), these studies examine information transfer by calculating how a disclosing firm's peer firms' prices change on the date of the announcement. Another area of research examines the correlation of returns across firms, either contemporaneously (Barberis, Shleifer, and Wurgler 2005) or with a lead-lag relationship (Lo and MacKinlay 1990; Hou 2007). In these areas, research designs focus on the price change from the beginning to end of the window and thus do not capture the information in the price changes occurring during the window.[5] In contrast, our method examines lead-lag prices at short time scales, detecting directed information flow from this overlooked price variation.

parallel across many processors. To put a time series size of 100,000 into perspective, four 6.5-hour trading days consist of 93,600 seconds.

[5] A noteworthy exception is the intraperiod timeliness (IPT) metric which captures how quickly a price moves during a window to the final price (Butler, Kraft, and Weiss 2007; Bushman, Smith, and Wittenberg-Moerman 2010; McMullin, Miller, and Twedt 2018). However, to our knowledge, IPT has yet to be used to examine information transfers.

Prior research examines the connections between the options and equity markets. While theory indicates prices of equity options should simply be a function of the underlying equity price (Hull 2012), recent papers document the options market can impact equity prices. For example, Roll, Schwartz, and Subrahmanyam (2010), Johnson and So (2012), and Ge, Lin, and Pearson (2016) calculate an options to stock volume ratio and document how this ratio can predict stock returns. Avellaneda and Lipkin (2003) and Golez and Jackwerth (2012) document pinning, instances where the underlying's price is pinned at the strike of an option with high open interest. These studies motivate estimating TEs in both directions across these markets. By examining information flows between the equity market and options markets, we are able to test whether TE detects information flows between securities with values that should be determined by the same set of information. Results of our analysis will thus serve both as construct validity tests of TE as well as provide insight into the literature examining the equity and options markets.

Our information flow examination approach differs from prior research. This research typically first identifies firms with a common information signal (e.g., stock split, earnings announcement, forecast, regulatory filing, etc.), and then examines whether the stock prices of these firms changes in a consistent manner. Our approach does not require identifying a common information signal. Rather, we detect the flow of information by exploiting the information signals embedded in prices. That is, we use prices themselves as the source of information. The advantage of this approach is that it allows us to detect information flow without regard to meaning and to measure information arrival at high frequencies. We note that our method could be used by researchers to first detect high amounts of information flow and then search typical information sources to provide insight into the potential meaning into the detected information flow. While this second step is not within the main scope of this study, our last set of analysis examining news articles, investor attention, and EDGAR filings are akin to this guided search. This "reverse" order is inline with the part of the information theory field concerned with detecting information flow (e.g., He and Tong (2008)).

2.1 Information Theory

Information theory is the scientific field concerned with quantifying, storing, and communicating information through noisy channels.[6] Shannon (1948) establishes the basic elements of communication from which the information theory field was born.[7] Among other insights, this paper established a formal definition of information. Given that the word "information" is already widely used in accounting, finance, and economics, it is important to note that in this field information does not capture actual meaning, but is defined as a statistical quantity. The information I in event x is defined with respect to the likelihood of the event occurring. More formally:

$$I(x) = log_2(\frac{1}{p(x)}) = -log_2 p(x) \tag{1}$$

This definition implies that rarer events contain more information. Likewise, events expected to occur with 100% probability contain no information, and do not reduce uncertainty. While x is the occurrence of a single event, we define X as a random variable over the set of possible events Ω.[8] The entropy H of X can be thought of as the amount of "disorder" or "uncertainty" in the process and is calculated as the weighted average information over the set of events, weighted by the probability of the events, or:

$$H(X) = -\sum_{x \in \Omega} p(x) \log_2 p(x). \tag{2}$$

Before we formally define transfer entropy, there are three more information theoretic quantities we need to define. The first is the conditional entropy (CE). CE is the entropy, or

[6] Unlike most fields, the initial founding of the information theory field is traced to a single paper (Shannon 1948). Stone (2015) provides an intuitive introduction to information theory.

[7] Shannon (1948) was initially published as an article titled "A Mathematical Theory of Communication." It was later republished in 1963 as a book titled, "The Mathematical Theory of Communication."

[8] For simplicity, we use discrete events to introduce these concepts.

uncertainty, remaining after conditioning on some process Y. It's helpful first to define the entropy after conditioning on a single event y:

$$H(X|y) = -\sum_{x \in \Omega_x} p(x|y) \log_2 p(x|y). \tag{3}$$

Conditional entropy is simply the average over all possible events in process Y,

$$H(X|Y) = \sum_{y \in \Omega_y} p(y) H(X|y). \tag{4}$$

CE captures the amount of information or uncertainty remaining in X after conditioning on Y. The second information theoretic quantity we need to define is mutual information (MI). MI measures the amount of information shared across two processes. MI can be thought of as a measure of statistical dependence or a non-linear form of correlation and is symmetric. More formally,

$$MI(X:Y) = H(X) - H(X|Y) = H(Y) - H(Y|X). \tag{5}$$

Interpreting this equation, recall that $H(X)$ is the amount of information (or uncertainty) in X. $H(X|Y)$ is the remaining amount of information in X after conditioning on Y. Thus, $H(X) - H(X|Y)$ captures the portion of the information in X that is also in Y, or the mutually shared information. When defined with respect to the probabilities of discrete events, MI is:

$$MI(X:Y) = \sum_{x \in \Omega_x} \sum_{y \in \Omega_y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}. \tag{6}$$

Combining the CE and MI concepts allows us to define the third quantity, conditional mutual information (CMI). CMI simply conditions the two terms of MI on a third random variable Z,

$$CMI(X:Y|Z) = H(X|Z) - H(X|(Y,Z)). \tag{7}$$

While first term $H(X|Z)$ captures the remaining uncertainty in X after conditioning on Z, the second term $H(X|(Y,Z))$ captures the remaining uncertainty in X after conditioning on both Z and Y. Thus the difference between these two terms captures the extent to which information in Y reduces uncertainty about X, all after conditioning on the information in Z.

With these information theory concepts in place, we are ready to defined transfer entropy (TE). TE captures the information flow from one sequential random variable Y to another sequential variable X by measuring the reduction in uncertainty about the future state of X from knowing current and/or past states of Y, conditional on current and/or past states of X[9]. TE can be seen as a form of CMI where the three processes are X_{t+1}, X_t, and Y_t,

$$\text{TE}_{Y\to X} = CMI(X_{t+1}:Y_t|X_t) = H(X_{t+1}|X_t) - H(X_{t+1}|(Y_t,X_t)). \tag{8}$$

The first term of this equation $H(X_{t+1}|X_t)$ represents the uncertainty remaining about the future of X, after conditioning on the current value of X. The second term $H(X_{t+1}|(Y_t,X_t))$ represents the remaining uncertainty about the future of X after conditioning on current values of both X and Y. This means the difference in the two terms measures the reduction in uncertainty or flow of information from the current value of Y to future values of X, while conditioning on any contemporaneously shared information in both X and Y.

TE was first formally defined by Schreiber (2000) and Paluš, Komárek, Hrnčíř, and Štěrbová (2001). Independently, these two papers discovered that information flow from Y to X could be determined after condition on the past information in X. A similar concept was developed much earlier by Granger (1969) in the context of causal inference and is

[9] One can think of this set up as a Markov chain where the past state(s) of X and Y inform the future state of X

commonly known as Granger Causality. Barnett, Barrett, and Seth (2009) show TE is equivalent to Granger Causality for Gaussian variables. Thus, using Granger Causality to estimate information transfers would be comparable to TE estimates if price data were Gaussian. Unfortunately, prices at small time scales are likely non-Gaussian, meaning to estimate information transfers at time scales similar to where theory predicts information flows occur requires using TE.[10]

3 Estimating and Computing Transfer Entropy

3.1 Estimating Transfer Entropy

There are multiple approaches for estimating transfer entropy between systems evolving in time. The simplest approach for estimating transfer entropy is binning time series data elements. This involves specifying a bin width and then partitioning the time series into the bins for estimating conditional probabilities. The challenge is that the bin width is a free parameter and transfer entropy estimates vary significantly based the choice of this parameter. While there are approaches to correctly identify the number of bins to use for a dataset in a histogram (Freedman and Diaconis 1981), the binning approach can still bias the probability distribution estimation even when including corrective terms (Darbellay and Vajda 1999) and adding adaptivity to the binning process (Fraser and Swinney 1986; Grassberger 1988). Consequently, prior research applies additional methods for estimating transfer entropy including both kernel density estimators (Moon, Rajagopalan, and Lall 1995) and nearest neighbor based estimators (Vicente, Wibral, Lindner, and Pipa 2011). Among nearest neighbor based estimators, computing transfer entropy via mutual information difference using the Kraskov estimator for mutual information (Kraskov, Stögbauer, and Grassberger

[10] Prior research uses Granger Causality to measure economic connectedness among financial companies (Billio, Getmansky, Lo, and Pelizzon 2012) and to identify return leader–follower pairs (Scherbina and Schlusche 2018) across monthly and weekly intervals, respectively. At these times scales, the linearity and Gaussian assumptions may be more valid than at sub-minute time scales.

2004) is among the most widespread, having been adopted in the widely-used TRENTOOL toolbox (Lindner, Vicente, Priesemann, and Wibral 2011) for Matlab and the Transfer Entropy package for R (Torbati and Lawyer 2018). This approach for estimating transfer entropy also performed favorably in a comparison of different transfer entropy estimation methods (Khan, Bandyopadhyay, Ganguly, Saigal, Erickson, Protopopescu, and Ostrouchov 2007).

The Kraskov estimator estimates mutual information; the mutual information difference can in turn be used to estimate transfer entropy as summarized here. Given the two time series of continuous random variables X and Y, the mutual information of these two series, measured in bits, is given by

$$\mathrm{MI}(X,Y) = \int\int \mu(x,y) \log_2 \frac{\mu(x,y)}{\mu_x(x)\mu_y(y)} dxdy \tag{9}$$

where $\mu(x,y)$ is the probability density of a random variable $Z = (X,Y)$ and μ_x and μ_y are the marginal probability densities of X and Y respectively. Mutual information, unlike transfer entropy, is symmetric (e.g., $\mathrm{MI}(X,Y) = \mathrm{MI}(Y,X)$). However, it can be shown by rearranging equation (8) that TE can be obtained from the difference of two mutual information computations as follows (Kaiser and Schreiber 2002):

$$\mathrm{TE}_{Y\to X} = T(X_{t+1}|X_t, Y_t) = \mathrm{MI}\left(X_{t+1}, [X_t, Y_t]\right) - \mathrm{MI}(X_{t+1}, X_i) \tag{10}$$

where the direction of information flow is from Y to X. Gencaga, Knuth, and Rossow (2015) demonstrate this estimation approach results in an overall lower bias when compared to alternative approaches. From this perspective, transfer entropy from Y to X is the difference between two mutual information measures: the mutual information of the X_{t+1} and the $Z_t = [X_t, Y_t]$ state and the mutual information of the X_{t+1} and X_t state.

The Kraskov estimator for mutual information is a follows (Kraskov, Stögbauer, and Grassberger 2004):

$$\text{MI} = \psi(k) - 1/k - < \psi(n_x) + \psi(n_y) > + \psi(N) \tag{11}$$

where k is a free parameter that specifies the size of the nearest neighborhood, ψ is the digamma function(Galassi 2018), n_x denotes the number of points within the x distance to its k-th neighbor, n_y denotes the number of points within the y distance to its k-th neighbor, $< \ldots >$ denotes the average of the quantity inside the angled brackets, and N is the length of the time series. The max-norm is used for computing distance in the Kraskov estimator

$$||\mathbf{x}||_\infty = \max\left(|x_0|, |x_1|, \ldots, |x_{n-1}|\right) \tag{12}$$

where $\mathbf{x}$ is a vector with elements $(x_0, x_1, \ldots, x_{n-1})$. When computing mutual information difference for transfer entropy estimation, several of the mutual information component terms cancel out, leaving just the difference of the averages.

Estimating transfer entropy from information difference requires choosing an embedding dimension for the time series as well as fixing the free parameter in the Kraskov estimator. The embedding dimension reflects the how many past values in a time series comprise a dependency for the current state and sets to what order Markov process the time series is. Thus, an embedding dimension of one would indicate that the state at time $t+1$ from a time series would only depend on the state at the previous time t. An embedding dimension larger than one would indicate a larger lookback in time for the dependency state to $t-1, t-2$, etc. An efficient market is consistent with an embedding dimension or lookback of one; the current value of the equity price would be the only relevant value that can affect the next price realization in an efficient market.

The free parameter in the Kraskov estimator, k, specifies the number of neighbors used for estimation and can impact the transfer entropy estimate. However, transfer entropy

estimates using the Kraskov estimator will converge for any k (Wang, Kulkarni, and Verdú 2009). In general, larger k values will reduce the estimate's variance while smaller k values will be more accurate (Khan, Bandyopadhyay, Ganguly, Saigal, Erickson, Protopopescu, and Ostrouchov 2007). We demonstrate this empirically in Section 3.2 where time series are generated with a known transfer entropy and then transfer entropy is estimated using the mutual information difference via the Kraskov estimator. Our analysis uses $k = 1$ for reporting results except in certain example cases.

There is no closed form solution for determining whether a TE is statistically significant. Thus, unlike other estimators where assumptions about asymptotics provide ways to estimate the standard error of the estimator, the suggested approach for determining significance is to randomly shuffle the Y time series a thousand times, recomputing TE after each shuffle (Bossomaier, Barnett, Harré, and Lizier 2016). In untabulated tests, when examining the equity and options data below, we find that this reshuffling results in TE estimates with a mean of zero and standard deviation of between 0.002 and 0.005. Thus, within our dataset, finding TEs with values greater than 0.01 would indicate the TE is at least two standard deviations above what would be expected if the Y vector contained pure noise.

3.2 Computing Transfer Entropy using Simulated Data

Efficiently computing transfer entropy via mutual information difference and the Kraskov estimator requires a fast method to find nearest neighbors to a given point and counting the numbers of points within a certain radius. We adopt the code framework of the Transfer Entropy package for R (Torbati and Lawyer 2018) which uses a kd-tree approach based on the arbitrary nearest neighbor (ANN) library (Mount and Arya 2018) for finding nearest neighbors. In this section, we examine the properties of the Kraskov estimated transfer entropy using two coupled time series as explored by (Gencaga, Knuth, and Rossow 2015) which are first order Markov chains:

$$y_{i+1} = 0.5y_i + n_i^1 \quad (13)$$
$$x_{i+1} = 0.6x_i + cy_i + n_i^2 \quad (14)$$

where $c \in [0.01, 1]$, and n_i^1, n_i^2 are random numbers from the standard normal distribution. The benefit of this approach is that the transfer entropy between the two time series is known. For these time series, the transfer entropy from X to Y, TE_{xy}, is analytically zero while the transfer entropy from Y to X, TE_{yx} is between 0 and 0.55 depending on the choice of the coupling parameter c. The analytic transfer entropy for Eqns. 13–14 is shown in Figure 2a as a function of different coupling parameter choices. The transfer entropy estimated from mutual information difference using the Kraskov estimator is also shown for two different choices of k parameters. As noted in Section 3.1, transfer entropy estimates converge regardless of the choice of k; however, larger values of k show less variance in the estimate while smaller values of k are more accurate.

To illustrate the convergence of transfer entropy estimates for differing k values in the Kraskov estimator, Figure 2b shows the estimated transfer entropy where $c = 1.0$ in Eqns. 13–14 for differing values of of k and different time series sizes. As the time series lengths increase, the differences in transfer entropy estimates between the different values of k disappear.

While the time series in Eqns. 13–14 are first order Markov chains and show no dependence on time series values prior to the present value, transfer entropy can still be estimated using an embedding dimension or lookback larger than one. However, doing so decreases the accuracy of the transfer entopy estimation for cases with nonzero transfer entropy such as TE_{yx} for Eqns. 13–14. This is illustrated in Figure 2c where transfer entropy estimation based on different lookback choices is plotted against the analytic transfer entropy.

When estimating transfer entropy from time series data sampled from fast processes such as those originating in financial data, there is a risk of underestimating transfer entropy as illustrated here. If the time series in Eqns. 13–14 are downsampled to mimic insufficiently

fast sampling of coupled autoregressive processes, the estimated transfer entropy will be substantially lower than the actual value for those processes as shown in Figure 2d.

4 Transfer Entropy Across Equity and Option Markets

Equity and options markets present a useful test case for determining whether transfer entropy estimates can detect information flows in markets. We expect information flows to occur in both directions across these markets. Since option value is derived from the underlying's value, prices of both should be influenced by the same set of information. Given this, we expect these securities to be strongly correlated. However, given that one security is a derivative of the other, its natural to expect that price discovery in the option market originates from the equity market, meaning information in the underlying's price should transfer to the option prices.

Prior research documents trading activity in the options market can impact equity prices. One such stream of literature examines the ratio of option market volume to equity market volume and documents how this ratio predicts equity prices (Roll, Schwartz, and Subrahmanyam 2010; Johnson and So 2012; Ge, Lin, and Pearson 2016). Another stream documents equity price pinning. In these pinning cases, the equity price finishes within a few cents of an option strike price and can be explained as the consequence of delta hedging in the presence of a large option open interest at that strike price (Avellaneda and Lipkin 2003). Pinning phenomena is frequently observed in options trading, appearing much more frequently than would be expected by random equity fluctuations with conclusive evidence given in (Ni, Pearson, and Poteshman 2005; Golez and Jackwerth 2012).

Our expectation is also supported anecdotally by two observations from options trading practice. First, traders can react to new information by trading options just as easily as trading the underlying itself. Reacting in this manner could make the option price the lead information source in price discovery for the underlying. Second, a large options order at a single strike frequently influences the underlying price itself, again illustrating how the option

price leads price discovery for the underlying. In summary, we posit that while options and their underlying are tightly correlated, there is frequently information moving not just from the underlying price to its options' prices, but also from the options' prices to the underlying price.

Transfer entropy is a unique tool for testing this conjecture. Unlike correlation and mutual information, transfer entropy gives the direction in which the information is moving and reveals the lead in price discovery. Further, it provides a quantitative comparison of the amount of information flowing in each direction. Finally, by empirically measuring the Markov chain order of the financial data via the embedding dimension which gives the peak transfer entropy, we can provide insight into market efficiency.

4.1 Data

We obtain data from Bloomberg to examine the performance of estimates of TE in security markets. Specifically, we collect bid and ask prices of 78 equity securities and all call options on these securities every two seconds during trading hours from September 7, 2018 until September 21, 2018, the day these options expire.[11] Table 1 lists the ticker, company name, number of options and the price range over our sample period for these 78 equities. In total, our sample contains 3,475 securities. The price time series for each security is approimately 204,000 periods long.[12] We selected these 78 so our sample contains firms from number of industries, with variation in the type information environment, while still having fairly active trading in both markets.

[11] We also collect last price, but do not tabulate analysis of this price as it updates only when a trade is executed. Thus, bid and ask prices are more likely to capture market participants responses to information.

[12] Using the Microsoft Excel Bloomberg add-in, we retrieve and save the realtime `LAST_PRICE`, `BID`, and `ASK` fields approximately every two seconds using a VBA script. Our sample is limited to 78 equity securities and their corresponding call options since Bloomberg limits the number of realtime data fields that can be displayed to 3,500 securities. Our Bloomberg terminal receives data with a 15 minute delay which is present and the same for all securites. We observe some minor variation in the number of periods in the time series due to initial delay in when the first data are delivered when the VBA script is executed each day.

We also collect data to capture information arrival and investor attention over this time period. From Bloomberg historical datasets, we collect the number of news articles published (Bloomberg field: NEWS_HEAT_PUB_DNUMSTORIES) and the readership interest (Bloomberg field: NEWS_HEAT_READ_DMAX) for each day in our sample period (Ben-Rephael, Da, and Israelsen 2017).[13] We also collect the number of forms each company files on the U.S. Securities and Exchange Commission's EDGAR system, from July 1, 2018 to September 21, 2018. We compute an abnormal measure of EDGAR filings by dividing the number of filings filed during our time period by the average number of filings filed every ten trading days between July 1, 2018 and September 6, 2018. While the number of news articles and EDGAR filings capture arrival of information, readership interest captures abnormal institutional investor attention. These three variables allow us to assess whether TE is higher for firms with potentially more information arriving and investor attention.

4.2 Analysis

We begin our analysis of the equity and options data by computing TE between SPY, an exchange traded fund that tracks the S&P 500, and call options on SPY. We focus our initial analysis on SPY for three reasons. First, there is likely plenty of new information impacting the price of SPY as its value is a function of many firms. Second, there are many options written with SPY as the underlying. Third, markets for SPY and options on SPY are liquid. These three aspects increase the likelihood that information is being transferred between these two markets and that we will be able to detect it.

We calculate both directions of TE between SPY and options on SPY at a time scale of two seconds, the fastest time scale our data permit, and k-nearest neighbor parameter

[13] NEWS_HEAT_READ_DMAX is the maximum value of the Bloomberg field NEWS_HEAT_USER_ACTIVITY_RT which is a measure of readership interest in a company relative to the prior 30 days. It is based on the number of times Bloomberg terminal users call up stories with the ticker attached and the number of times they search for current news specifically about a given ticker. A score of 0 indicates readership is not widespread or is below the 30-day average. Scores 1-4 indicate readership is unusually high, 4 being the top of the range.

equal to one, two, and three. Plotting the estimated TEs and the options' strike prices in Figure 3a reveals an interesting pattern. We find a positive TE between SPY and in-the-money options; approximately between 0.06 and 0.11 for $\text{TE}_{SPY \to O}$ and 0.02 and 0.075 for $\text{TE}_{O \to SPY}$. Comparing $\text{TE}_{SPY \to O}$ to $\text{TE}_{O \to SPY}$, we find stronger information transfer from SPY to the options. As the in-the-money options become closer to being at-the-money, we find TE peaks, such that strongest information transfers occur with options that are at-the-money. For TE in both directions, we find TEs drop sharply to zero as options become more out-of-the-money.

As mentioned above, the value of an equity and its options should be based on a similar set of information. Much of this information set is likely impounded into price simultaneously. Using the absolute value of returns as a measure of information entering price, we examine how similar contemporaneous price discovery is to directed information flows by calculating the correlation of absolute value of returns across the markets. We posit that the strength of this simultaneous reaction to some information is correlated with the delayed flow of other information. Thus, we expect to find a pattern similar the pattern in Figure 3a for the correlation of contemporaneous unsigned returns. Figure 3b depicts these correlations between SPY and each option. These reveal a similar peak for at-the-money options, increase in strength of connection as in-the-money options are closer to being at-the-money, and drop in connection for out-of-the-money options. In summary, we interpret the similarity across Figure 3a and Figure 3b as evidence TE captures directed information flow.

In the TE estimation, lookback captures the number of prior periods' prices modeled to reduce uncertainty about future price.[14] Finding TEs with a lookback of two or greater (i.e., price at t and $t-1$) are lower than TEs estimated with a lookback of one (i.e., price at t only) indicates information transfer occurs from price from t to $t+1$ and that market incorporate prices efficiently. We estimate TE between SPY and options on SPY with a lookback of two

[14] This lookback is a parameter in modeling the time series as a Markov chains. A lookback of one is a first order Markov chain, two is a second order Markov chain, etc.

and report these estimates in Figure 3c. While we find the pattern of TE with lookback of two (Figure 3c) is similar to that estimated when using a lookback of one (Figure 3c) the TE with a lookback of two is much smaller, suggesting information is efficiently incorporated into SPY and options on SPY.[15]

The time scale, or sampling frequency, is the time in between price measurements. The time scale with the highest TE indicates when information flow is the strongest. While we collected the data a two second sampling frequency, two seconds may not be the peak timescale for information transfer across SPY and options on SPY. By estimating TE at various time scales, we find peak $\text{TE}_{SPY \to O}$ at a time scale of twelve seconds. Figure 3d shows the estimated TE using time series data at two second and twelve seconds for both TE directions. At the 12 second timescale, the maximum at-the-money TE from the underlying to the option is nearly 50% higher than when measuring TE at the two second sampling timescale. For SPY, the peak timescales are 12 and 14 seconds from underlying to option and from option to underlying, respectively. TEs for SPY at these timescales are shown in Figure 3e. This reveals that even after identifying the peak TE for both directions, $\text{TE}_{SPY \to O}$ is about twice as high as $\text{TE}_{O \to SPY}$.

Next we compute TE between the other 77 equities and their respective options. We perform a search for the lookback and timescale at which TE between the equity and at-the-money options peaks. Specifically, for each security examined, we identify the maximum TE for those options which were at-the-money at some point during our sample period. We compute TEs with lookback parameters of one through three and timescales from two seconds to twenty seconds in two second increments. In Table 2, we report peak TEs from the underlying to the option in the column titled TE_{uo} and TEs from the option to the underlying in the column titled TE_{ou}. Associated with each of these columns is the timescale which gives the peak transfer entropy in each case. In this search, we found the peak transfer

[15] This is also evidence that the price time series are first order Markov chains.

entropy always occurs with a lookback of one, supporting the efficient market hypothesis for these data.[16]

We provide the full plot of each ticker's TEs as a function of strike price in the online appendix. Many of the estimated TE patterns show a pronounced jump in the vicinity of the at-the-money options. Out-of-the-money options tend to have a very low transfer entropy in either direction. To summarize these many figures, we plot in Figure 4a both directions of TE between options and their underlying. For the 78 tickers examined along with their respective options, all show some transfer entropy from the option to the underlying. A good example of this is the transfer entropy for AMZN, shown in Figure 4b. For AMZN, while the TE is predominantly in the direction from the underlying to the option, there is a strong signal from the option to the underlying. Even though this is contrary to the conventional theoretical treatment of options in that the option price should not affect the underlying price, it is consistent with empirical evidence from option trading as discussed in Section 4.

As illustrated in Figure 4a, for most of the 78 tickers examined the TE from the underlying to the options is larger than from the options to the underlying. However, there were some notable exceptions: Abiomed (ABMD), a medical device company, and Illumina (ILMN), a biotechnology company, were both outliers where the TE from the option to the underlying was significantly larger than from the underlying to the option. The timescale for the information transfer from option to underlying was also faster than the timescale for information transfer from underlying to option in both cases. Some of the other outliers where more TE moves from the option to the underlying than from the underlying to the option include ISRG, MSCI, TDG, GWW, AZO,ULTA, CMG, CHTR, GOOG, and GOOGL.

A survey of the timescales at which the peak TE was observed is provided in Figure 4c. The most common timescale observed clusters around 14 seconds for nearly all tickers with some exceptions. Among those exceptions, there are more tickers which show a measurably faster timescale of information transfer in the direction from the option to the underlying.

[16] All results set the free parameter, k, in the Kraskov estimator to one.

Figure 4d shows an example from NFLX, where the peak TE time scale from the option to the underlying occurs at 8 seconds whereas peak TE occurs at a time scale of 12 seconds from the underlying to the option.

In our final set of analyses, we compare estimated TEs with measures of information arrival and investor attention. If TE indeed captures information flows, we expect that TE should be higher when more information arrives about a firm and when investors pay more attention to the firm. In Figure 5a, we plot the quantity of news articles written about particular firm during the data collection period with the estimated TE. We find TE and news articles exhibit a positive correlation. In fact, the relation appears exponential (note, y-axis is base-10 log scale). Figure 5a contains some notable outliers. For example, there are a number of firms which experience a high $TE_{O \to U}$, but have a low number of news stories. This may be due to information arriving for these firms via sources other than traditional media. There is only one firm, GE, that experiences low $TE_{O \to U}$ but has a high number of news stories.

This trend continues when plotting the TE against measures of readership interest. The readership interest variable measures abnormal investor attention by measuring news searching and reading activity for specific stocks on Bloomberg terminals (Ben-Rephael, Da, and Israelsen 2017) and captures an important aspect of the price formation process. Upon learning of information, investors may search for news articles about a firm to determine if the news is stale or to put the new information in context. Investors likely search for additional information prior to making investment decisions. Greater amounts of this activity likely occur when more information arrives. In Figure 5b, we plot the average `NEWS_HEAT_READ_DMAX` with the estimated TE. We find a similar pattern indicating a positive correlation between readership and TE. Like in the previous figure, there are notable outliers. For example, ISRG, CMG, ABMD, GWW, and ILMN all have high peak levels of $TE_{O \to U}$, but have low readership.

Regulations require firms to file disclosures with the SEC on EDGAR. These disclosures often provide new information to the market. Using the number of filings filed with the SEC, adjusted for filing activity over the prior two months, we create Figure 5c. In this figure we do not find that more EDGAR filings result in higher information flow. This may be explained by prior research's findings that even important periodic filings cause a minimal market reaction (Li and Ramesh 2009). Also, our EDGAR measure captures all filings over our sample period. While some filings may have information content, many likely do not. In that case, an alternative explanation may be that our measure of firms' EDGAR activity may be too noisy to capture information arrival.

5 Conclusion

Prices contain information and change as new information arrives. Theories of market efficiency posit that the process whereby information enters occurs rapidly. However, prior research examining this price adjustment process focuses on windows between a couple of days to many months. Moreover, this research design focuses on the price change from the beginning to the end of the window, ignoring the price movements in between these endpoints. By exploiting advances from the field of information theory to detect information flow, we propose using TE to measure the directed flow of information between prices at sub-minute time scales. With TE we are able to extracts the information signal from the typically ignored noisy price movements between end points of these windows.

To examine whether transfer entropy captures information flows in securities markets, we examine the connections between equity price and prices of options on the equity, two fundamentally connected markets. We document stronger information flow from the equity market to the option market, relative to the information flow from the option market to the equity market; however, the option to equity information flow is non-zero. This is consistent with option prices being determined by the price of the underlying security, but with some price discovery occurring in the options market and then transferring to the equity market.

We also document transfer entropy is higher when more news stories are published about a firm or when abnormal institutional investor interest is higher.

In conclusion, our study presents the estimation requirements for and evidence supporting the use of TE to measure sub-minute information transfers. Future research may use this measure to understand the typical speed and efficiency of information flows in any market. With an improved method for uncovering information flows, future research into the price adjustment process could use this measure to detect and analyze market information networks both within and across markets. Additionally, this metric of information flow could be used by future research to better ascertain the information content of individual disclosures, or to help identify the source of the information being transferred.

References

Avellaneda, M. and M. Lipkin (2003, 12). A market-induced mechanism for stock pinning. *Quantitative Finance 3*(6), 417–425.

Baginski, S. P. (1987, October). Management Forecasts of Earnings. *Journal of Accounting Research 25*(2), 196–216.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research 6*(2), 159–178.

Barber, B. M. and J. D. Lyon (1997, December). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics 43*, 341–372.

Barberis, N., A. Shleifer, and J. Wurgler (2005, February). Comovement. *Journal of Financial Economics 75*(2), 283–317.

Barnett, L., A. B. Barrett, and A. K. Seth (2009, December). Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. *Physical Review Letters 103*(23), 125–4.

Ben-Rephael, A., Z. Da, and R. D. Israelsen (2017, May). It Depends on Where You Search: Institutional Investor Attention and Underreaction to News. *The Review of Financial Studies 30*(9), 3009–3047.

Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon (2012, June). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics 104*(3), 535–559.

Bossomaier, T., L. Barnett, M. Harré, and J. T. Lizier (2016). *An Introduction to Transfer Entropy*. Cham, Switzerland: Springer International Publishing.

Bushman, R. M., A. J. Smith, and R. Wittenberg-Moerman (2010, August). Price Discovery and Dissemination of Private Information by Loan Syndicate Participants. *Journal of Accounting Research 48*(5), 921–972.

Butler, M., A. Kraft, and I. Weiss (2007). The effect of reporting frequency on the timeliness of earnings: The cases of voluntary and mandatory interim reports. *Journal of Accounting and Economics 43*(2-3), 181–217.

Darbellay, G. A. and I. Vajda (1999, May). Estimation of the information by an adaptive partitioning of the observation space. *IEEE Transactions on Information Theory 45*(4), 1315–1321.

Fama, E. F., L. Fisher, M. Jensen, and R. Roll (1969). The adjustment of stock prices to new information. *International Economic Review 10*(1), 1–21.

Foster, G. (1981). Intra-industry information transfers associated with earnings releases* 1. *Journal of Accounting and Economics 3*(3), 201–232.

Fraser, A. M. and H. L. Swinney (1986, Feb). Independent coordinates for strange attractors from mutual information. *Phys. Rev. A 33*, 1134–1140.

Freedman, D. and P. Diaconis (1981, Dec). On the histogram as a density estimator:l2 theory. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 57*(4), 453–476.

Freeman, R. and S. Tse (1992). An earnings prediction approach to examining intercompany information transfers* 1. *Journal of Accounting and Economics 15*(4), 509–523.

Galassi, M. e. a. (2018). Gnu scientific library reference manual.

Ge, L., T.-C. Lin, and N. D. Pearson (2016, May). Why does the option to stock volume ratio predict stock returns? *Journal of Financial Economics 120*(3), 1–22.

Gencaga, D., K. H. Knuth, and W. B. Rossow (2015). A recipe for the estimation of information flow in a dynamical system. *Entropy 17*(1), 438–470.

Golez, B. and J. C. Jackwerth (2012, December). Pinning in the S&P 500 futures. *Journal of Financial Economics 106*(3), 566–585.

Granger, C. W. J. (1969, August). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica 37*(3), 424–438.

Grassberger, P. (1988). Finite sample corrections to entropy and dimension estimates. *Physics Letters A 128*(6), 369 – 373.

Han, J. C. and J. J. Wild (1990, April). Unexpected Earnings and Intraindustry Information Transfers: Further Evidence. *Journal of Accounting Research 28*(1), 211–219.

Han, J. C., J. J. Wild, and K. Ramesh (1989, January). Managers' Earnings Forecasts and Intra-Industry Information Transfers. *Journal of Accounting and Economics 11*(1), 3–33.

Harré, M. and T. Bossomaier (2009, July). Phase-transition–like behaviour of information measures in financial markets. *EPL (Europhysics Letters) 87*(1), 18009–6.

He, T. and L. Tong (2008, October). Detection of Information Flows. *IEEE Transactions on Information Theory 54*(11), 4925–4945.

Hou, K. (2007, June). Industry Information Diffusion and the Lead-lag Effect in Stock Returns. *The Review of Financial Studies 20*(4), 1113–1138.

Hull, J. C. (2012, January). *Options, Futures, and Other Derivatives*. Boston, MA: Prentice Hall.

James, R. G., N. Barnett, and J. P. Crutchfield (2016, June). Information Flows? A Critique of Transfer Entropies. *Santa Fe Institute Working Paper*, 1–6.

Johnson, T. L. and E. C. So (2012, November). The option to stock volume ratio and future returns. *Journal of Financial Economics 106*(2), 262–286.

Kaiser, A. and T. Schreiber (2002). Information transfer in continuous processes. *Physica D: Nonlinear Phenomena 166*(1), 43 – 62.

Khan, S., S. Bandyopadhyay, A. R. Ganguly, S. Saigal, D. J. Erickson, V. Protopopescu, and G. Ostrouchov (2007, Aug). Relative performance of mutual information estimation methods for quantifying the dependence among short and noisy data. *Phys. Rev. E 76*, 026209.

Kim, Y., M. Lacina, and M. S. Park (2008, September). Positive and Negative Information Transfers from Management Forecasts. *Journal of Accounting Research 46*(4), 885–908.

Kiyono, K., Z. R. Struzik, and Y. Yamamoto (2006, February). Criticality and Phase Transition in Stock-Price Fluctuations. *Physical Review Letters 96*(6), 563–4.

Koo, D. S., J. J. Wu, and P. E. Yeung (2017, August). Earnings Attribution and Information Transfers. *Contemporary Accounting Research 34*(3), 1547–1579.

Kothari, S. P. and J. Warner (1997). Measuring long-horizon security price performance. *Journal of Financial Economics 43*(3), 301–339.

Kraskov, A., H. Stögbauer, and P. Grassberger (2004, Jun). Estimating mutual information. *Phys. Rev. E 69*, 066138.

Li, E. X. and K. Ramesh (2009, July). Market Reaction Surrounding the Filing of Periodic SEC Reports. *The Accounting Review*, 1–38.

Lindner, M., R. Vicente, V. Priesemann, and M. Wibral (2011, Nov). Trentool: A matlab open source toolbox to analyse information flow in time series data with transfer entropy. *BMC Neuroscience 12*(1), 119.

Lo, A. W. and A. C. MacKinlay (1990, January). When are Contrarian Profits Due to Stock Market Overreaction? *The Review of Financial Studies 3*(2), 175–205.

MacKay, D. J. C. (2003, October). ***Information Theory, Inference, and Learning Algorithms***. Cambridge University Press.

Marschinski, R. and H. Kantz (2002, November). Analysing the information flow between financial time series. *The European Physical Journal B 30*(2), 275–281.

McMullin, J. L., B. P. Miller, and B. J. Twedt (2018, July). Increased mandated disclosure frequency and price formation: evidence from the 8-K expansion regulation. *Review of Accounting Studies*, 1–33.

Moon, Y.-I., B. Rajagopalan, and U. Lall (1995, Sep). Estimation of mutual information using kernel density estimators. *Phys. Rev. E 52*, 2318–2321.

Mount, D. and S. Arya (2018, October). ANN: A Library for Approximate Nearest Neighbor Searching.

Ni, S., N. Pearson, and A. Poteshman (2005, 10). Stock price clustering on option expiration dates. *Journal of Financial Economics 78*(1), 49–87.

Paluš, M., V. Komárek, Z. Hrnčíř, and K. Štěrbová (2001, March). Synchronization as adjustment of information rates: Detection from bivariate time series. *Physical Review E 63*(4), 520–6.

Pownall, G. and G. B. Waymire (1989, January). Voluntary Disclosure Choice and Earnings Information Transfer. *Journal of Accounting Research 27*(Current Studies on the Information Content of Accounting Earnings), 85–105.

Pyo, Y. and S. Lustgarten (1990, December). Differential intra-industry information transfer associated with management earnings forecasts. *Journal of Accounting and Economics 13*(4), 365–379.

Ramnath, S. (2002, December). Investor and Analyst Reactions to Earnings Announcements of Related Firms: An Empirical Analysis. *Journal of Accounting Research 40*(5), 1351–1376.

Reddy, Y. V. and A. Sebastin (2009, February). Interaction Between Forex and Stock Markets in India: An Entropy Approach. *Vikalpa 33*(4), 27–45.

Roll, R., E. Schwartz, and A. Subrahmanyam (2010, February). O/S The relative trading activity in options and stock. *Journal of Financial Economics 96*(1), 1–17.

Scherbina, A. and B. Schlusche (2018, November). Follow the Leader: Using the Stock Market to Uncover Information Flows Between Firms. *Review of Finance*, 1–71.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters 85*(2), 461–464.

Sensoy, A., C. Sobaci, S. Sensoy, and F. Alali (2014, November). Effective transfer entropy approach to information flow between exchange rates and stock markets. *Chaos, Solitons and Fractals: the interdisciplinary journal of Nonlinear Science, and Nonequilibrium and Complex Phenomena 68*(C), 180–185.

Shannon, C. E. (1948, October). A Mathematical Theory of Communication. *Bell System Technical Journal 27*, 379–423– 623–656.

Stone, J. V. (2015, February). *Information Theory: A Tutorial Introduction*. Sebtel Press.

Torbati, G. and G. Lawyer (2018, October). The Transfer Entropy Package.

Vicente, R., M. Wibral, M. Lindner, and G. Pipa (2011, Feb). Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience 30*(1), 45–67.

Walker, S. I., L. Cisneros, and P. C. W. Davies (2012, July). Evolutionary Transitions and Top-Down Causation. *Proceedings of Artificial Life 13*, 283–290.

Wang, Q., S. R. Kulkarni, and S. Verdú (2009, May). Divergence estimation for multidimensional densities via k-nearest-neighbor distances. *IEEE Trans. Inf. Theor. 55*(5), 2392–2405.

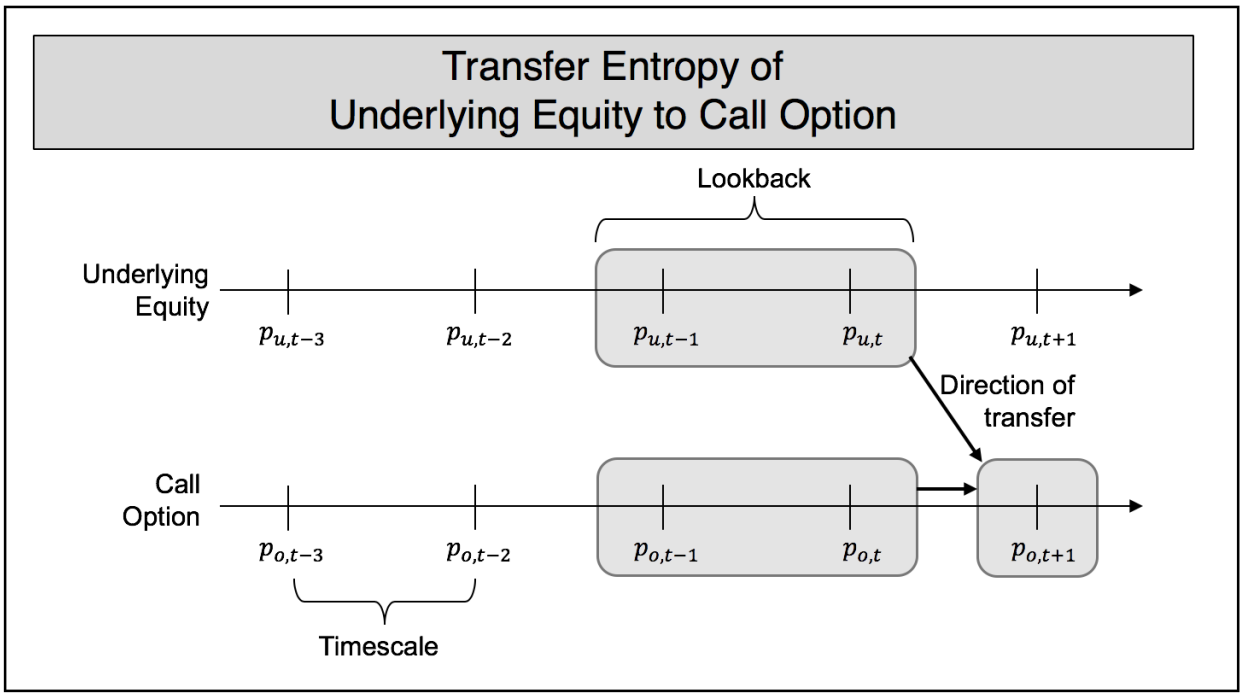

Figure 1 This figure depicts the key parameters in modeling transfer entropy. These parameters include: direction of information transfer, time scale, lookback and time series size. In this figure, the direction of information transfer is from the underlying equity to the call option The time scale is the amount of time between measuring p_{t-3} and p_{t-2} and is same between all price measurements. Price measurements for both equity and option are taken simultaneously. Lookback is the number of past prices that inform future price and is also referred to as the embedding dimension. In this figure, the lookback is two. Time series size is the number of price measurements.

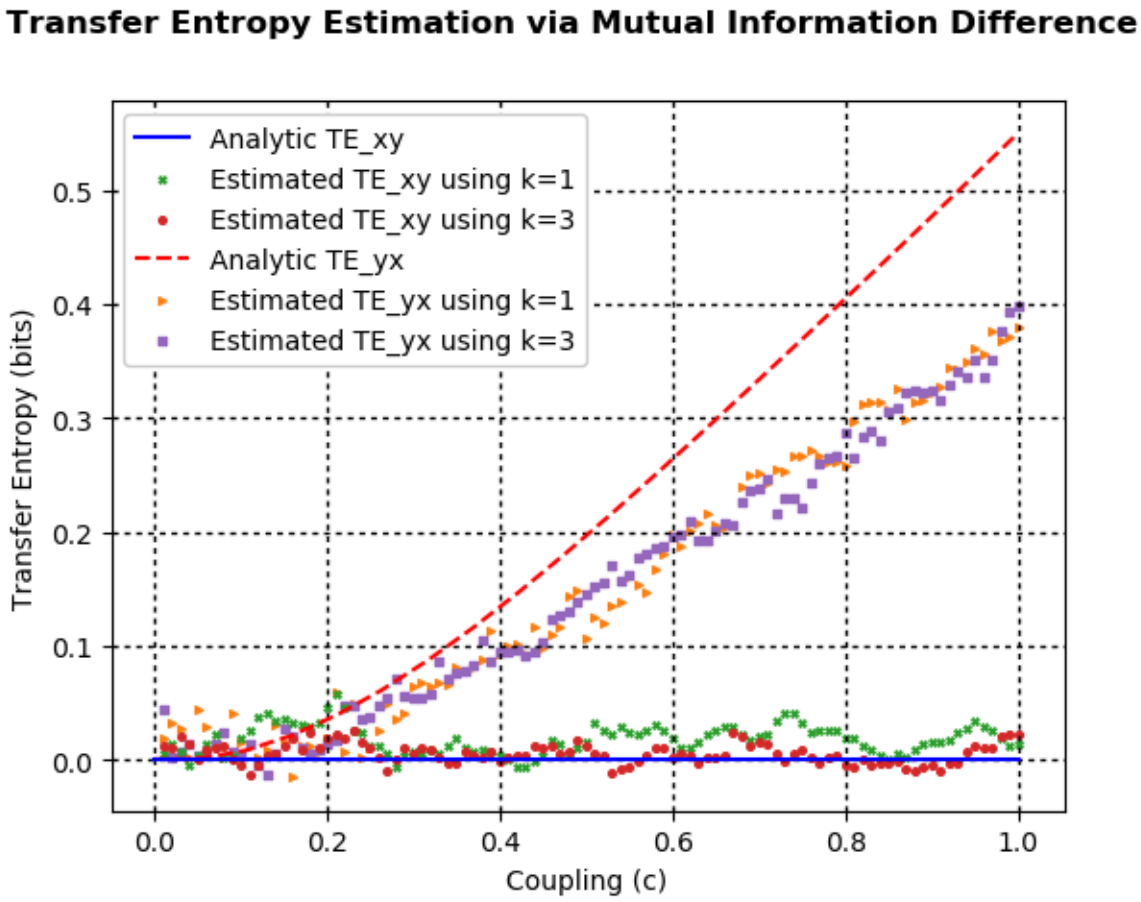

Figure 2a The analytic and estimated transfer entropy for the coupled processes in Eqns. 13–14 are shown here. Transfer entropy for Y to X, $\mathrm{TE}_{\mathrm{yx}}$ grows as the coupling constant c increases while transfer entropy for X to Y, $\mathrm{TE}_{\mathrm{xy}}$, is analytically zero. Transfer entropy is estimated from mutual information difference based on the Kraskov estimator. Two different choices of values for the free parameter in the Kraskov estimator, k, are illustrated to explore the impact this parameter can have on the estimate. In this example, the two time series each consisted of $20,000$ values.

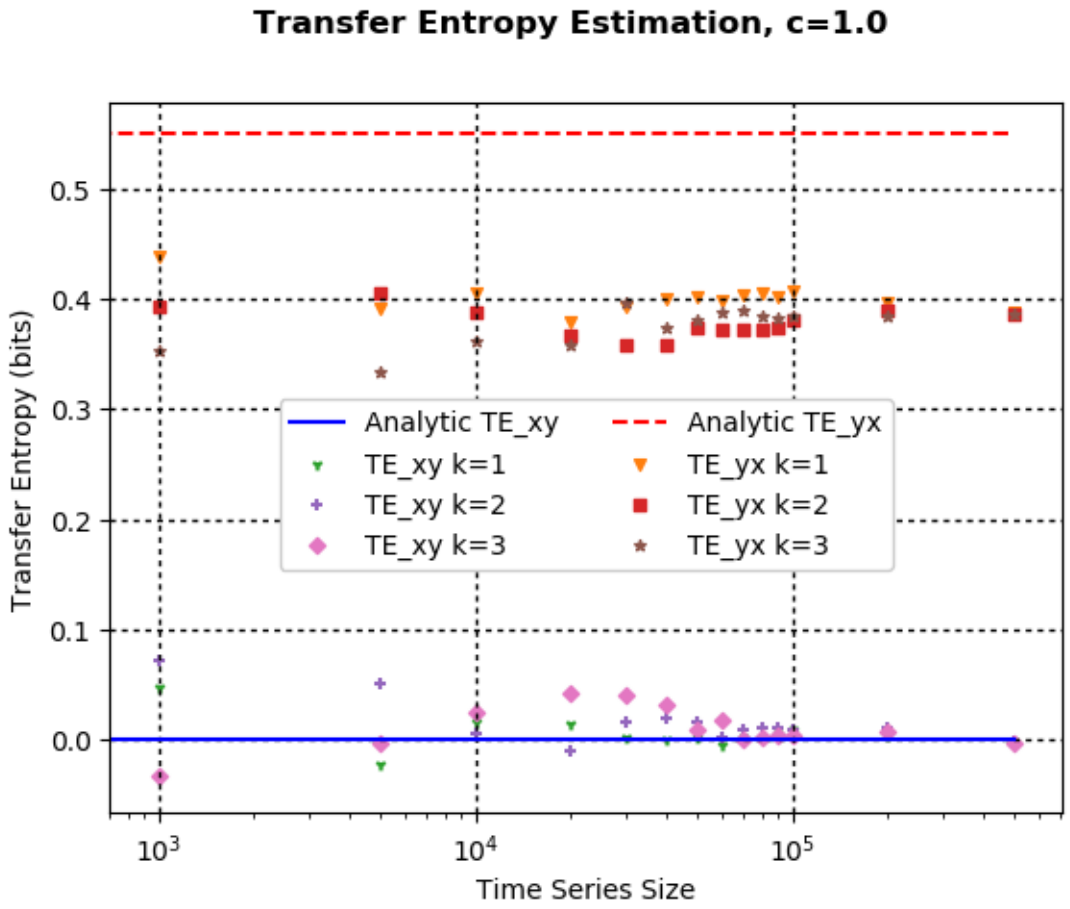

Figure 2b The analytic and estimated transfer entropy for the coupled processes in Eqns. 13–14 where $c = 1.0$ are shown here. Transfer entropy is estimated from mutual information difference based on the Kraskov estimator and different k values as well as time series lengths are used. As the time series length increases moving towards the right of the plot the spread between transfer entropy estimates using different choices of k disappear.

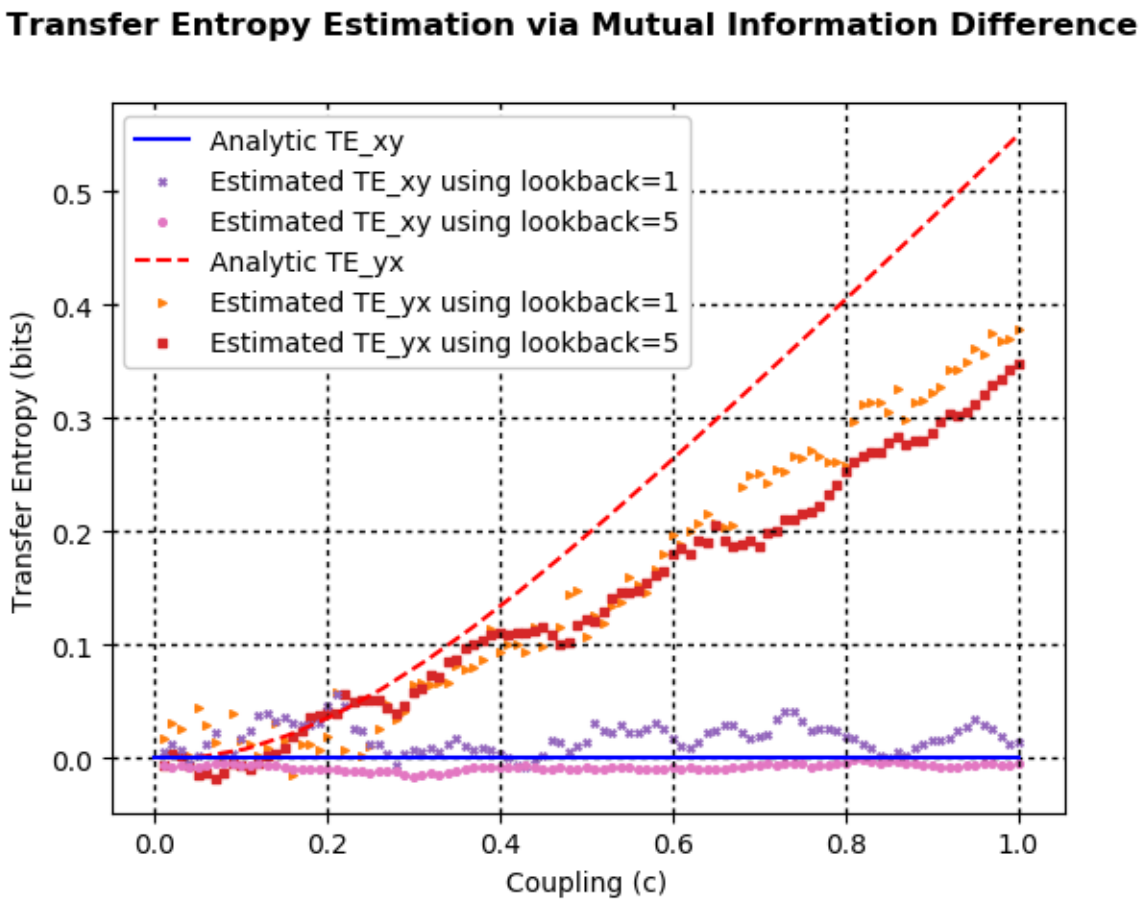

Figure 2c The analytic and estimated transfer entropy for the coupled processes in Eqns. 13–14 are shown here. The estimated transfer entropy used $k = 1$ in all cases but varied the lookback parameter. Using a larger lookback parameter decreased the accuracy of the transfer entropy estimation of TE_{yx} because the time series described in Eqns. 13–14 are not influenced by $t-1$, $t-2$, ... time series values. The two time series used here each consisted of $20,000$ values.

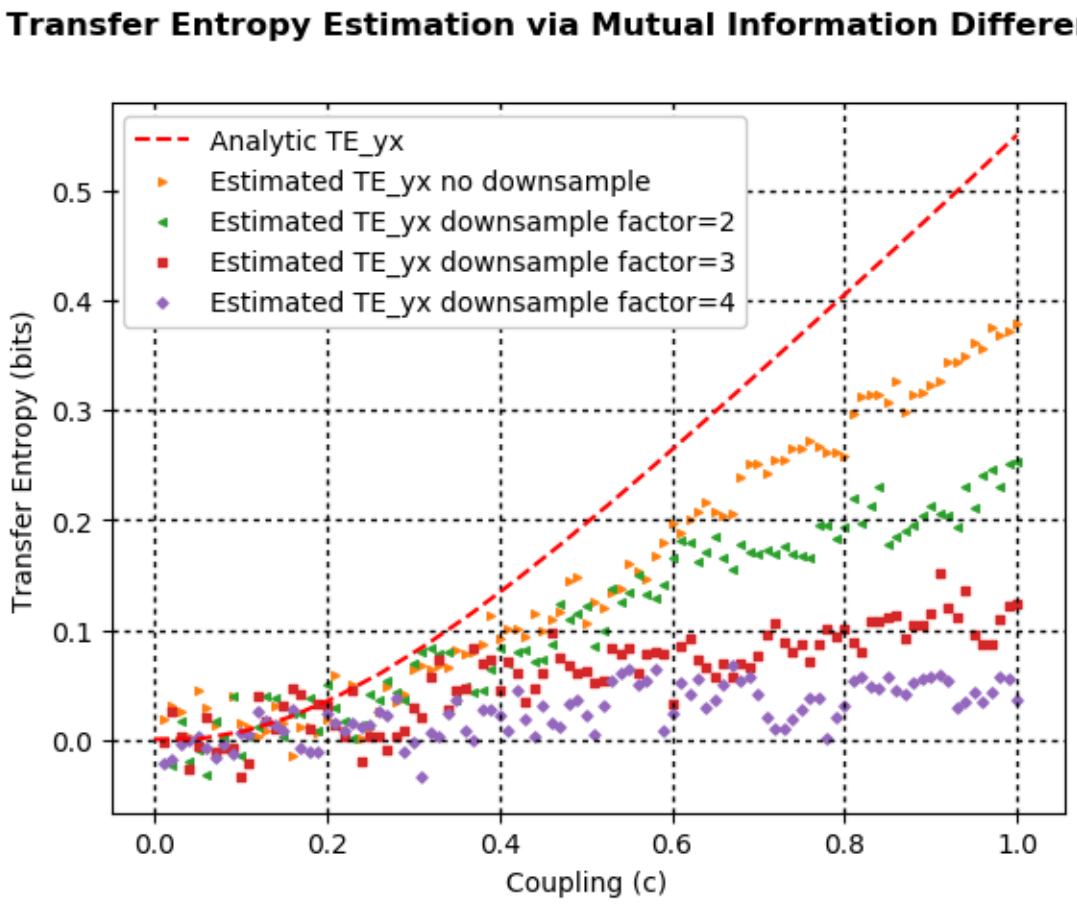

Figure 2d The analytic and estimated transfer entropy for the coupled processes in Eqns. 13–14 are shown here. The estimated transfer entropy used $k = 1$ in all cases but downsampled the processes to mimic what would happen in the event of insufficiently fast sampling of a system. As the downsampling is increased, the estimated transfer entropy also drops substantially. The time series lengths used for estimation in all cases was $20,000$ points after downsampling.

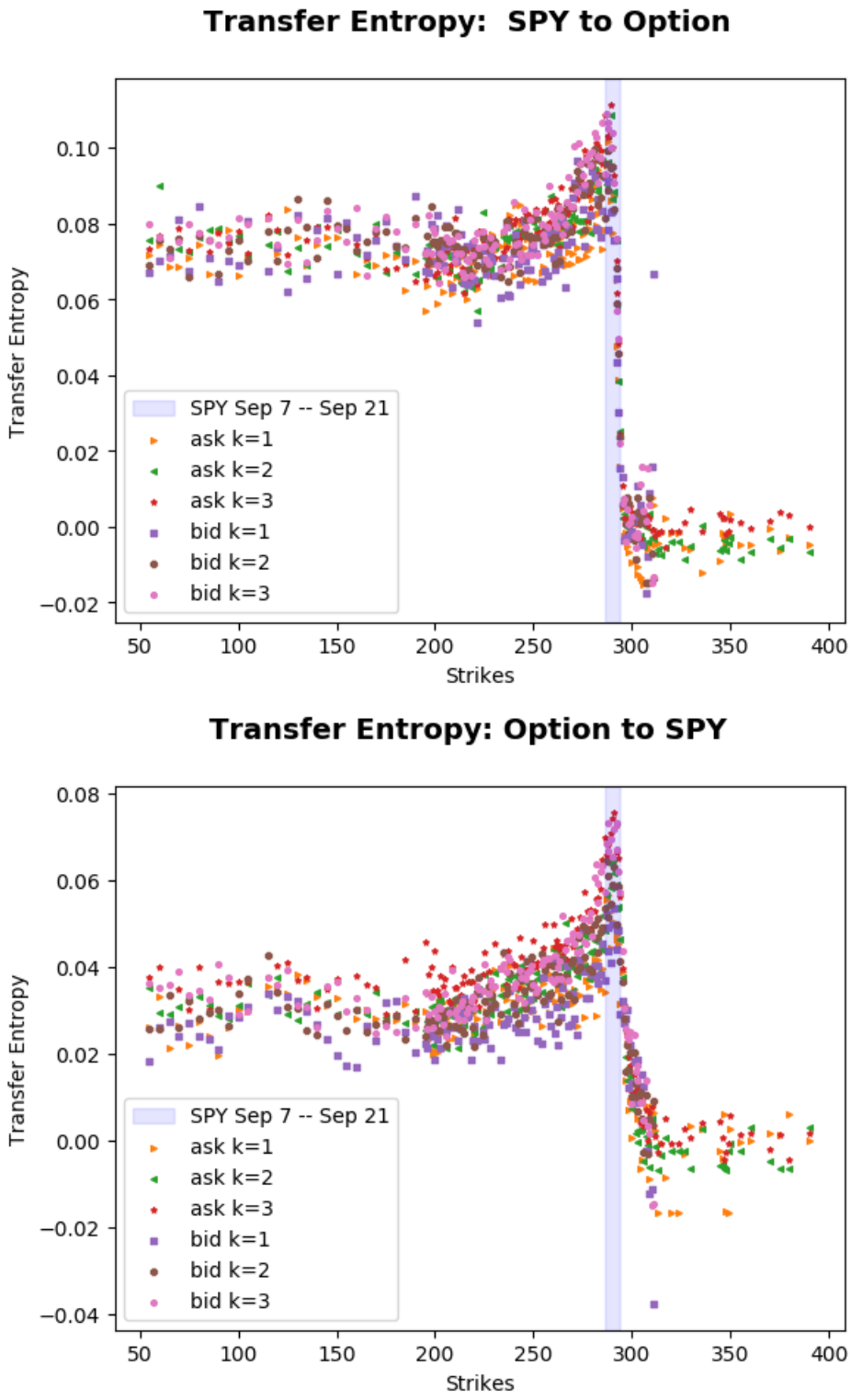

Figure 3a Estimated transfer entropy between SPY and its options for lookback set to be one and timescale set to be two seconds. On the top is the transfer entropy from the underlying to the option; on the bottom is the transfer entropy from the option to the underlying. The shaded region indicates the range over which the SPY price varied over the sample period. Results for several different k values used in the Kraskov estimator are presented.

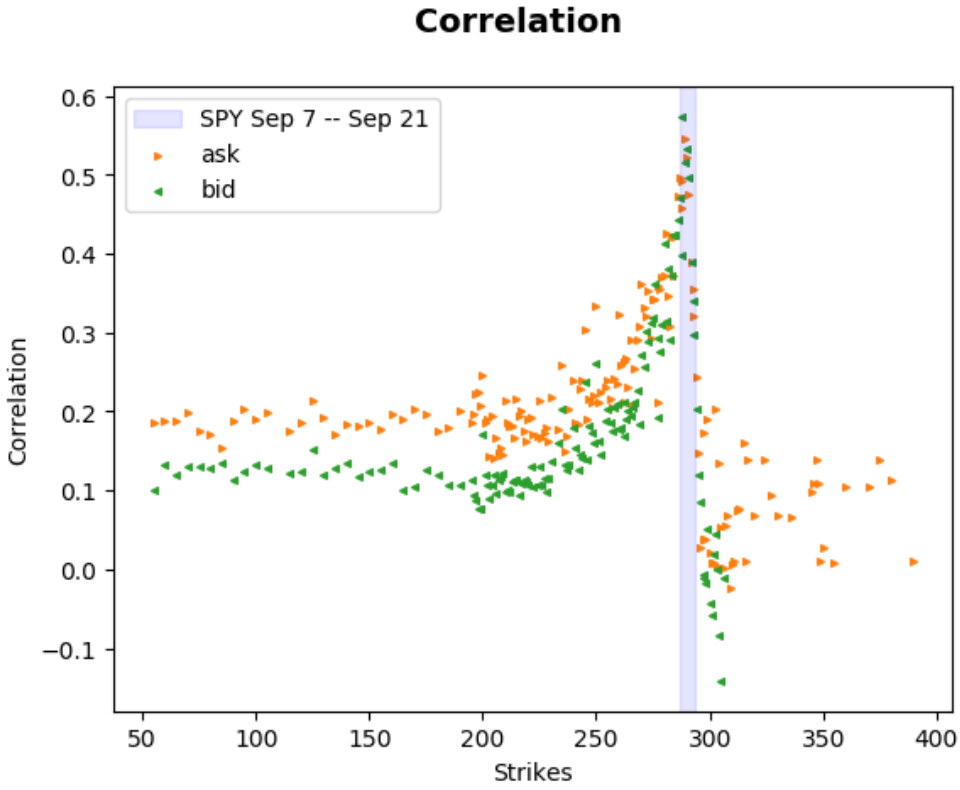

Figure 3b Correlation between the equity and option's absolute value of returns for the SPY ticker. The shaded region indicates the range over which the SPY price varied over the sampling period.

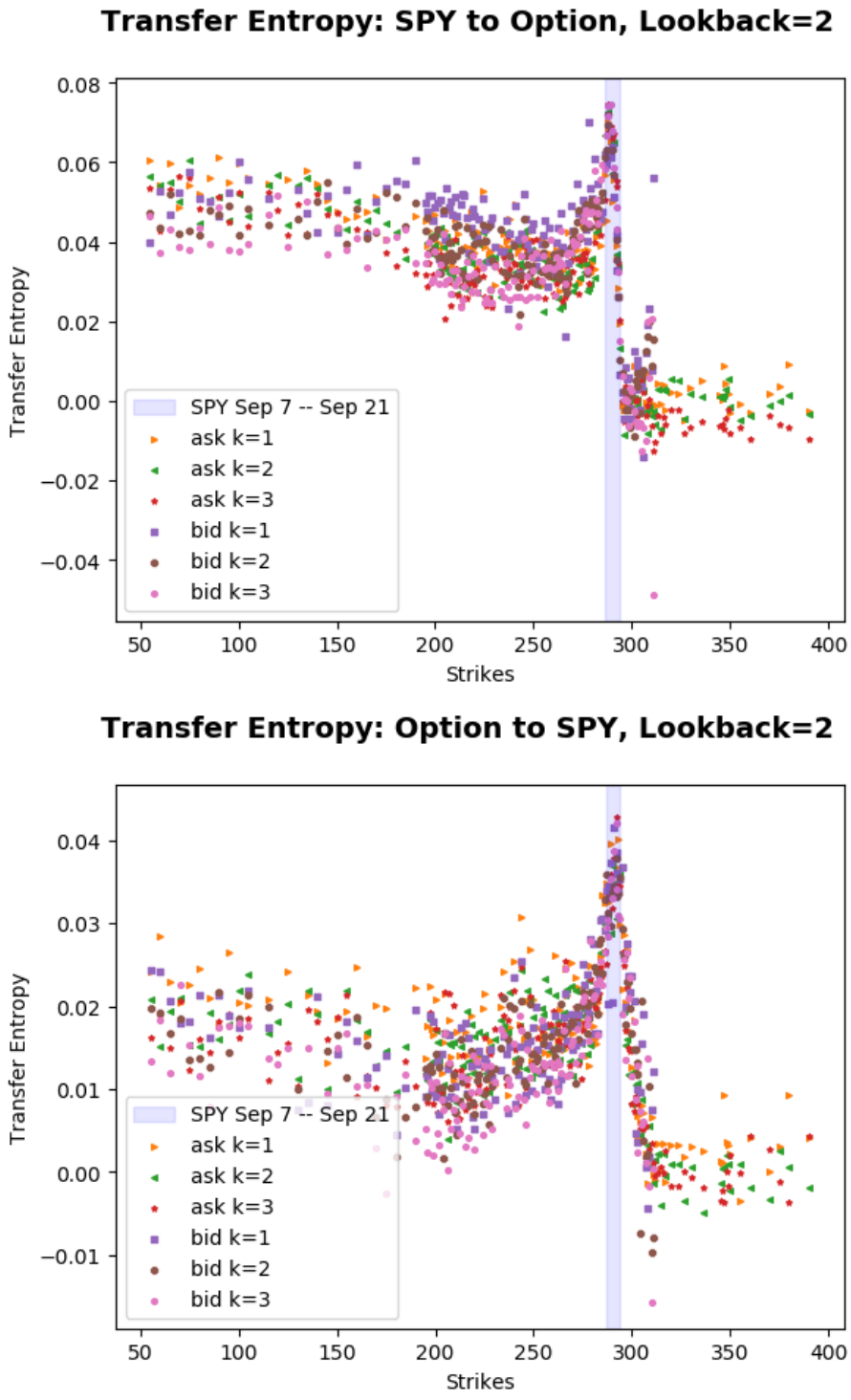

Figure 3c Estimated transfer entropy between SPY and its options for lookback set to be two and timescale set to be two seconds. On the top is the transfer entropy from the underlying to the option; on the bottom is the transfer entropy from the option to the underlying. The shaded region indicates the range over which the SPY price varied over the sample period. The transfer entropy is much less than that estimated for lookback set to be one (See Figure 3a) supporting that the time series are first order Markov chains.

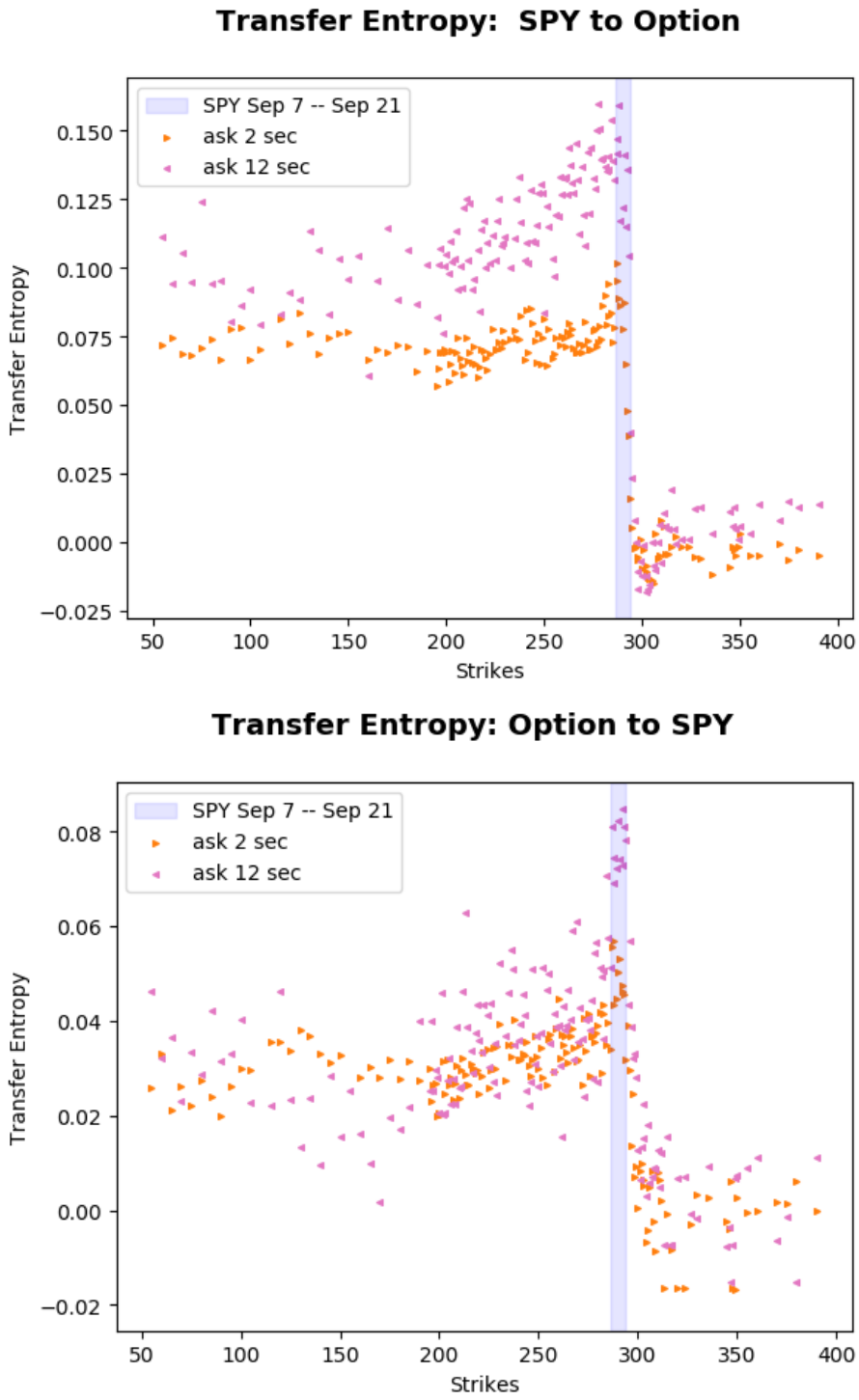

Figure 3d Estimated transfer entropy between SPY and its options where the price data has been sampled every two seconds and every 12 seconds. The 12 second timescale is the peak transfer entropy timescale for SPY to its options in the range between two seconds and 20 seconds. The shaded region indicates the range over which the SPY price varied over the sample period. The 12 second sampling timescale gives a significantly higher transfer entropy indicating that more information is moving between SPY and its options at this timescale than the others explored.

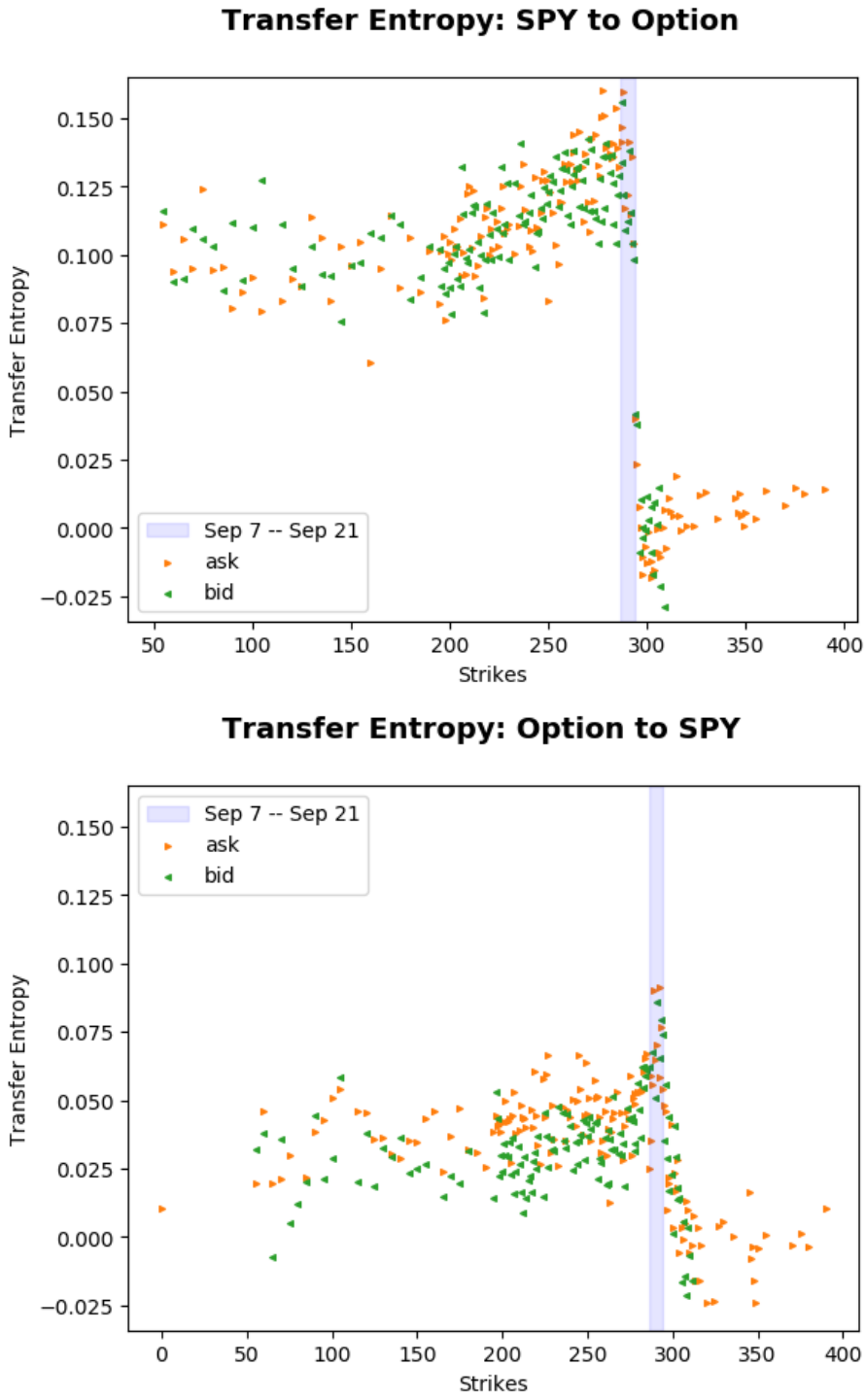

Figure 3e Estimated transfer entropy between SPY and its options. The maximum at-the-money transfer entropy from the underlying to the option is realized when sampling the data every 12 seconds and every 14 seconds for the transfer entropy from the option to the underlying. The shaded region indicates the range over which the SPY price varied over the sample period. For SPY, the maximum at-the-money transfer entropy occurs with a lookback of 1 suggesting that these data are consistent with the efficient market hypothesis.

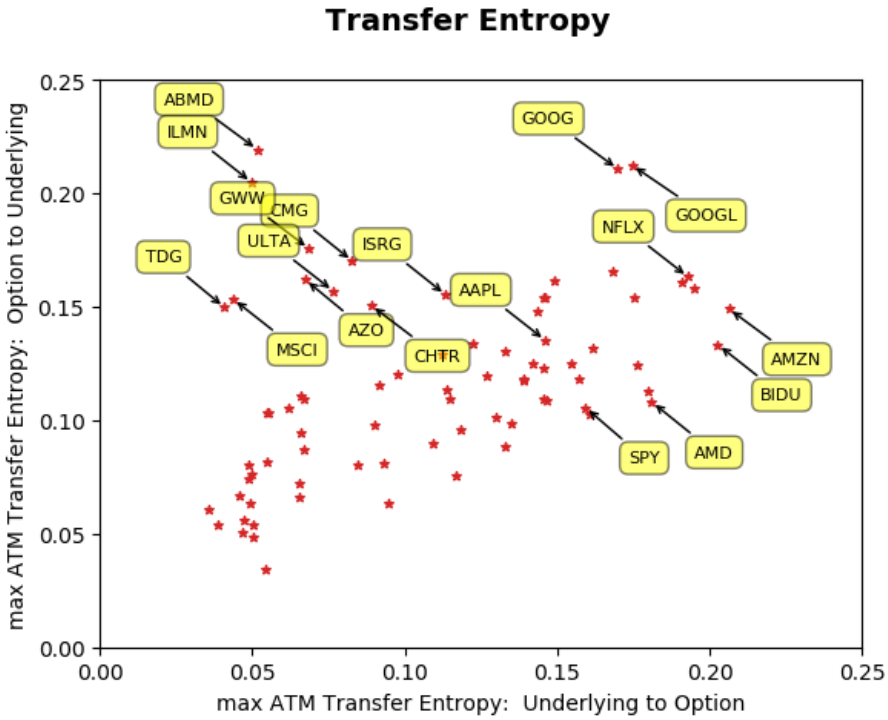

Figure 4a Estimated transfer entropy between the 78 tickers and their options; the domain is the transfer entropy from the underlying to the option while the range is the transfer entropy from the option to the underlying. Nearly all tickers show greater transfer entropy from the underlying to their respective options except a few notable outliers: ABMD, ILMN, ISRG, MSCI, TDG, GWW, AZO,ULTA, CMG, CHTR, GOOG, and GOOGL. The largest outliers, ABMD and ILMN, both service the medical industry. Complete data for all tickers are found in Table 2.

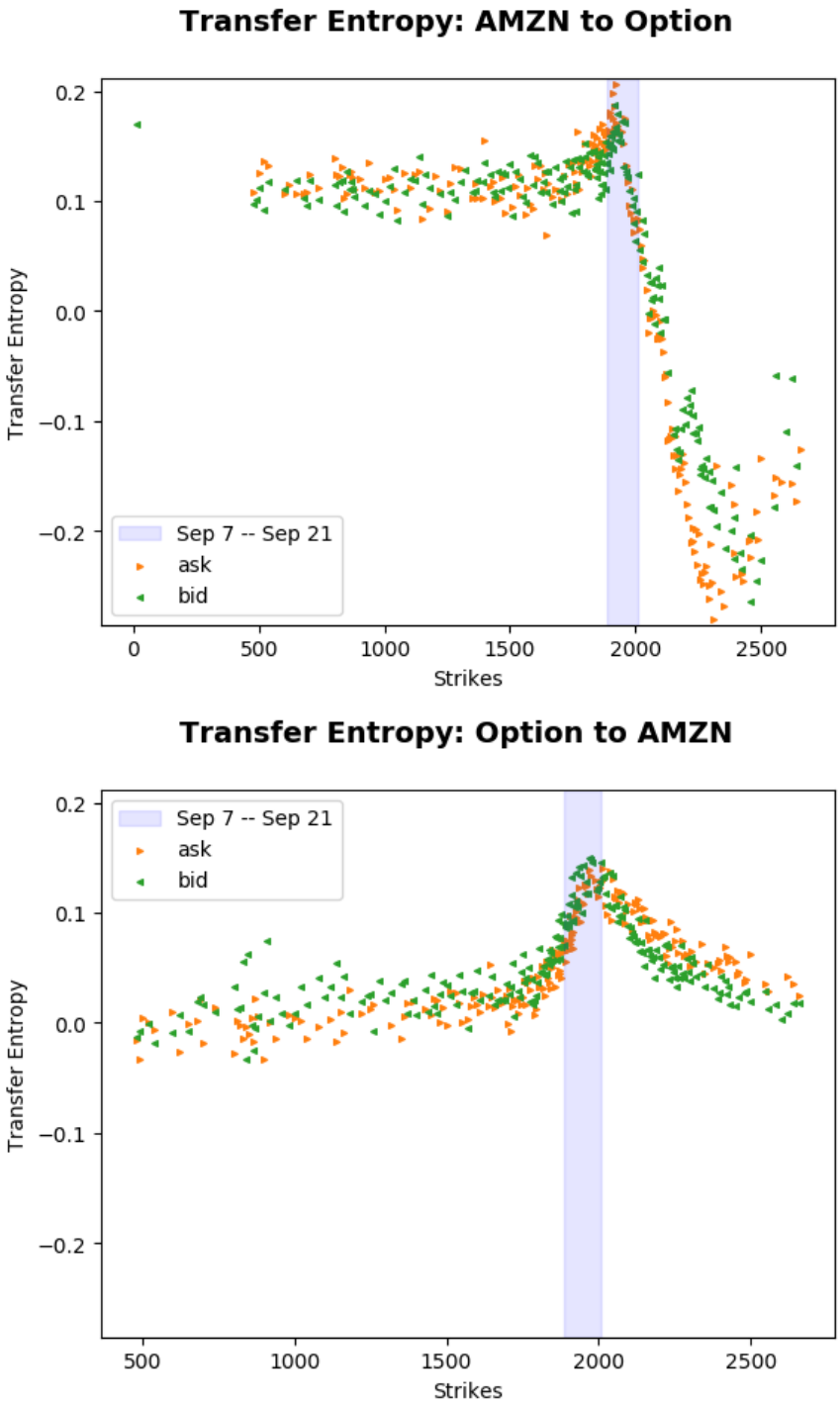

Figure 4b Estimated transfer entropy between AMZN and its options at timescales of 10 seconds and 14 seconds for the underlying to option and option to underlying, respectively. The shaded region indicates the range over which the AMZN price varied over the sample period. The transfer entropy reported here is representative of most of the tickers explored: a nonzero transfer entropy from the option to the underlying but a stronger signal in the conventional direction of underlying to option. Like most tickers explored, the peak transfer entropy occurs here for options that were at-the-money in the data collection period.

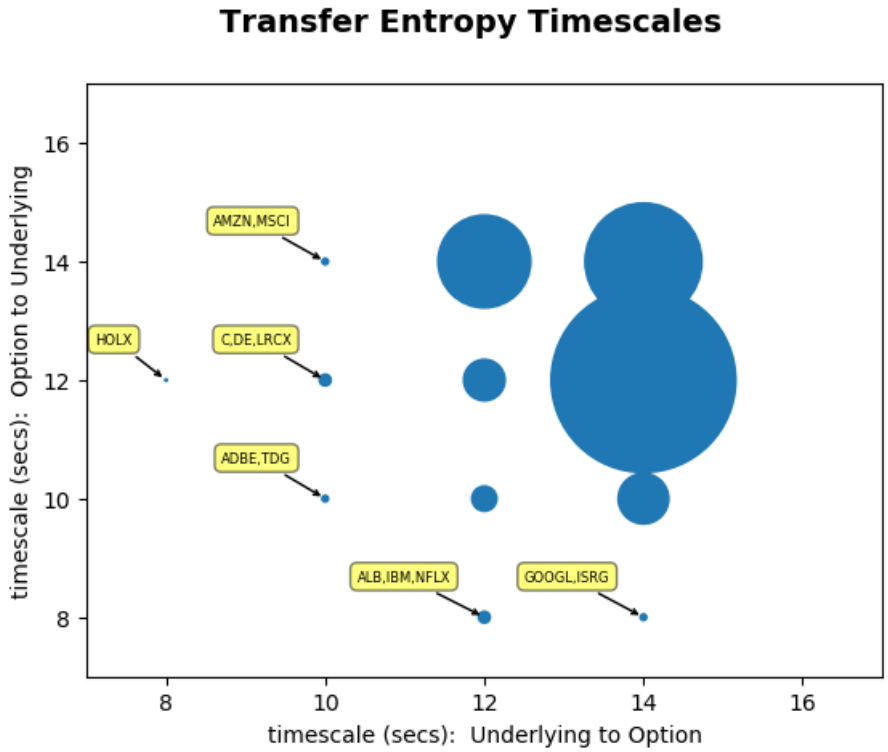

Figure 4c The timescales giving the largest at-the-money estimated transfer entropies for each of the 78 tickers and their options; the domain is the timescale for the transfer entropy from the underlying to the option while the range is the timescale for the transfer entropy from the option to the underlying. The size of the circle in the graph reflects the number of tickers sharing that same maximum transfer entropy timescale. Most tickers show greater transfer entropy around the 12 to 14 second timescales. When there is a timescale faster than 12 to 14 seconds, it is more likely that the faster timescale is for the transfer entropy from the option to the underlying rather than in the opposite direction. Complete data for all tickers are found in Table 2.

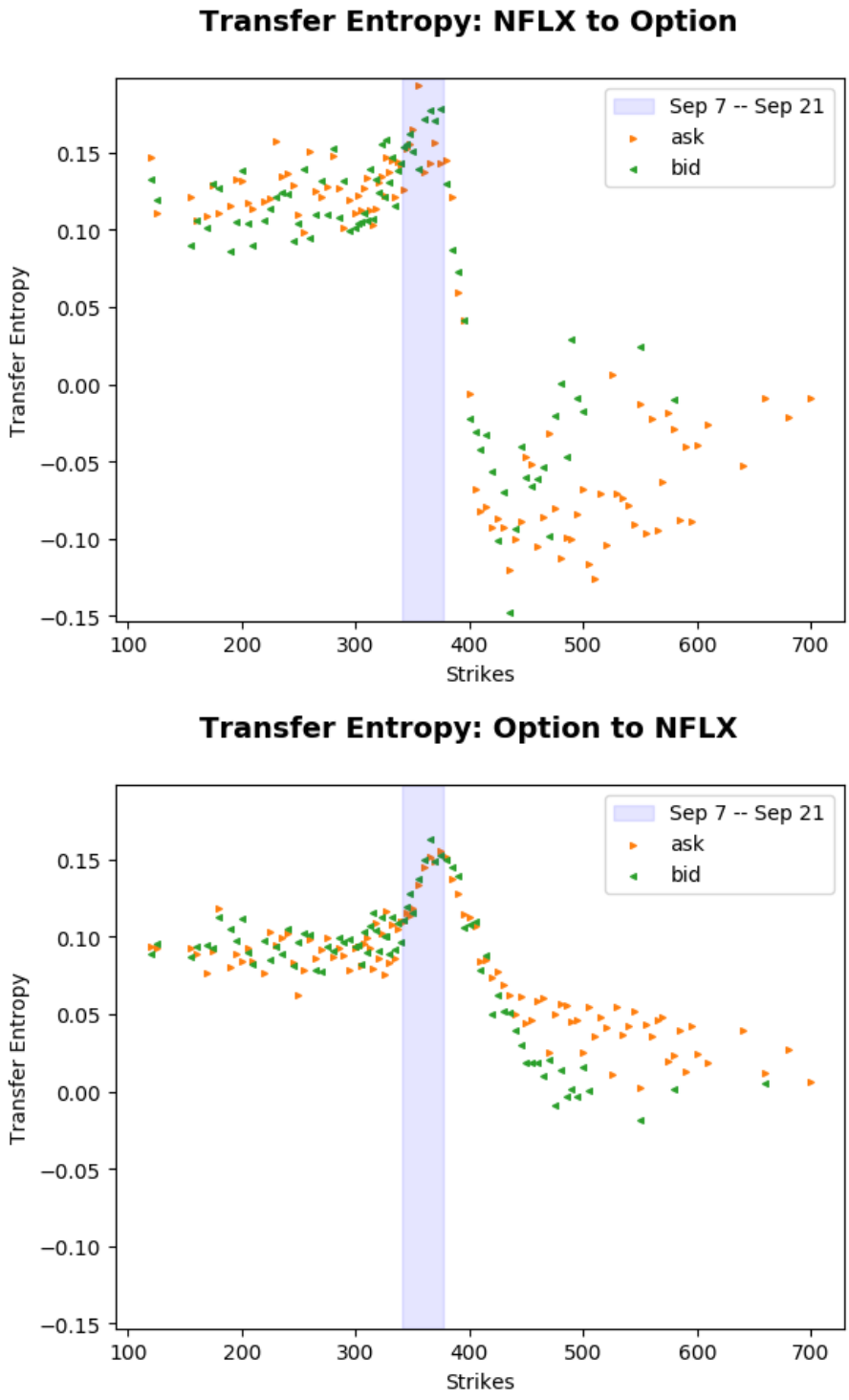

Figure 4d Estimated transfer entropy between NFLX and its options. The peak transfer entropy timescale from the option to the underlying was 8 seconds whereas the peak transfer entropy timescale from the underlying to the option was much slower at 12 seconds. Most tickers examined show similar timescales for both directions of the transfer entropy, but when this is not true, the timescale in the direction from the option to the underlying is most often faster.

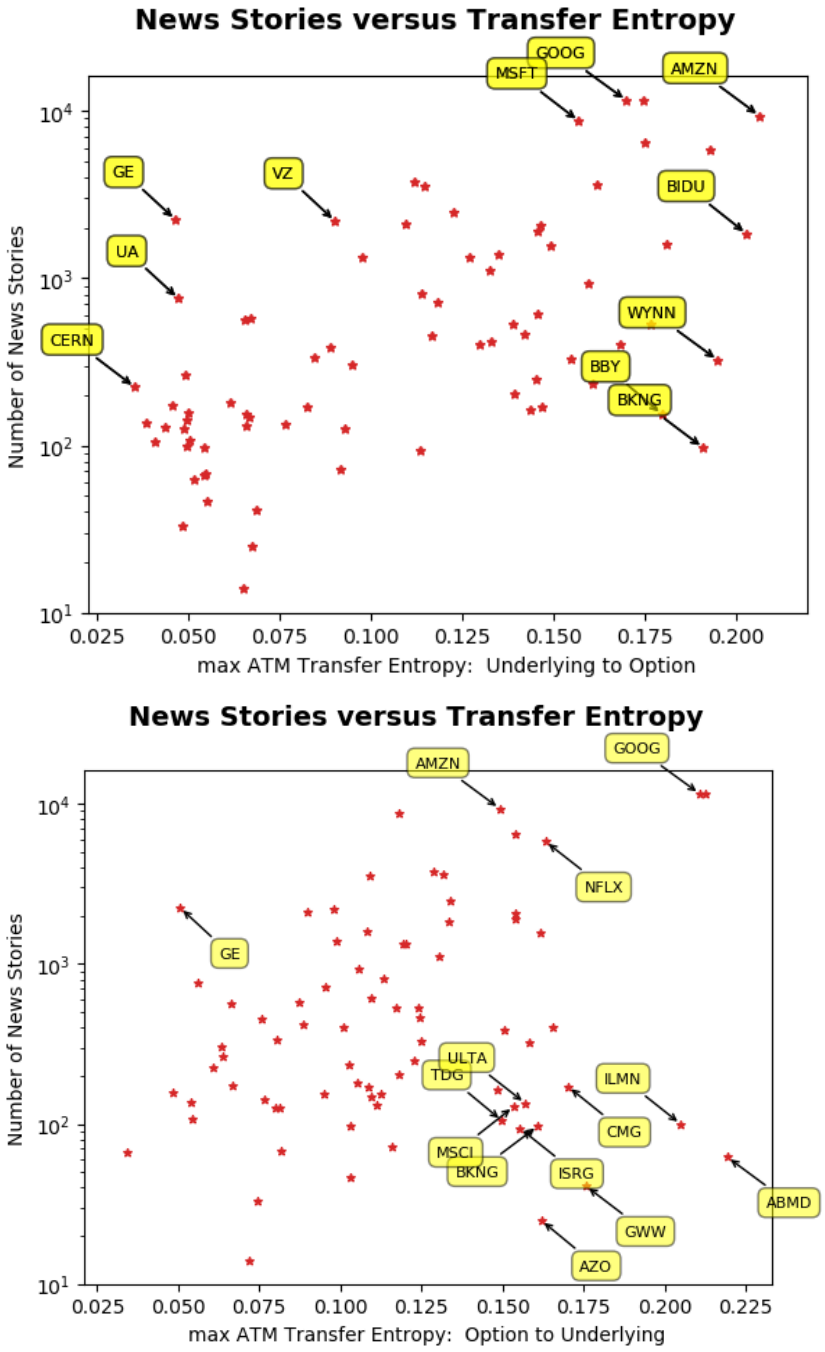

Figure 5a The daily number of stories published per ticker summed from 7 September through 21 September and plotted against the max at-the-money (ATM) transfer entropy. The top figure uses transfer entropy in the direction from the underlying to the option while the bottom figure uses transfer entropy in the direction from the option to the underlying. The daily number of stories and transfer entropy generally appear to show an exponential relationship with higher news stories resulting in higher transfer entropy with some exceptions such as GE and others. This exponential relationship is even sharper when using transfer entropy from the option to the underlying. Exceptions to this include some of the outliers from Figure 4a where in spite of a relatively small number of published news stories a high transfer entropy was found.

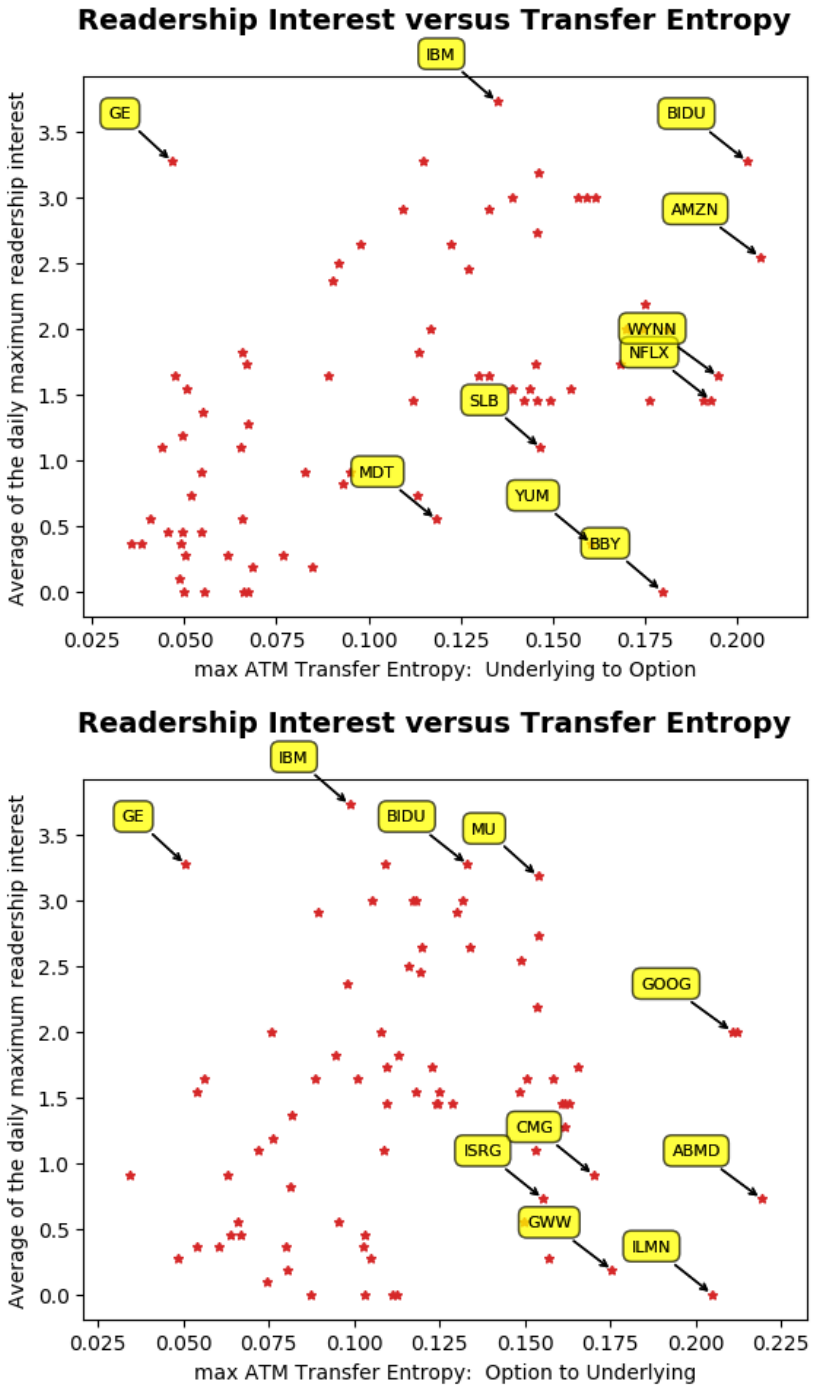

Figure 5b The average of the maximum value of readership interest in a company relative to the previous 30 days plotted against the max ATM transfer entropy. The top figure uses transfer entropy from the underlying to the option while the bottom uses transfer entropy from the option to the underlying. The readership interest is scored from 0 to 4, with a score of 1 through 4 indicating unusually high readership.

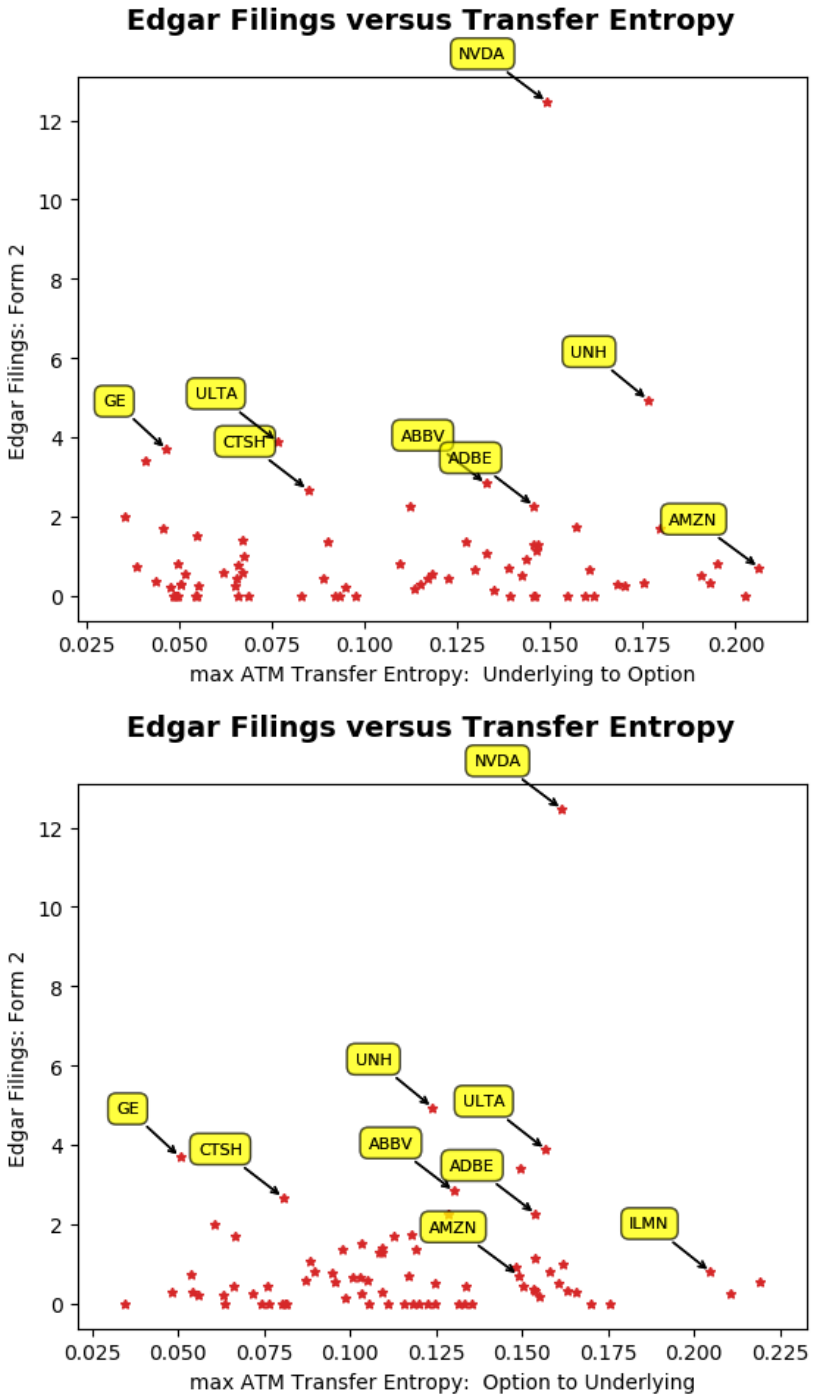

Figure 5c Abnormal EDGAR filing activity plotted against the max ATM transfer entropy. The top figure uses transfer entropy from the underlying to the option while the bottom uses transfer entropy from the option to the underlying.

Table 1 List of the companies and number of options included in this study. The min and max of the ask and bid during the period of 7 Sep 2018 through 21 Sep 2018 are also included.

Ticker	Company Name	Options	Min ask/bid	Max ask/bid
AAPL	Apple	62	215.3	228.35
ABBV	AbbVie Inc	47	90.39	96.41
ABMD	Abiomed	39	366.38	399.62
ADBE	Adobe Inc	49	257.12	277.68
ALB	Albemarle Corporation	21	95.78	107.43
AMZN	Amazon.com, Inc.	246	1887.53	2008.76
AVGO	Broadcom Inc	51	215.97	248.78
AZO	AutoZone, Inc.	50	707.95	783.47
BA	Boeing Co	65	337.26	372.58
BBY	Best Buy Co Inc	44	76.44	81.54
BIDU	Bidu, Inc.	50	208.5	234.5
BKNG	Booking Holdings Inc	133	1845	1964.59
BMY	Bristol-Myers Squibb Co	27	59.57	62.62
C	Citigroup Inc	45	68.91	75.24
CAG	Conagra Brands Inc	15	35.8	38.42
CAT	Caterpillar Inc.	53	140.01	157.72
CELG	Celgene Corporation	51	86.33	91.25
CERN	Cerner Corporation	15	62.01	66.05
CHTR	Charter Communications Inc	59	291	334.88
CMG	Chipotle Mexican Grill, Inc.	81	465.2	499.75
CTSH	Cognizant Technology Solutions Corp	20	75.1	77.86
DE	Deere & Company	43	144.02	156.8
EIX	Edison International	8	66.44	69.88
EW	Edwards Lifesciences Corp	21	140.07	154.28
FB	Facebook, Inc. Common Stock	49	159	167.26
FDX	FedEx Corporation	26	238.57	259.3
FMC	FMC Corp	10	81.02	91.08
GE	General Electric	29	12.17	12.99
GILD	Gilead Sciences, Inc.	47	71.49	75.93
GOOG	Alphabet Inc Class C	110	1153.63	1192.21
GOOGL	Alphabet Inc Class A	110	1158.25	1197
GT	Goodyear Tire & Rubber Co	12	22.48	24.85
GWW	W W Grainger Inc	28	345.2	361.69
HAL	Halliburton Company	24	35.75	40.81
HBAN	Huntington Bancshares Incorporated	8	15.41	16.55
HFC	HollyFrontier Corp	33	64.08	72.79
HOLX	Hologic, Inc.	20	37.89	41.67
HP	Helmerich & Payne, Inc.	19	61.53	67.89

IBM	IBM Common Stock	39	145.11	152.6
ILMN	Illumina, Inc.	45	342.81	360.82
INTC	Intel Corporation	41	44.06	47.49
ISRG	Intuitive Surgical, Inc.	45	534	572.99
KMB	Kimberly-Clark Corp	17	114.36	117.59
KORS	Michael Kors Holdings Ltd	24	71.76	74.68
KSU	Kansas City Southern	27	116.21	120.22
LMT	Lockheed Martin Corporation	51	320.08	345.88
LRCX	Lam Research Corporation	49	149.02	161.59
MDT	Mettler-Toledo International Inc.	34	95.51	98
MNST	Monster Beverage Corp	25	57.76	61.34
MO	Altria Group Inc	44	59.12	64.02
MSCI	MSCI Inc	16	172.78	184
MSFT	Microsoft Corporation	53	107.24	114.98
MU	Micron Technology, Inc.	40	40.69	47.22
NFLX	Netflix, Inc.	104	341.26	377.61
NVDA	NVIDIA Corporation	65	261.94	279.1
PM	Philip Morris International Inc.	51	77.17	83.85
PYPL	Paypal Holdings Inc	47	87.39	93.71
QCOM	Qualcomm, Inc.	29	69.62	76.04
SJM	The J.M. Smucker Company	12	107.53	111.83
SLB	Schlumberger Limited	41	59.25	62.27
SPG	Simon Property Group Inc	12	178.17	186.46
SPY	SPDR S&P 500 ETF Trust	167	286.72	293.93
STT	State Street Corp	8	85.18	90.21
TDG	TransDigm Group Incorporated	9	349.89	374.99
TRIP	Tripadvisor Inc Common Stock	44	48.34	54.13
TSN	Tyson Foods, Inc.	20	61.22	64.07
UA	Under Armour Inc Class A	15	17.03	19.19
ULTA	Ulta Beauty Inc	53	274.99	290.42
UNH	UnitedHealth Group Inc	48	257.45	271.2
VAR	Varian Medical Systems, Inc.	5	100.15	115.23
VLO	Valero Energy Corporation	44	108.1	117.82
VZ	Verizon Communications Inc.	30	53.31	55.42
WMT	Walmart Inc	45	94.16	97.66
WYNN	Wynn Resorts, Limited	44	126.97	139
XRX	Xerox Corp	10	26.85	28.76
YUM	Yum! Brands, Inc.	19	86.5	89.83
ZTS	Zoetis Inc	23	87.92	90.94

Table 2 The maximum at-the-money transfer entropy for each of the companies examined are listed here along with the timescales exhibiting the largest at-the-money transfer entropy.

Ticker	max ATM TE_{uo}	Timescale (secs)	max ATM TE_{ou}	Timescale (secs)
AAPL	0.146	14	0.135	12
ABBV	0.133	14	0.130	12
ABMD	0.052	12	0.219	10
ADBE	0.146	10	0.154	10
ALB	0.051	12	0.054	8
AMD	0.181	12	0.108	10
AMZN	0.206	10	0.149	14
AVGO	0.114	14	0.113	14
AZO	0.067	12	0.162	14
BA	0.162	14	0.132	12
BBY	0.180	14	0.113	14
BIDU	0.203	14	0.133	12
BKNG	0.191	14	0.161	12
BMY	0.117	12	0.076	14
C	0.109	10	0.090	12
CAG	0.038	14	0.054	12
CAT	0.139	12	0.117	14
CELG	0.145	14	0.123	14
CERN	0.036	12	0.061	10
CHTR	0.089	14	0.151	10
CMG	0.083	14	0.170	12
CTSH	0.085	14	0.081	12
DE	0.155	10	0.125	12
EIX	0.049	14	0.064	14
EW	0.067	12	0.109	14
FB	0.175	12	0.154	14
FDX	0.127	12	0.119	14
FMC	0.049	14	0.074	10
GE	0.047	12	0.051	10
GILD	0.146	14	0.109	12
GOOG	0.170	14	0.211	10

GOOGL	0.175	14	0.212	8
GT	0.050	14	0.048	10
GWW	0.069	14	0.176	14
HAL	0.095	14	0.063	12
HBAN	0.055	14	0.034	14
HFC	0.066	14	0.095	14
HOLX	0.046	8	0.067	12
HP	0.065	14	0.072	12
IBM	0.135	12	0.099	8
ILMN	0.050	14	0.205	12
INTC	0.115	14	0.109	12
ISRG	0.113	14	0.155	8
KMB	0.062	12	0.105	12
KORS	0.067	12	0.087	14
KSU	0.066	14	0.111	12
LMT	0.098	12	0.120	12
LRCX	0.144	10	0.148	12
MDT	0.118	12	0.096	12
MNST	0.055	14	0.082	12
MO	0.142	14	0.125	14
MSCI	0.044	10	0.153	14
MSFT	0.157	14	0.118	10
MU	0.146	12	0.154	12
NFLX	0.193	12	0.163	8
NVDA	0.149	12	0.162	12
PM	0.130	14	0.101	14
PYPL	0.168	14	0.166	14
QCOM	0.123	12	0.134	14
SJM	0.055	14	0.103	14
SLB	0.147	14	0.109	10
SPG	0.092	14	0.116	12
SPY	0.159	12	0.105	14
STT	0.133	12	0.088	10
TDG	0.041	10	0.150	10

TRIP	0.050	12	0.076	12
TSN	0.066	12	0.066	12
UA	0.047	14	0.056	10
ULTA	0.077	12	0.157	14
UNH	0.176	12	0.124	14
VAR	0.055	14	0.103	12
VLO	0.139	14	0.118	14
VZ	0.090	14	0.098	10
WMT	0.112	14	0.129	12
WYNN	0.195	12	0.158	14
XRX	0.093	14	0.081	14
YUM	0.161	14	0.103	12
ZTS	0.049	14	0.080	14

6 Online Appendix

This appendix presents plots of the estimated transfer entropies between each equity securities and the options on the equity as a function of the options' strike prices for each ticker in our sample. We depict the TE estimates with a timescale and lookback where TE between the equity and at-the-money options peaks. The shaded region indicates the range over which the equity price varied during the data sampling period.

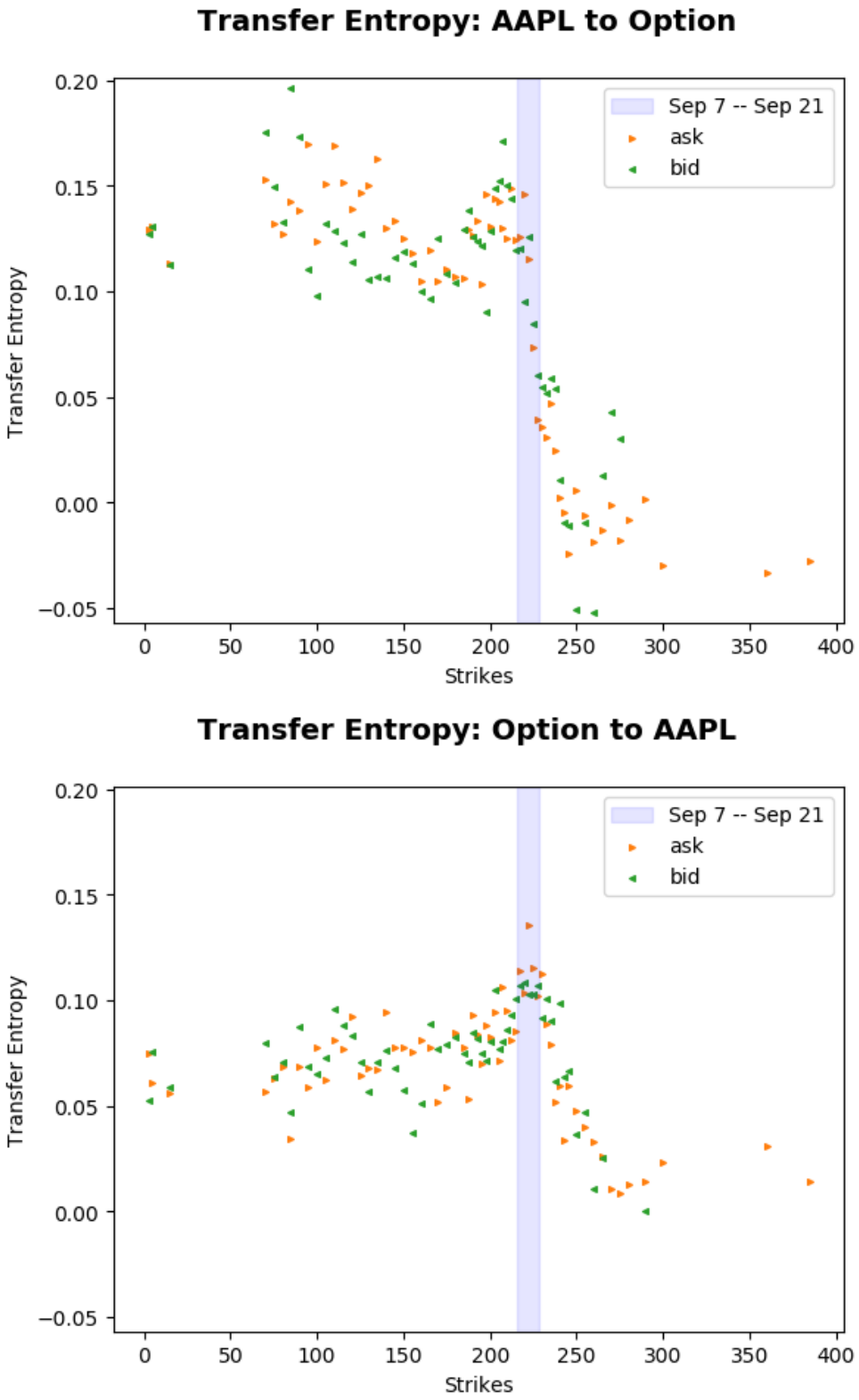

Figure 1 Transfer entropy between AAPL and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

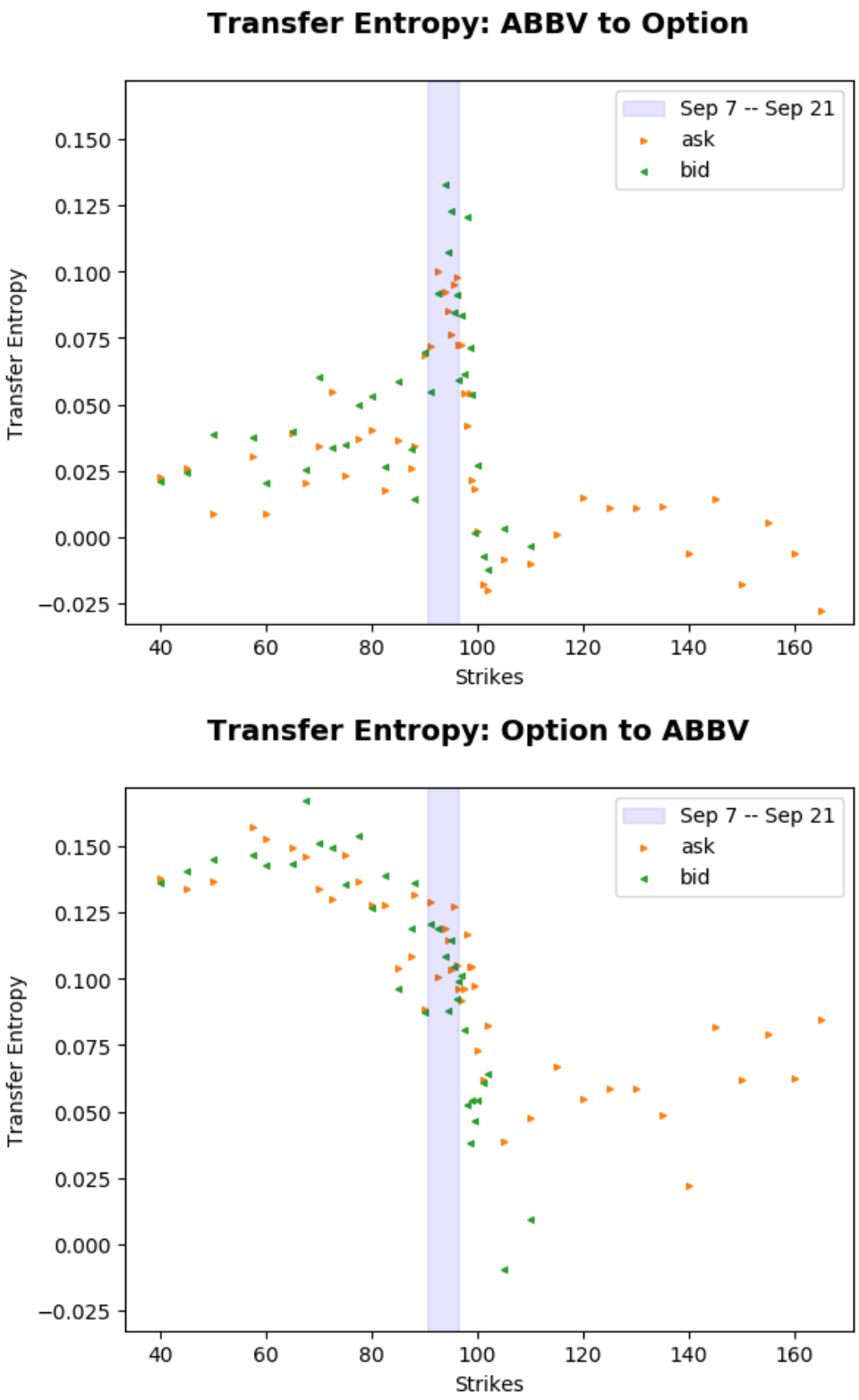

Figure 2 Transfer entropy between ABBV and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

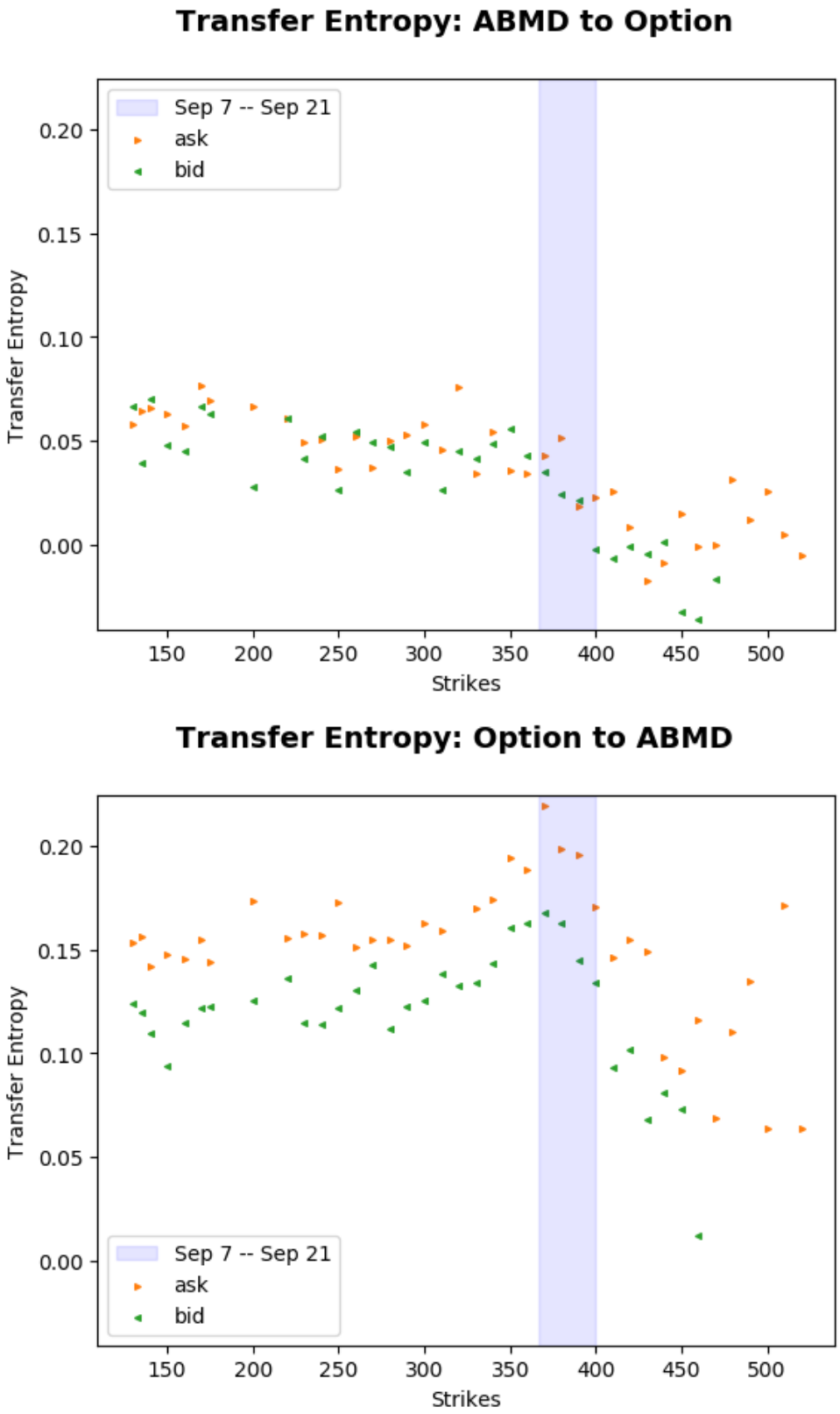

Figure 3 Transfer entropy between ABMD and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 10 seconds.

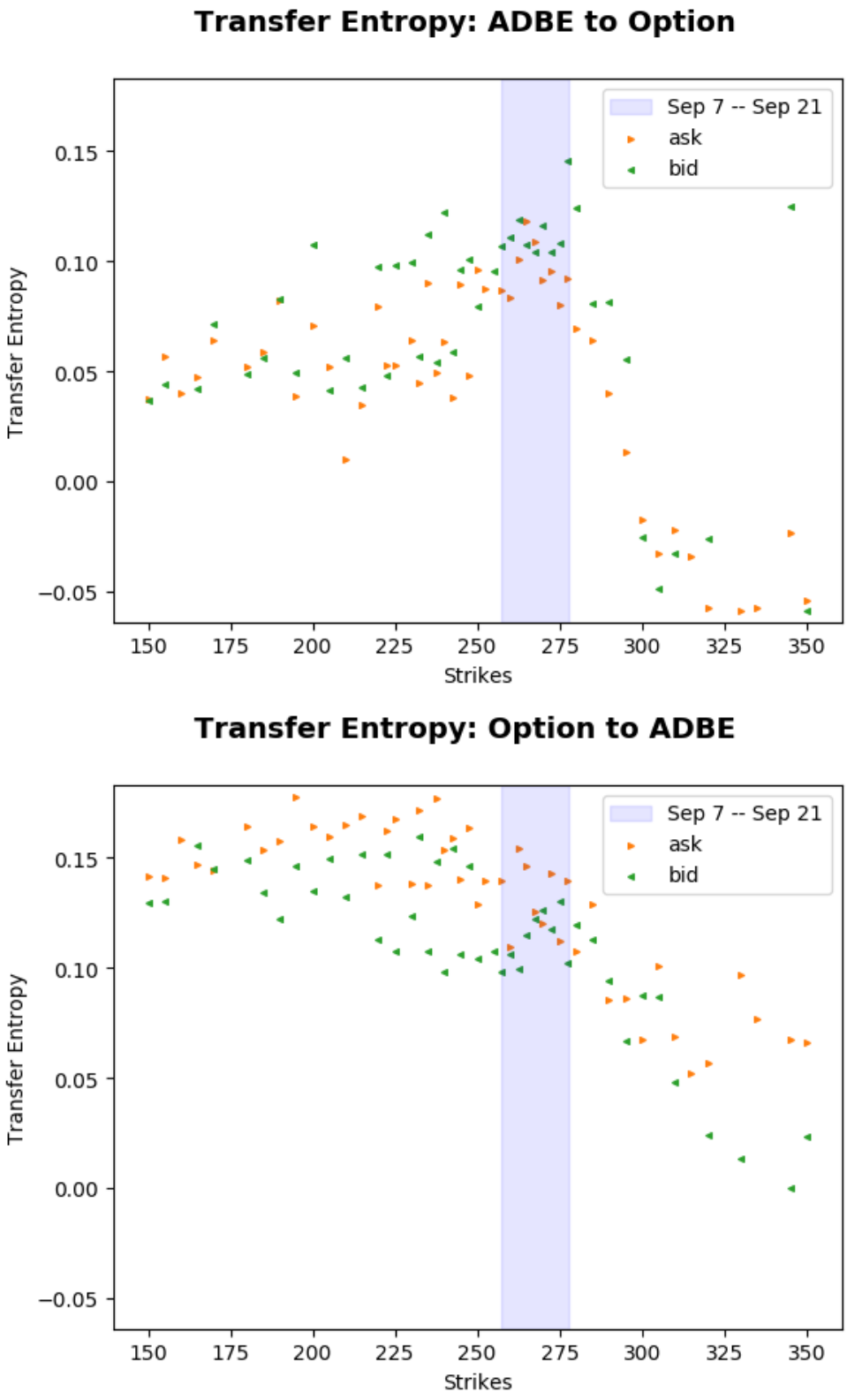

Figure 4 Transfer entropy between ADBE and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 10 seconds. Option to underlying timescale: 10 seconds.

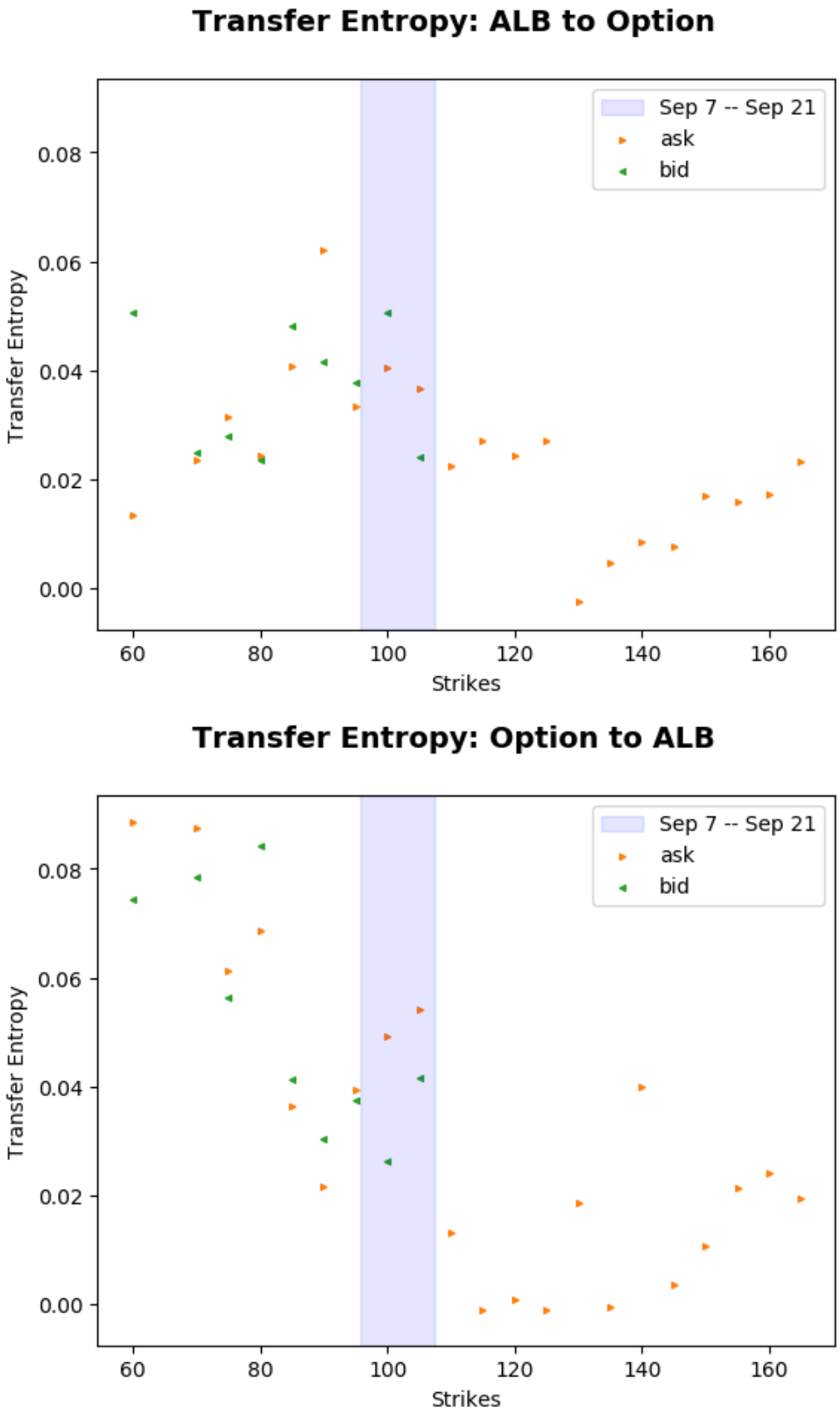

Figure 5 Transfer entropy between ALB and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 8 seconds.

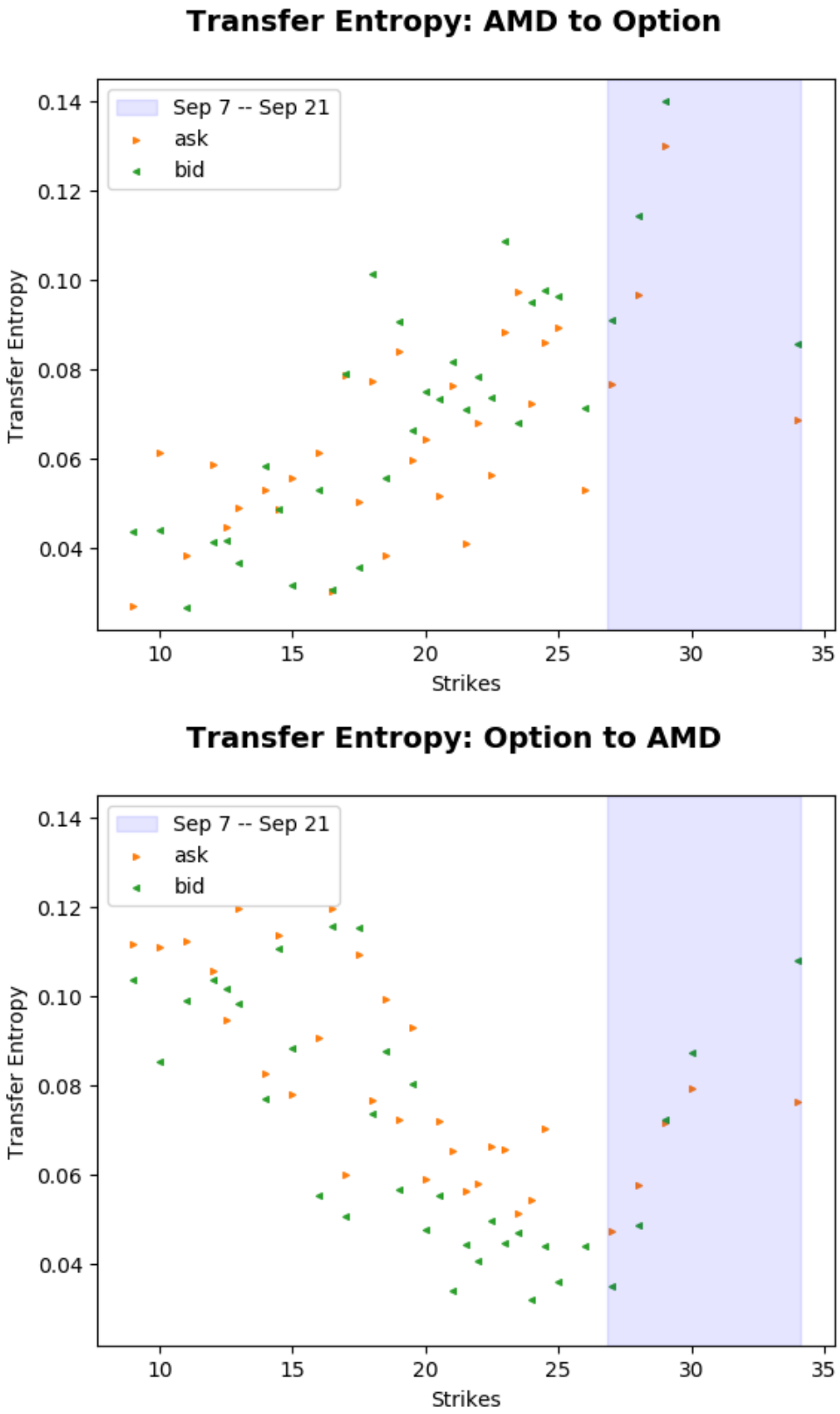

Figure 6 Transfer entropy between AMD and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 10 seconds.

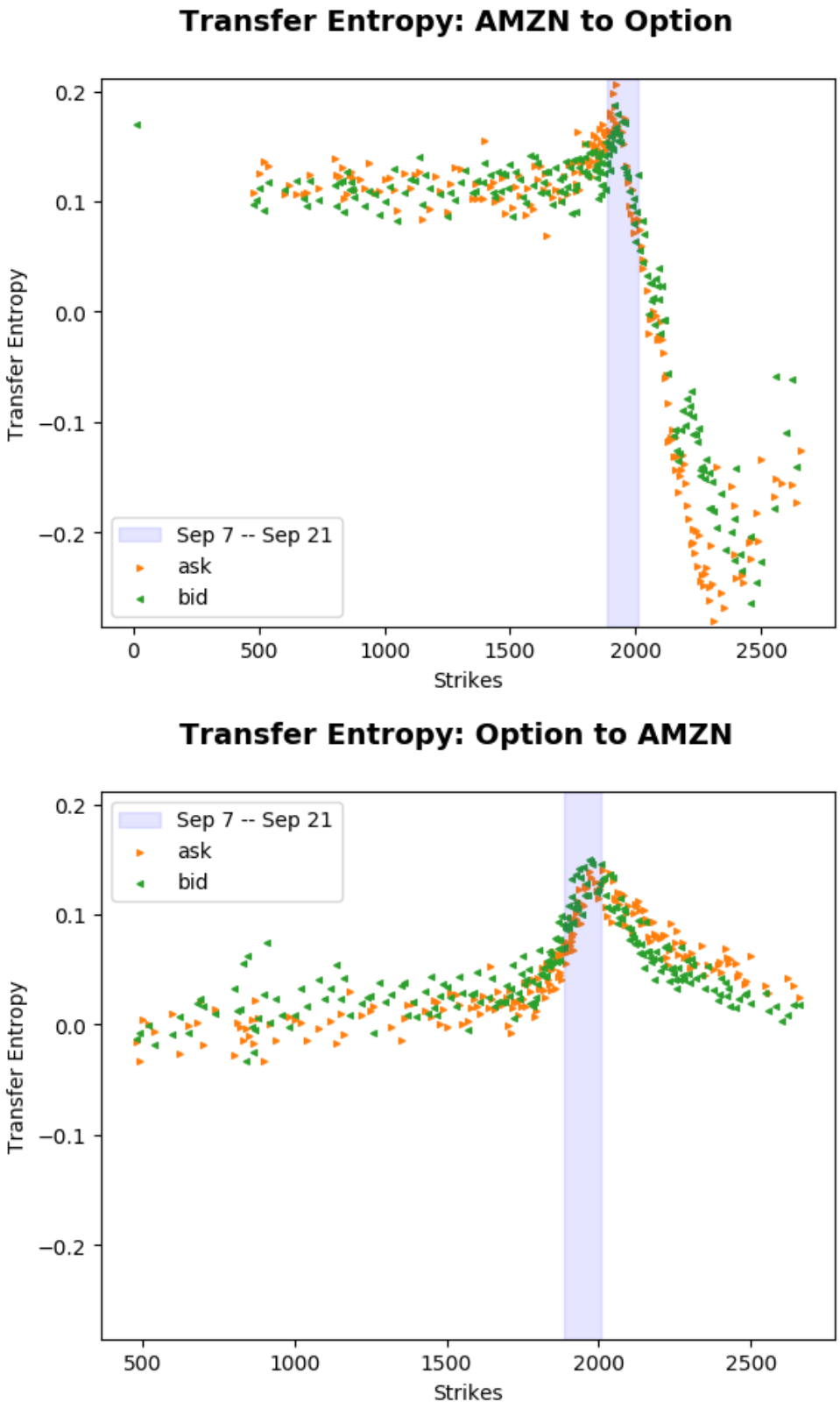

Figure 7 Transfer entropy between AMZN and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 10 seconds. Option to underlying timescale: 14 seconds.

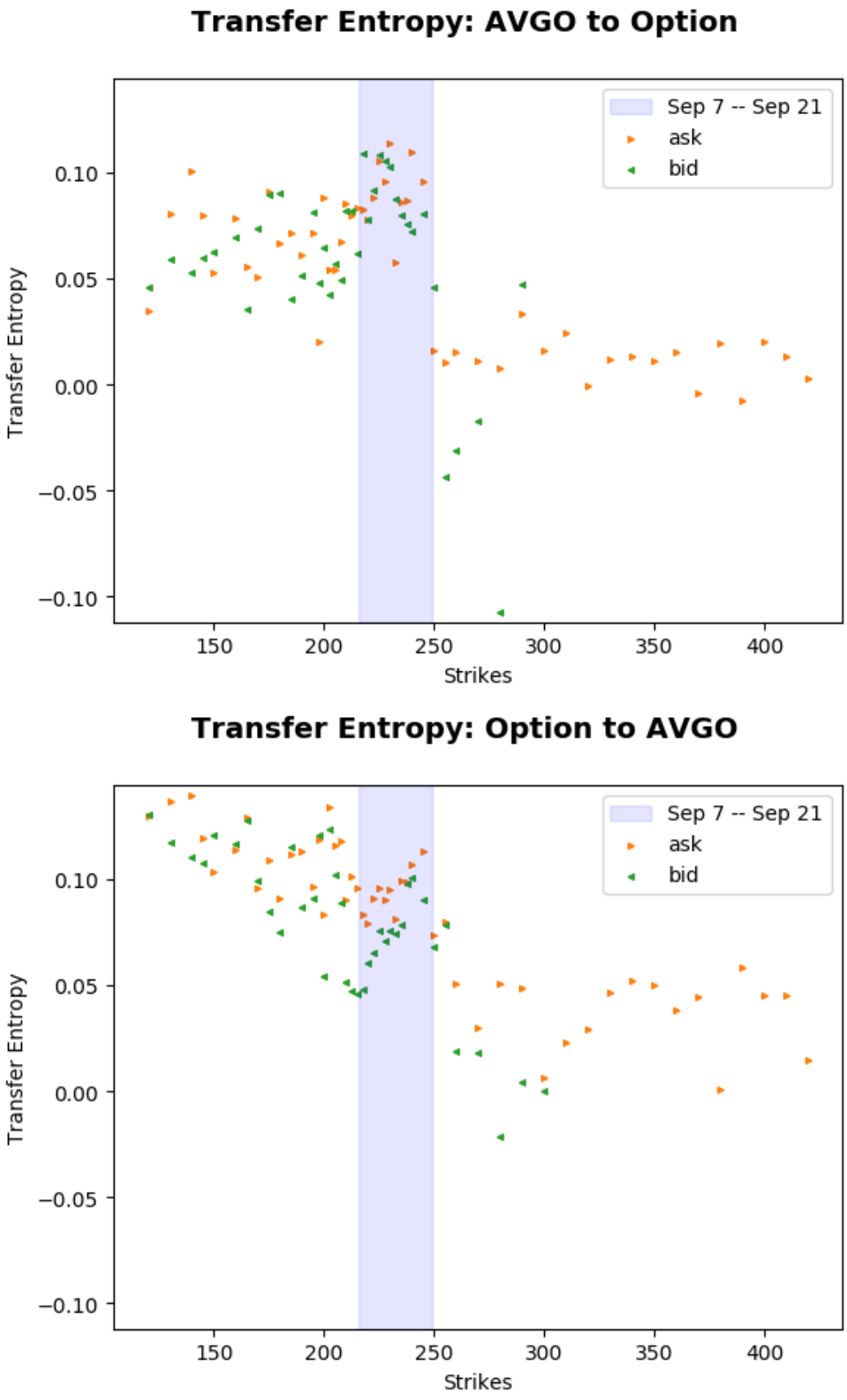

Figure 8 Transfer entropy between AVGO and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

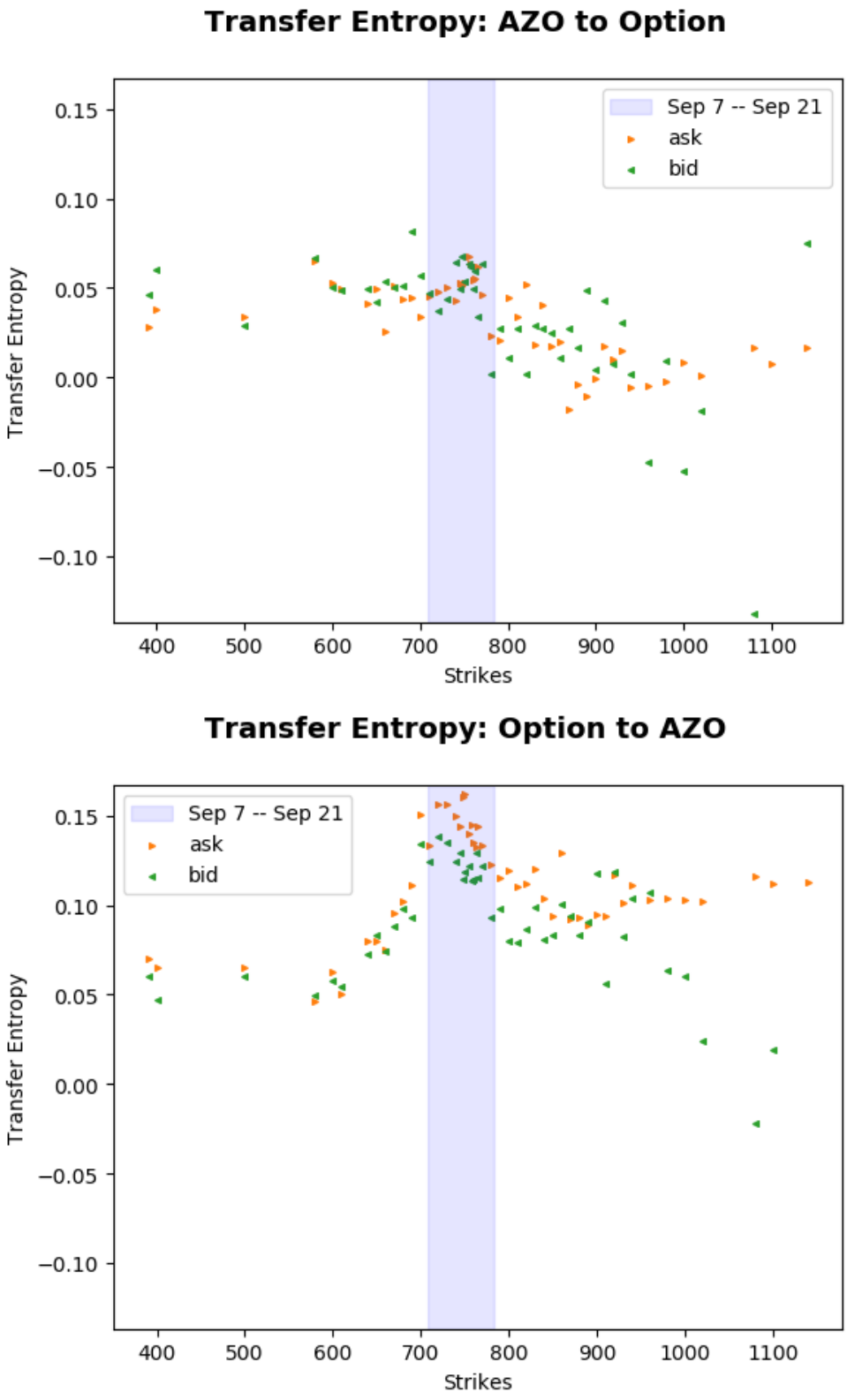

Figure 9 Transfer entropy between AZO and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

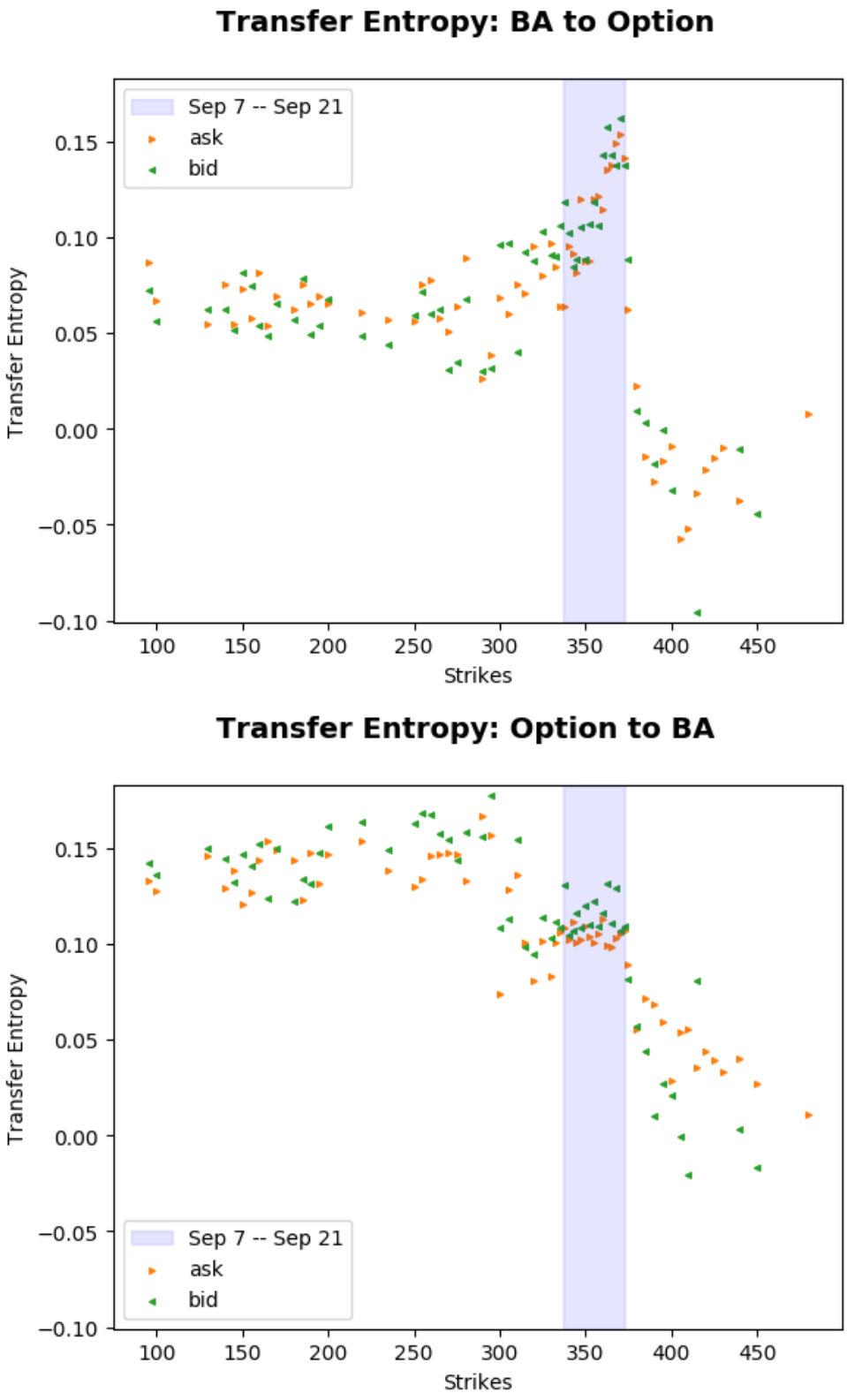

Figure 10 Transfer entropy between BA and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

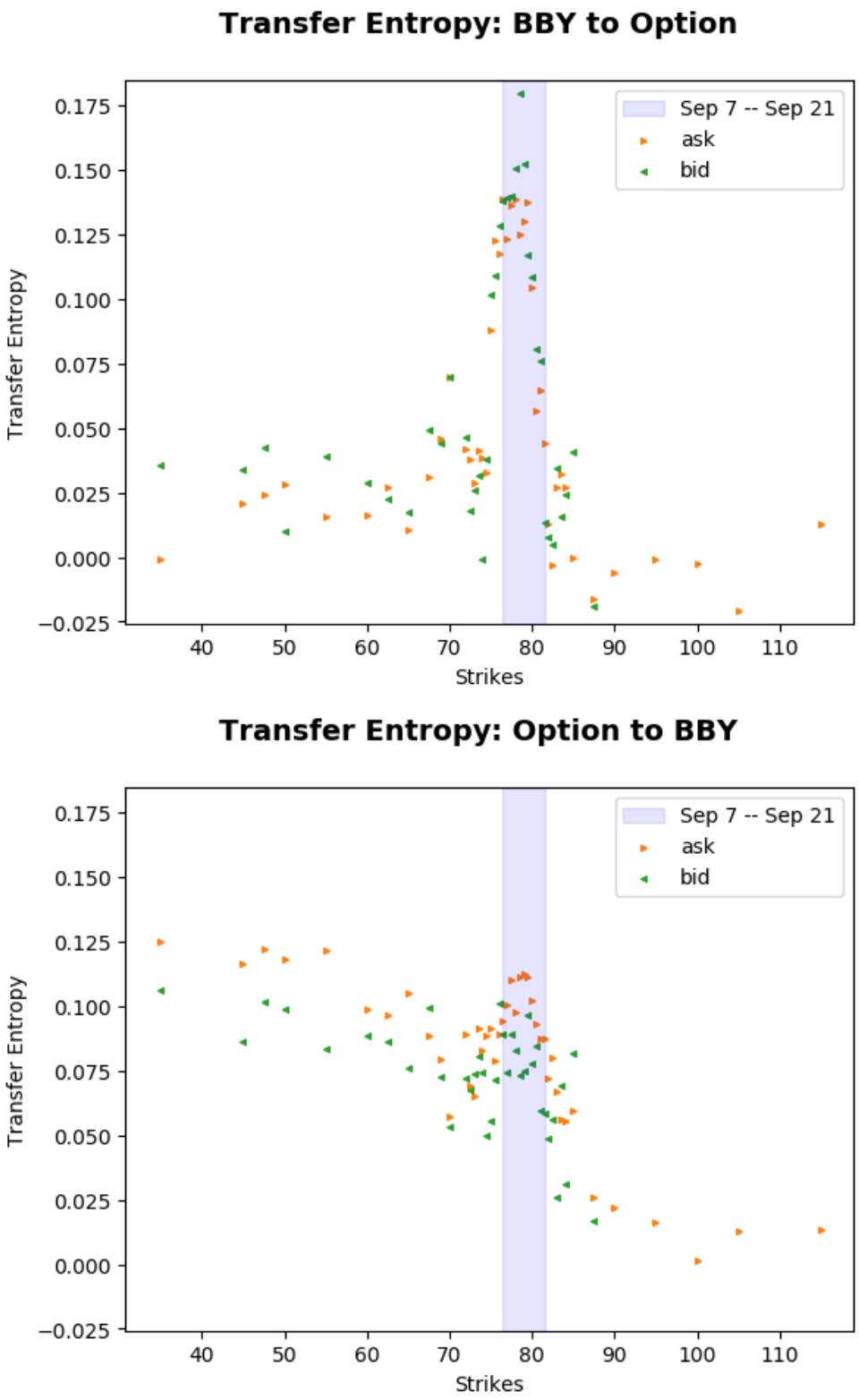

Figure 11 Transfer entropy between BBY and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

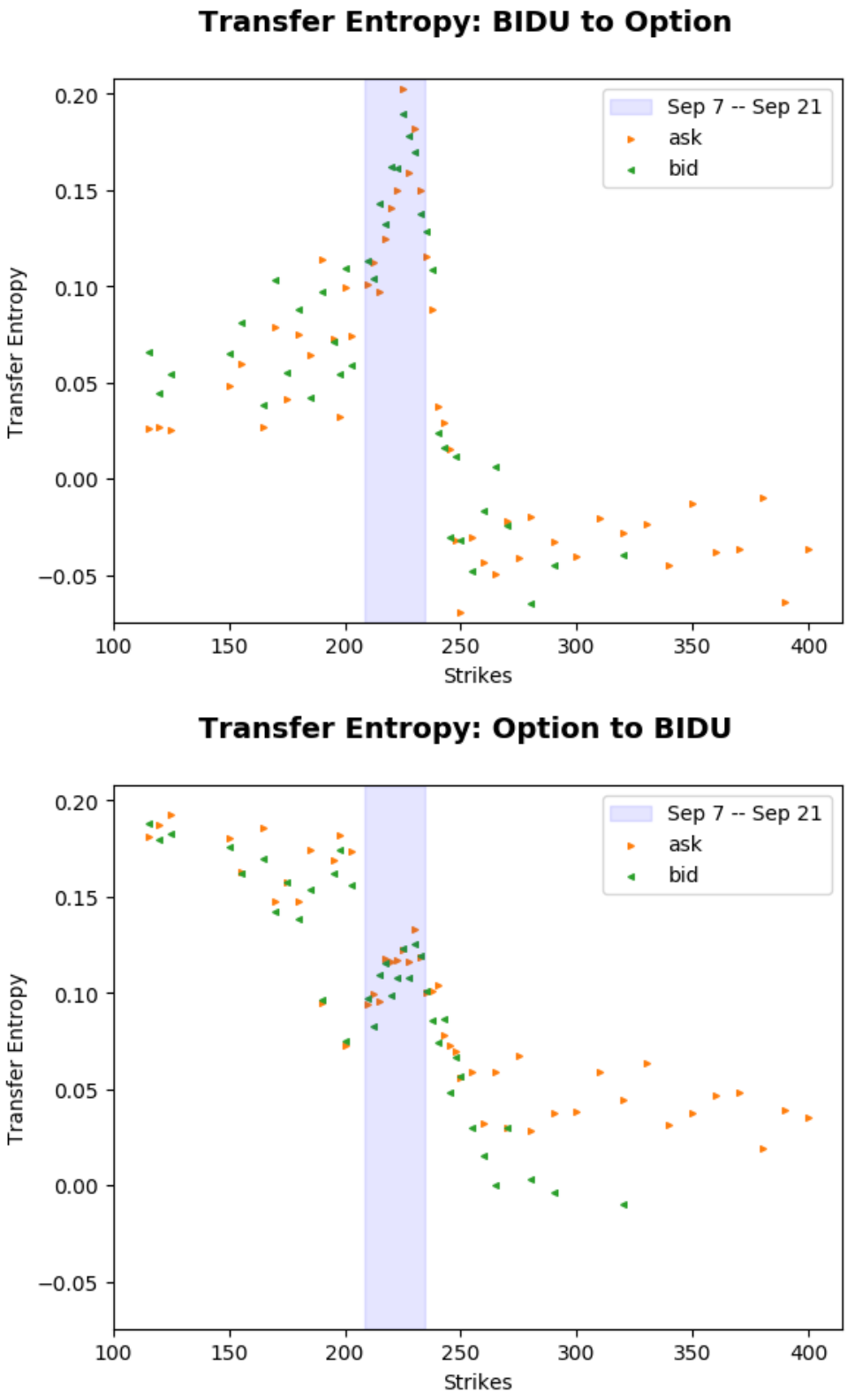

Figure 12 Transfer entropy between BIDU and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

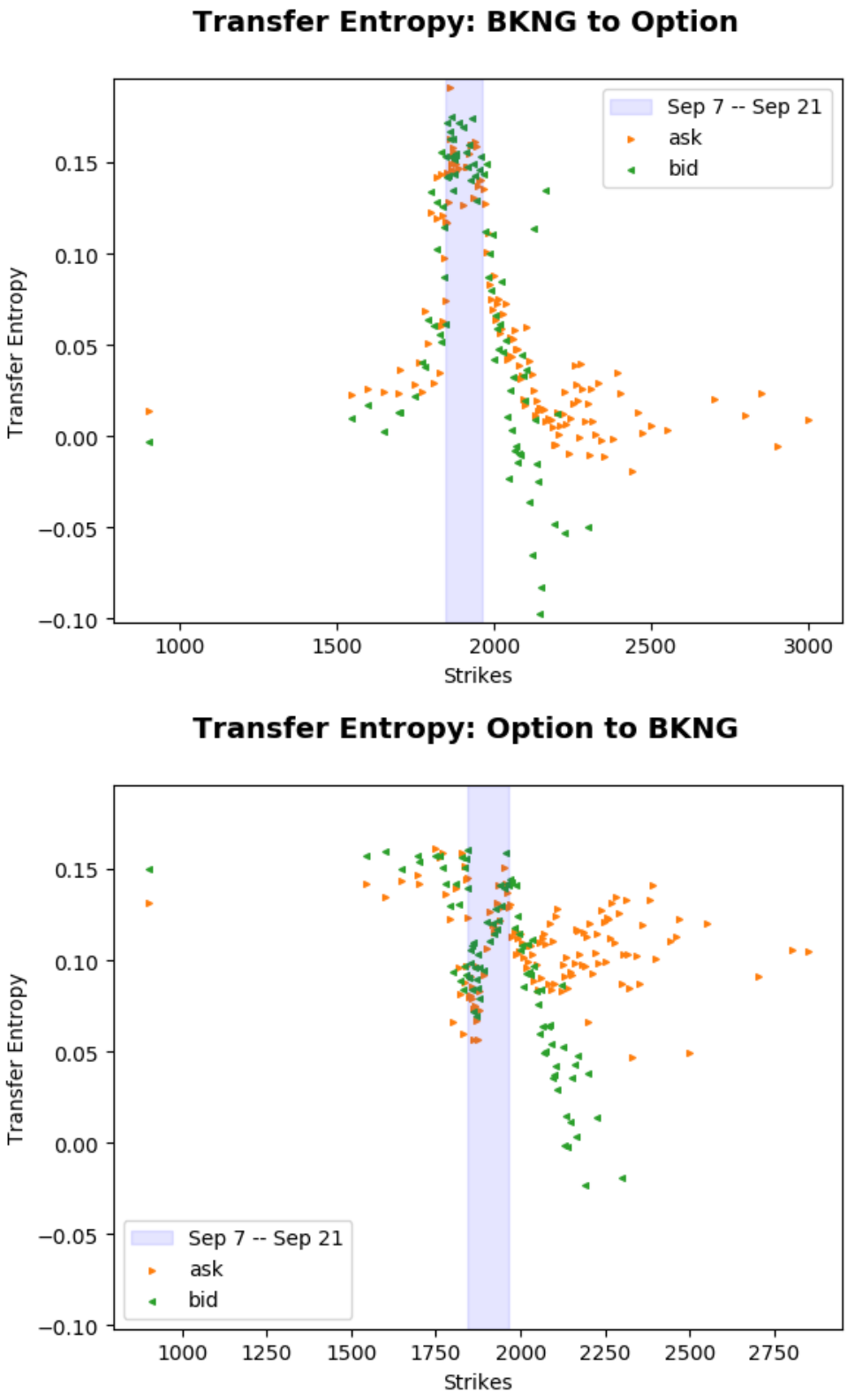

Figure 13 Transfer entropy between BKNG and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

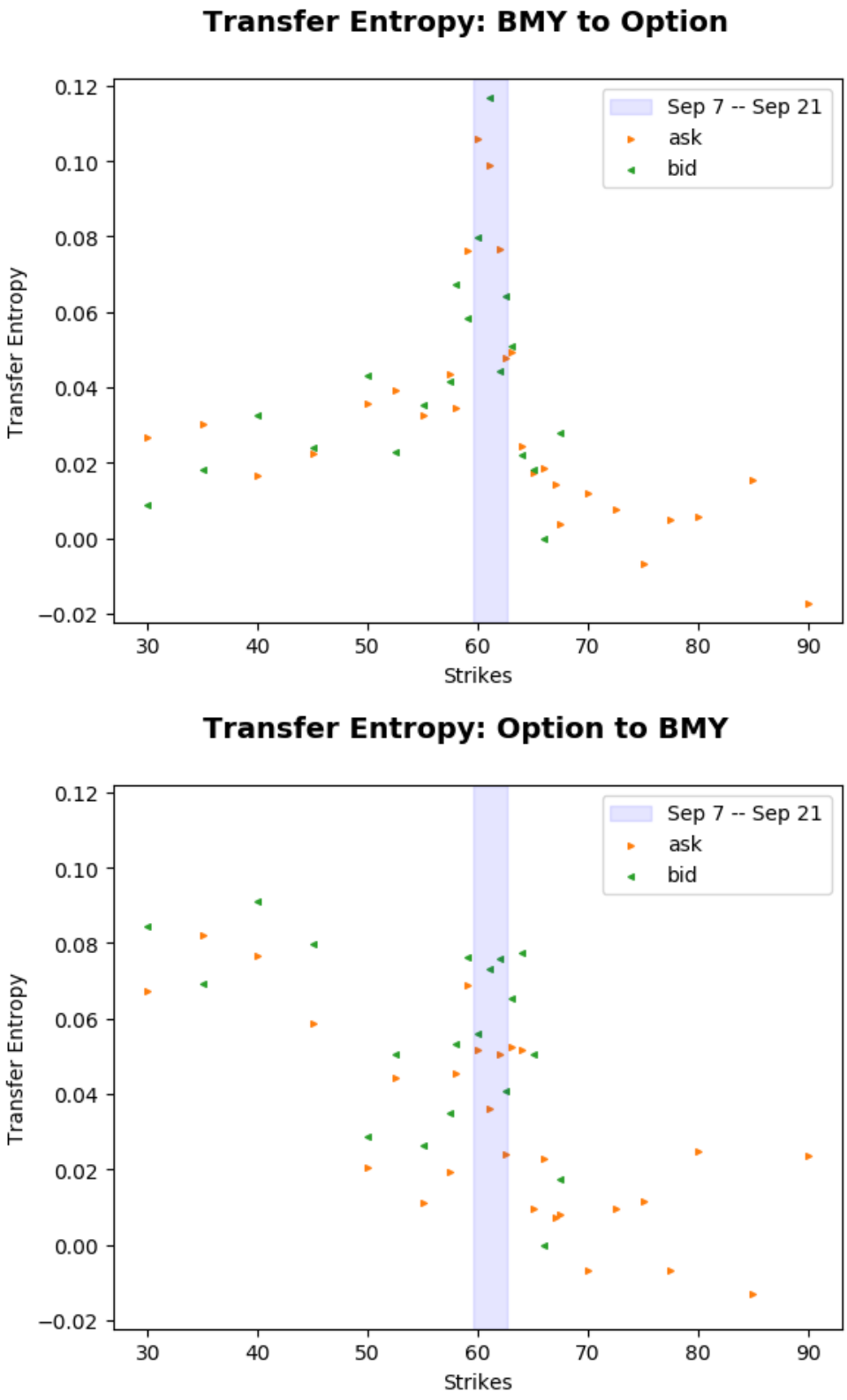

Figure 14 Transfer entropy between BMY and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

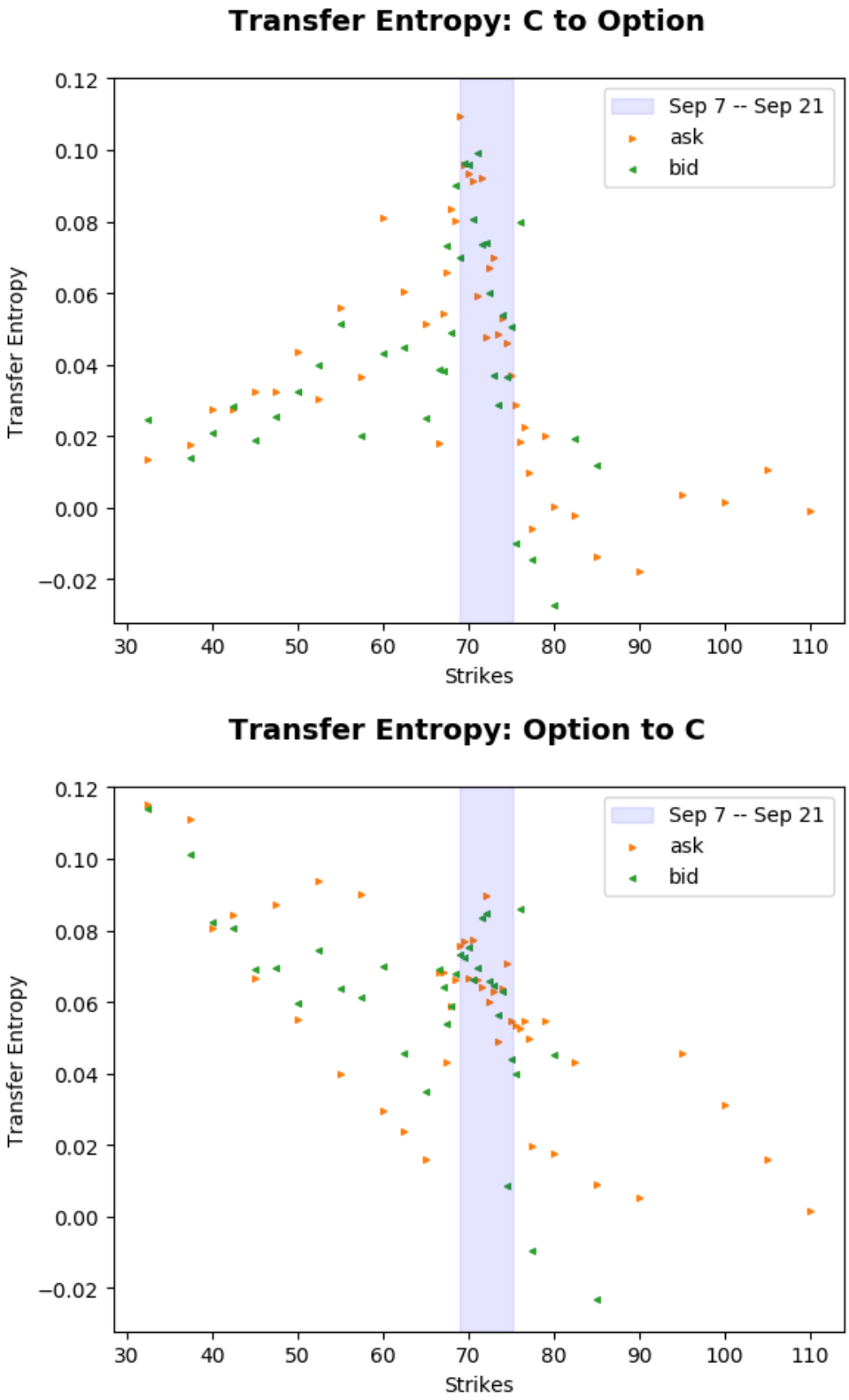

Figure 15 Transfer entropy between C and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 10 seconds. Option to underlying timescale: 12 seconds.

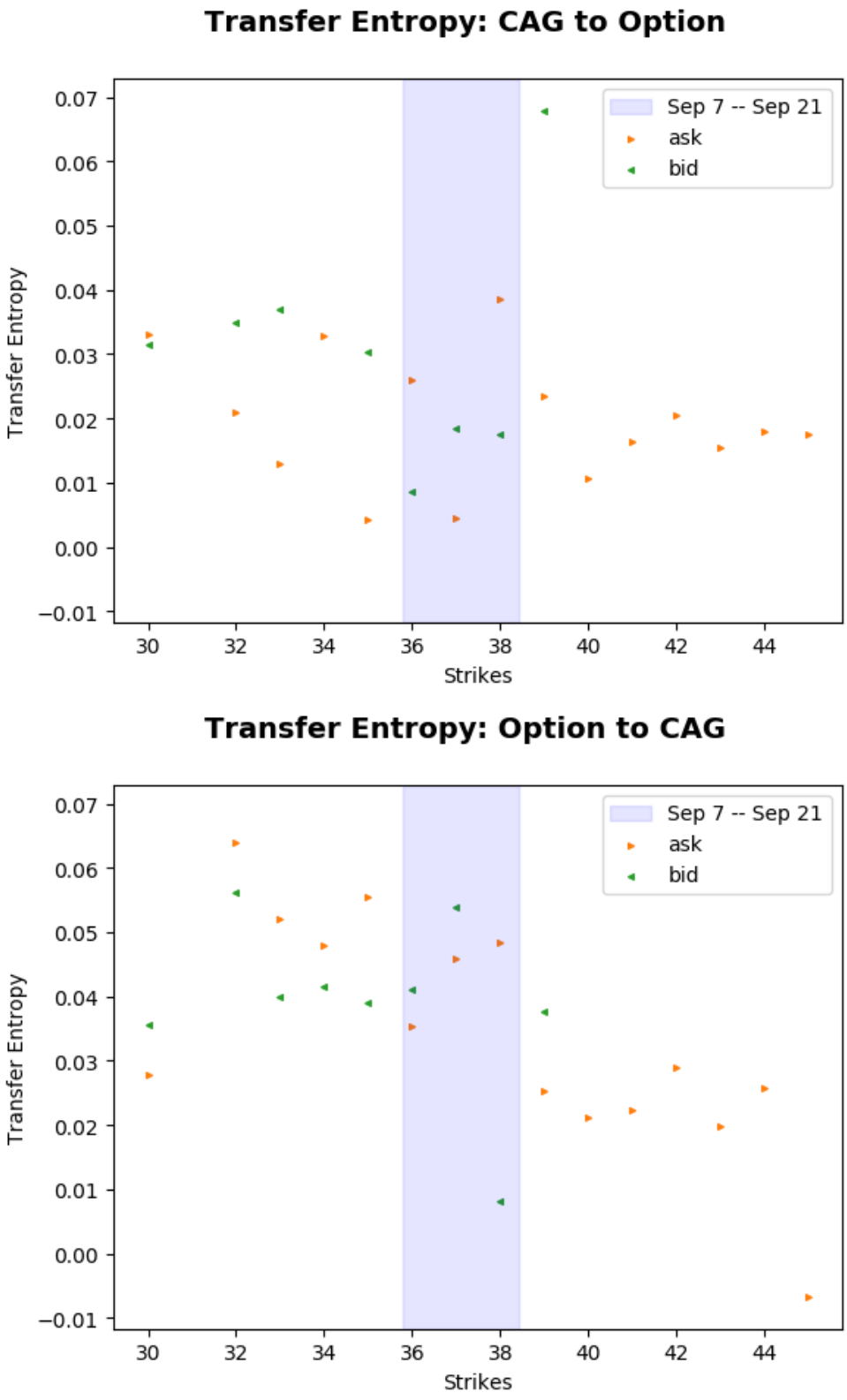

Figure 16 Transfer entropy between CAG and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

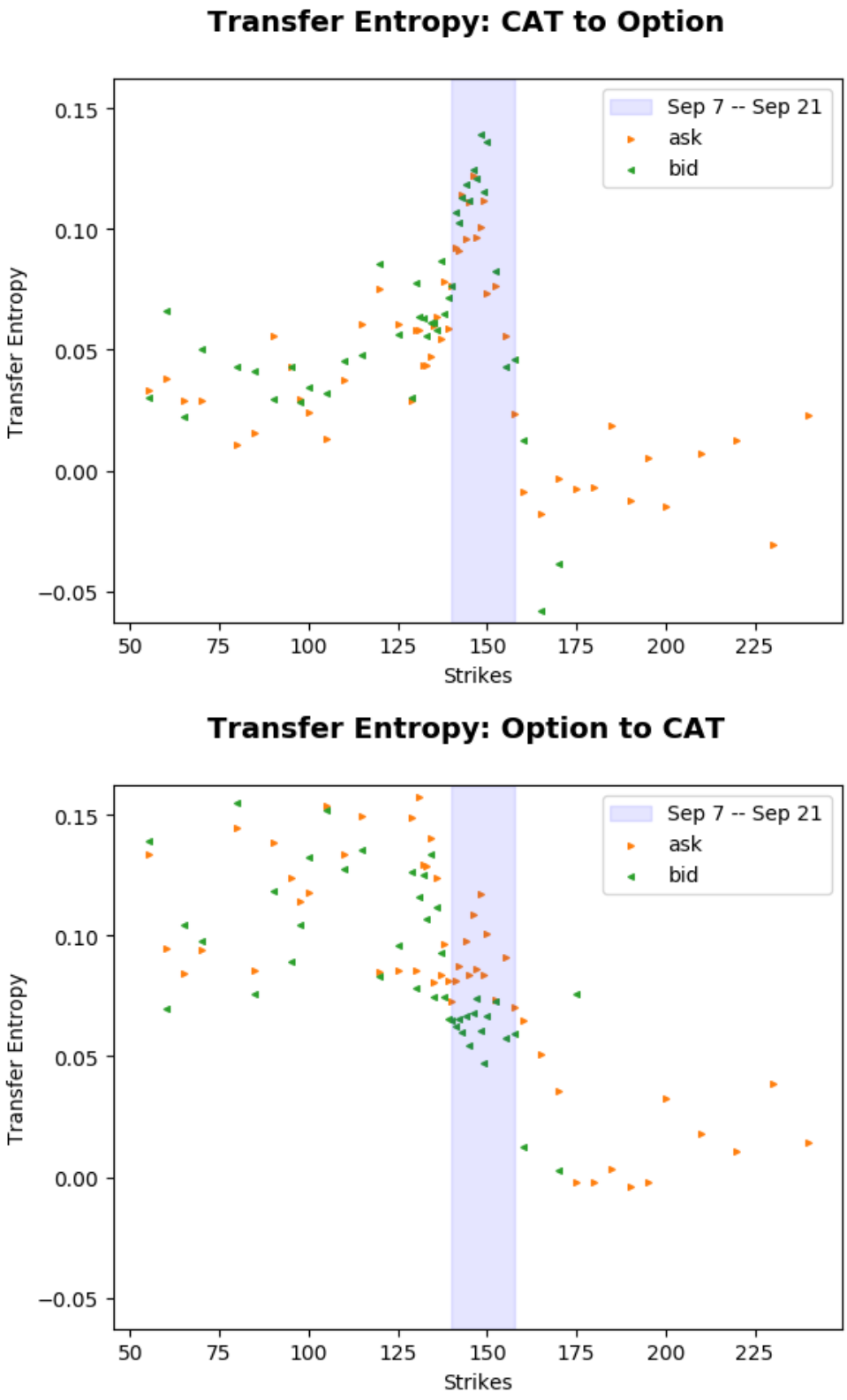

Figure 17 Transfer entropy between CAT and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

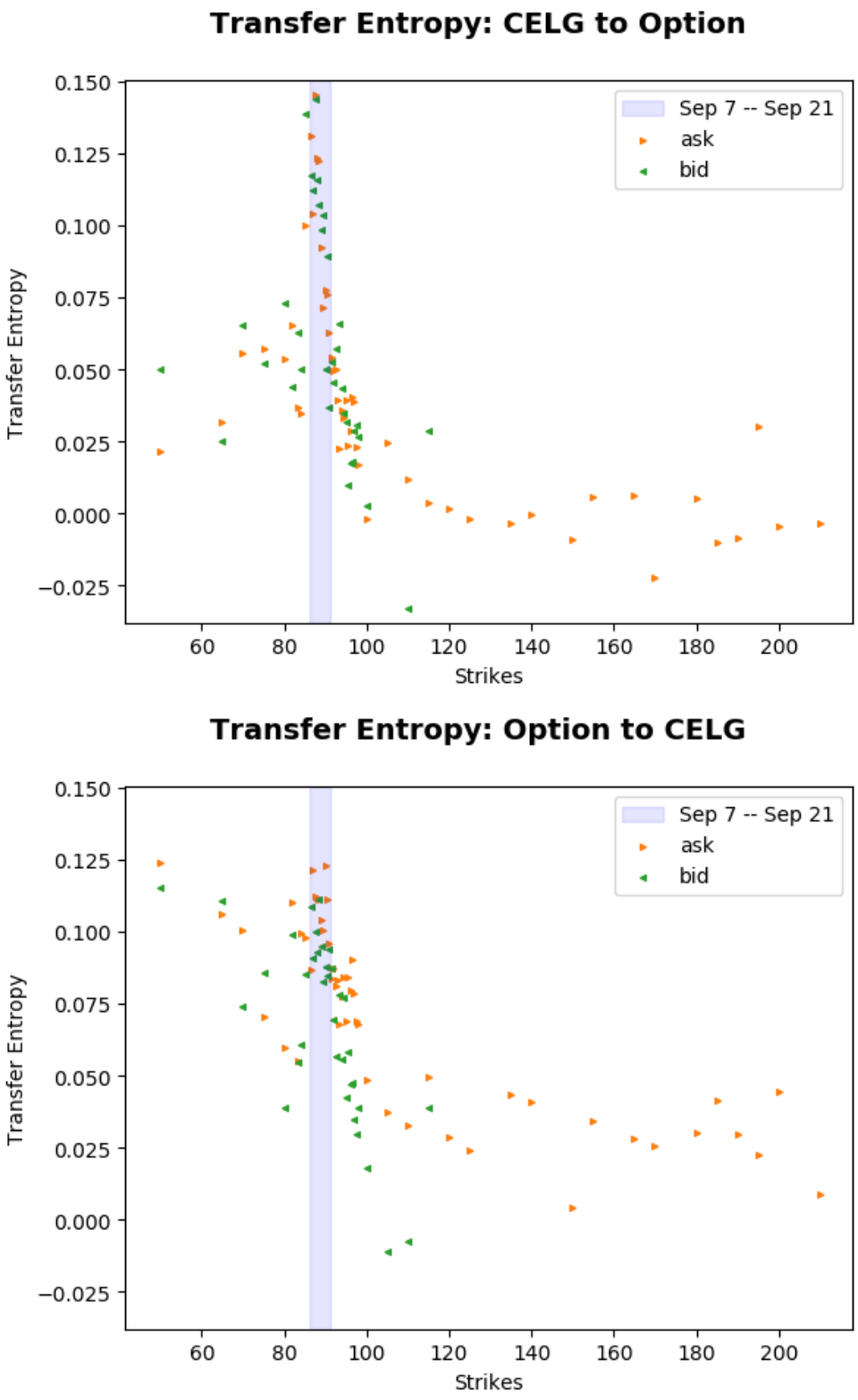

Figure 18 Transfer entropy between CELG and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

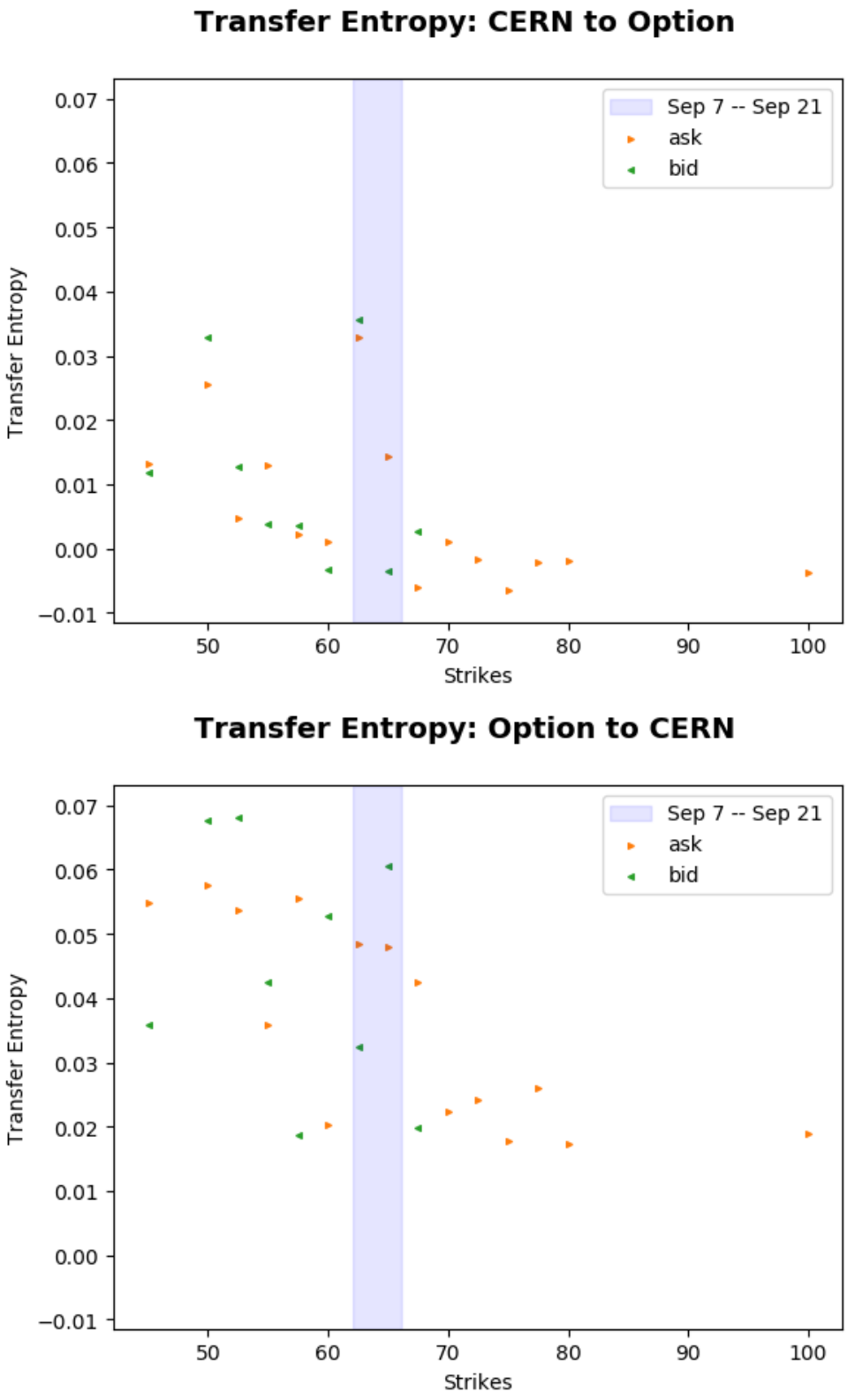

Figure 19 Transfer entropy between CERN and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 10 seconds.

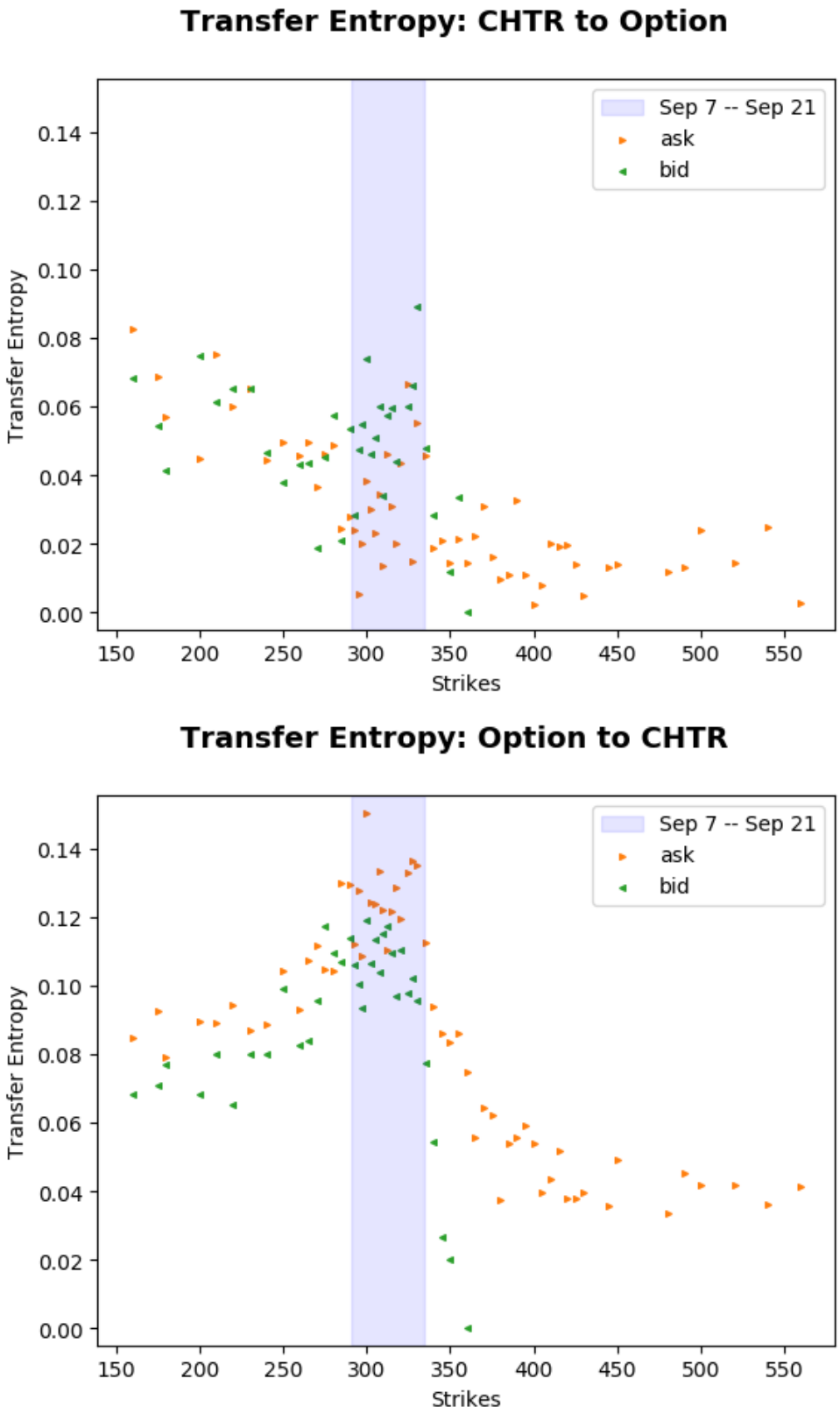

Figure 20 Transfer entropy between CHTR and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 10 seconds.

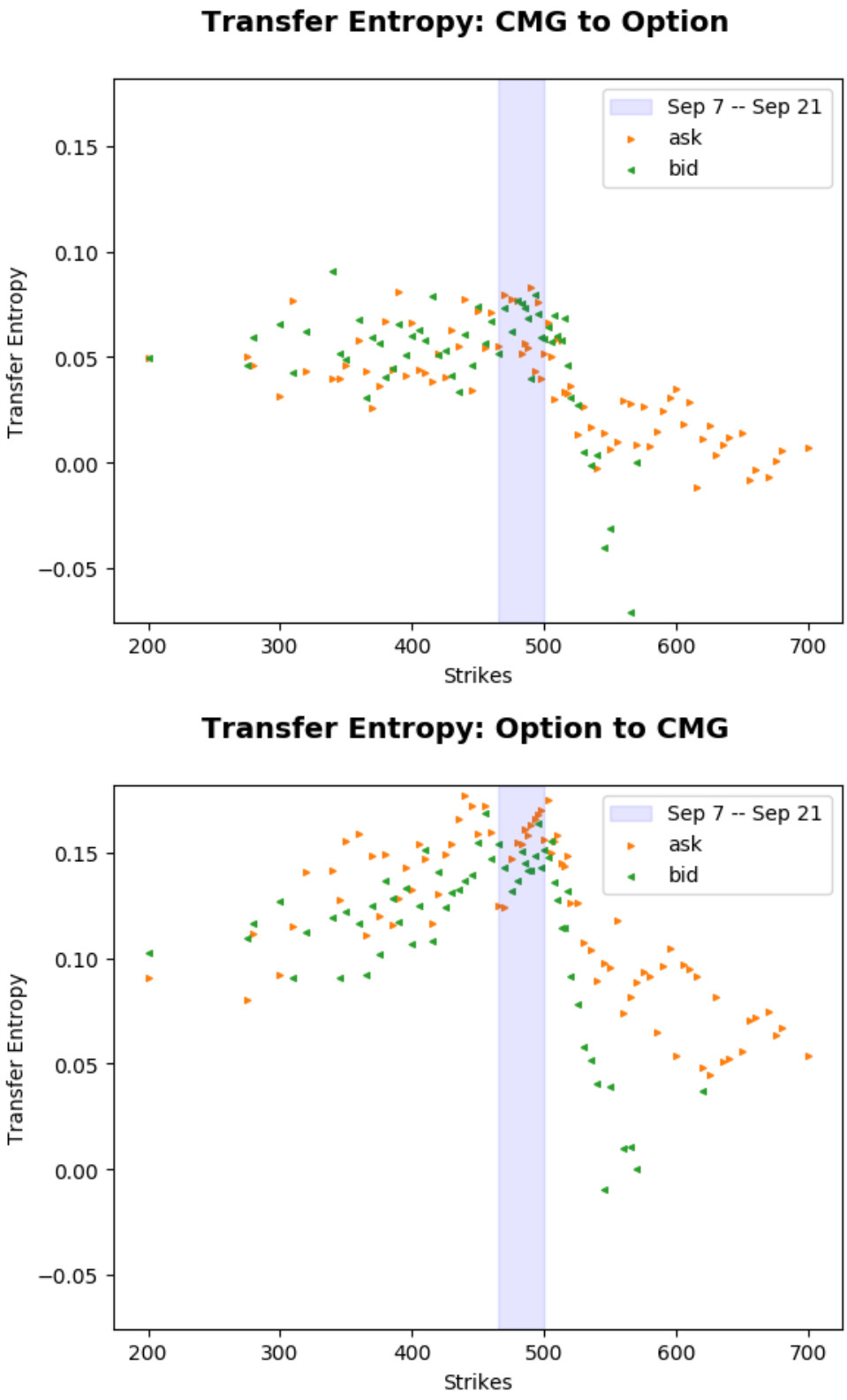

Figure 21 Transfer entropy between CMG and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

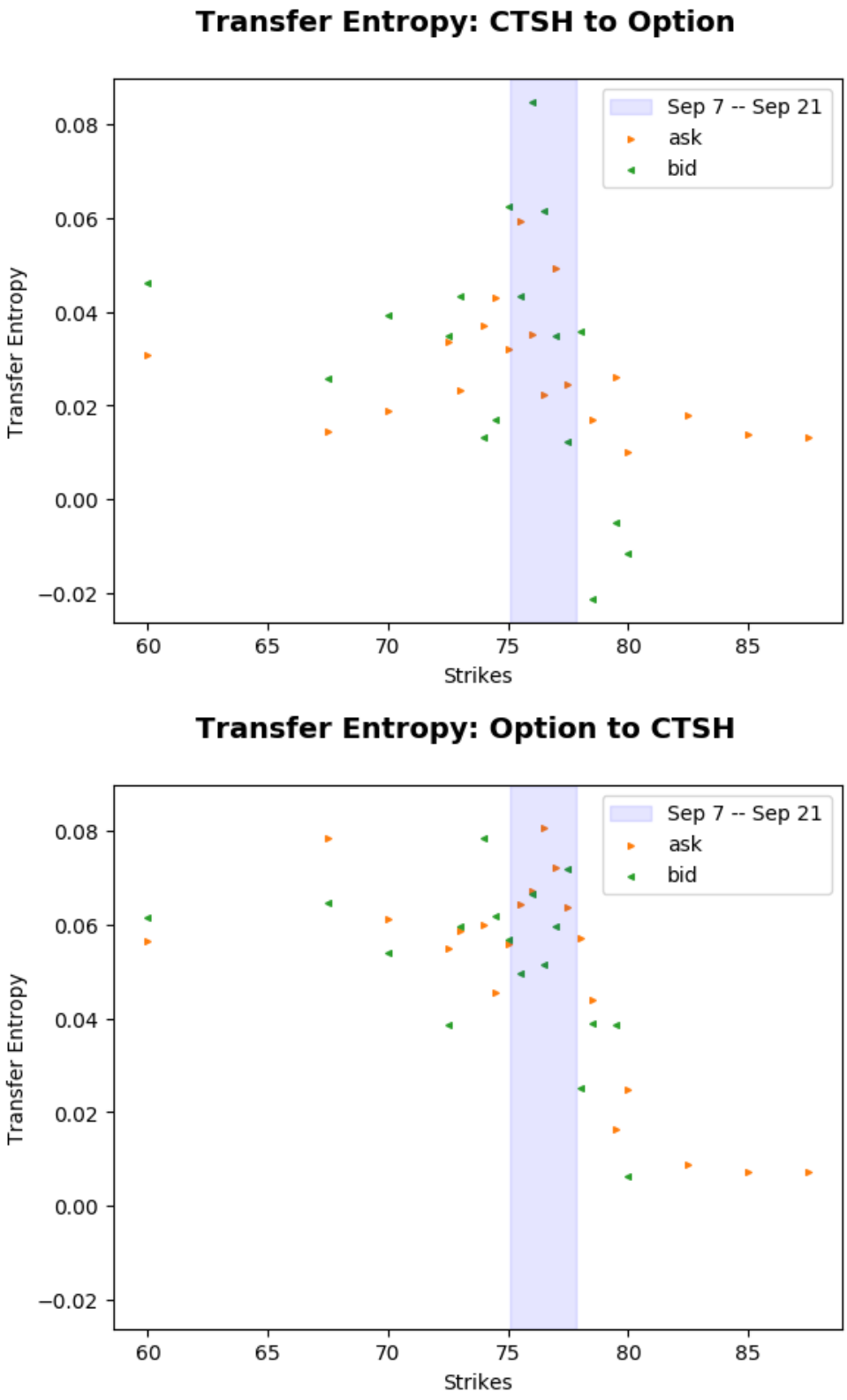

Figure 22 Transfer entropy between CTSH and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

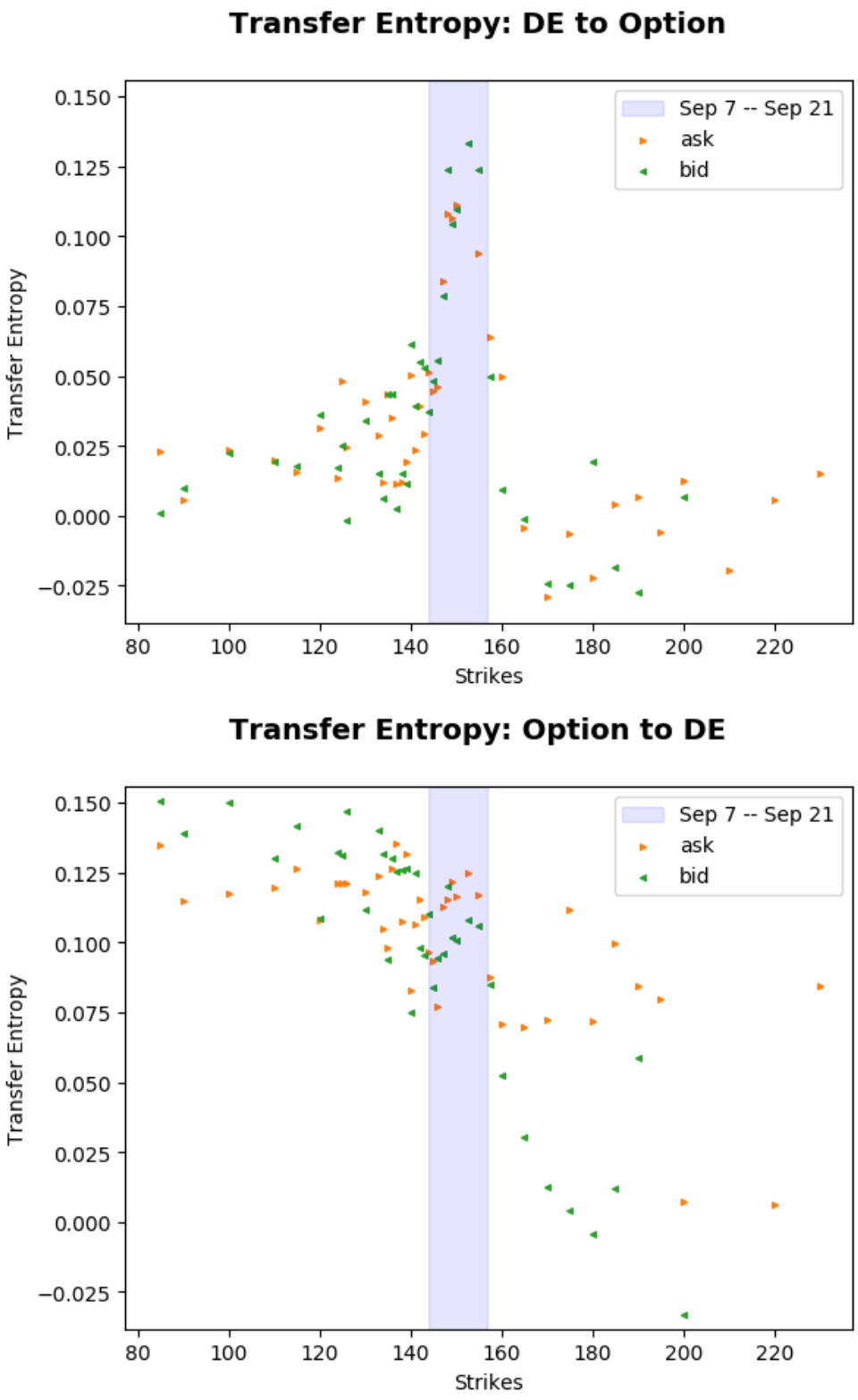

Figure 23 Transfer entropy between DE and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 10 seconds. Option to underlying timescale: 12 seconds.

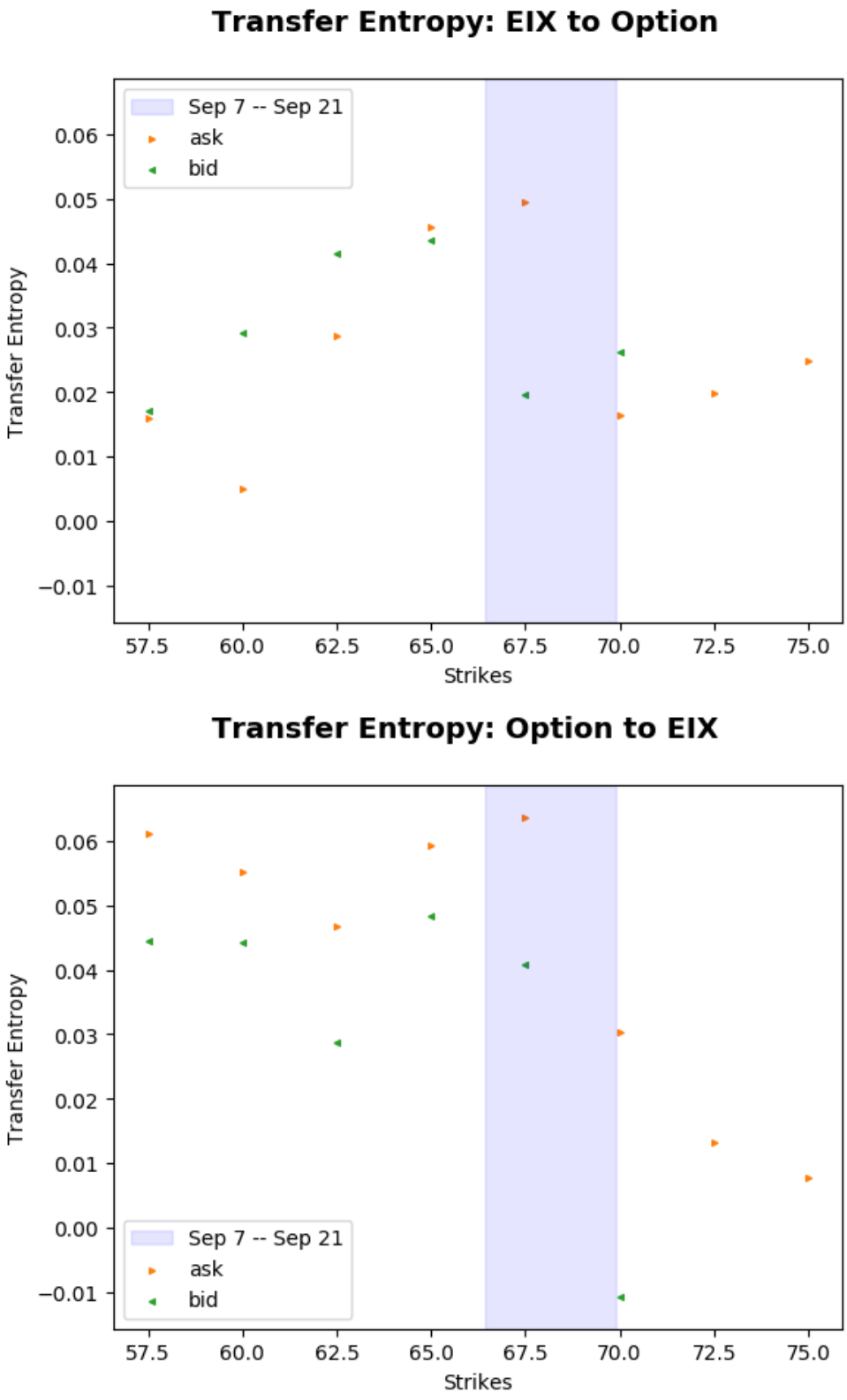

Figure 24 Transfer entropy between EIX and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

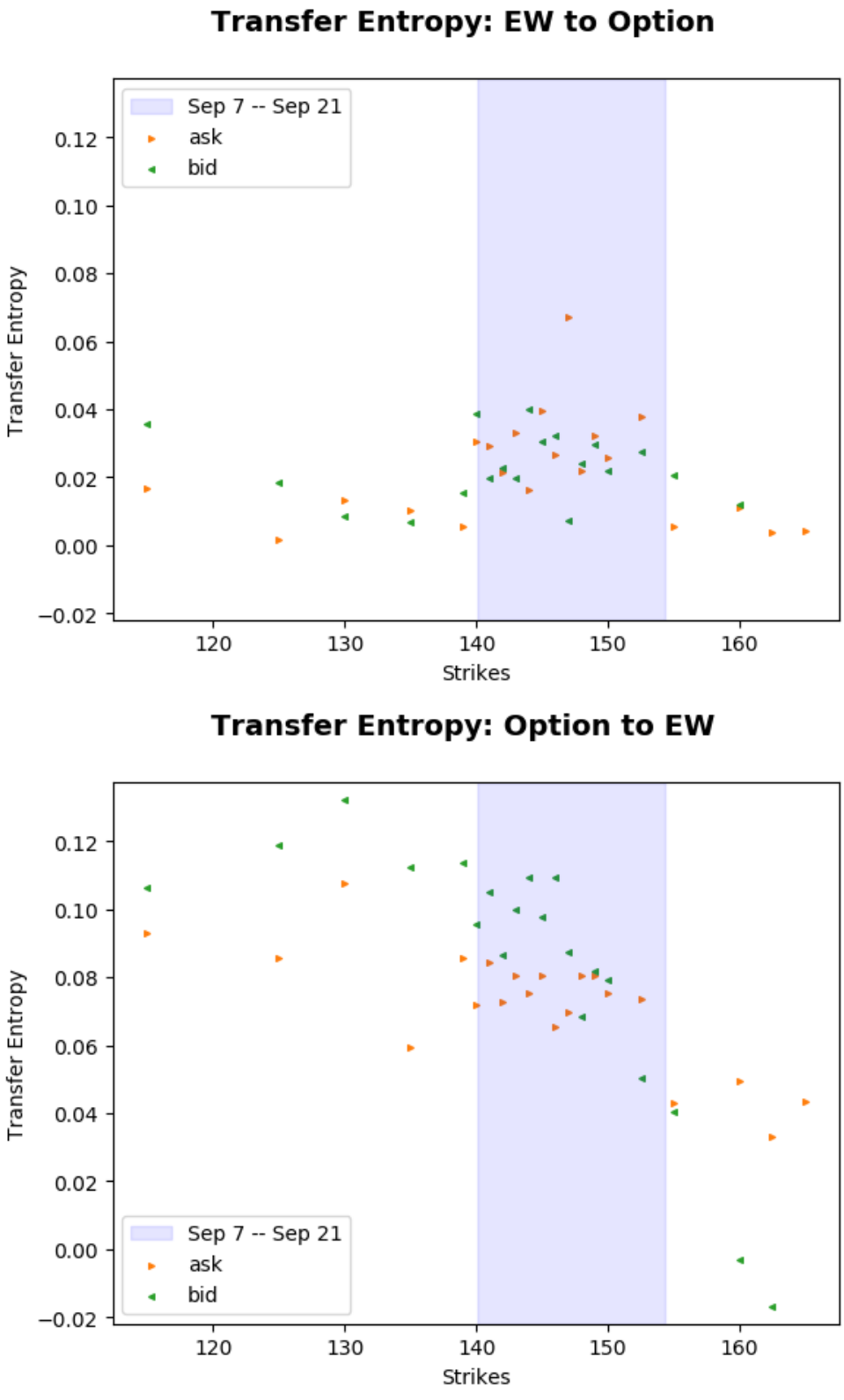

Figure 25 Transfer entropy between EW and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

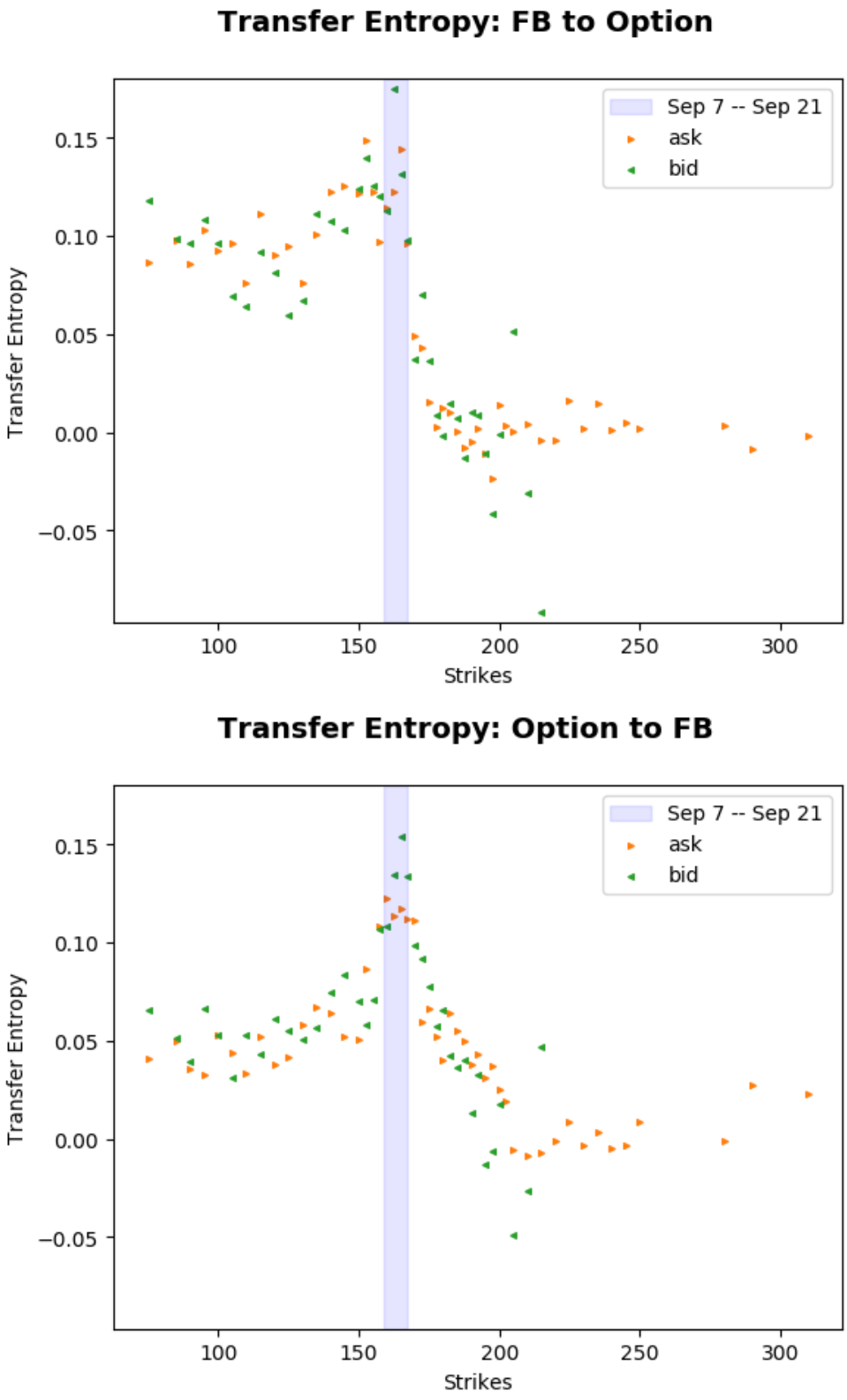

Figure 26 Transfer entropy between FB and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

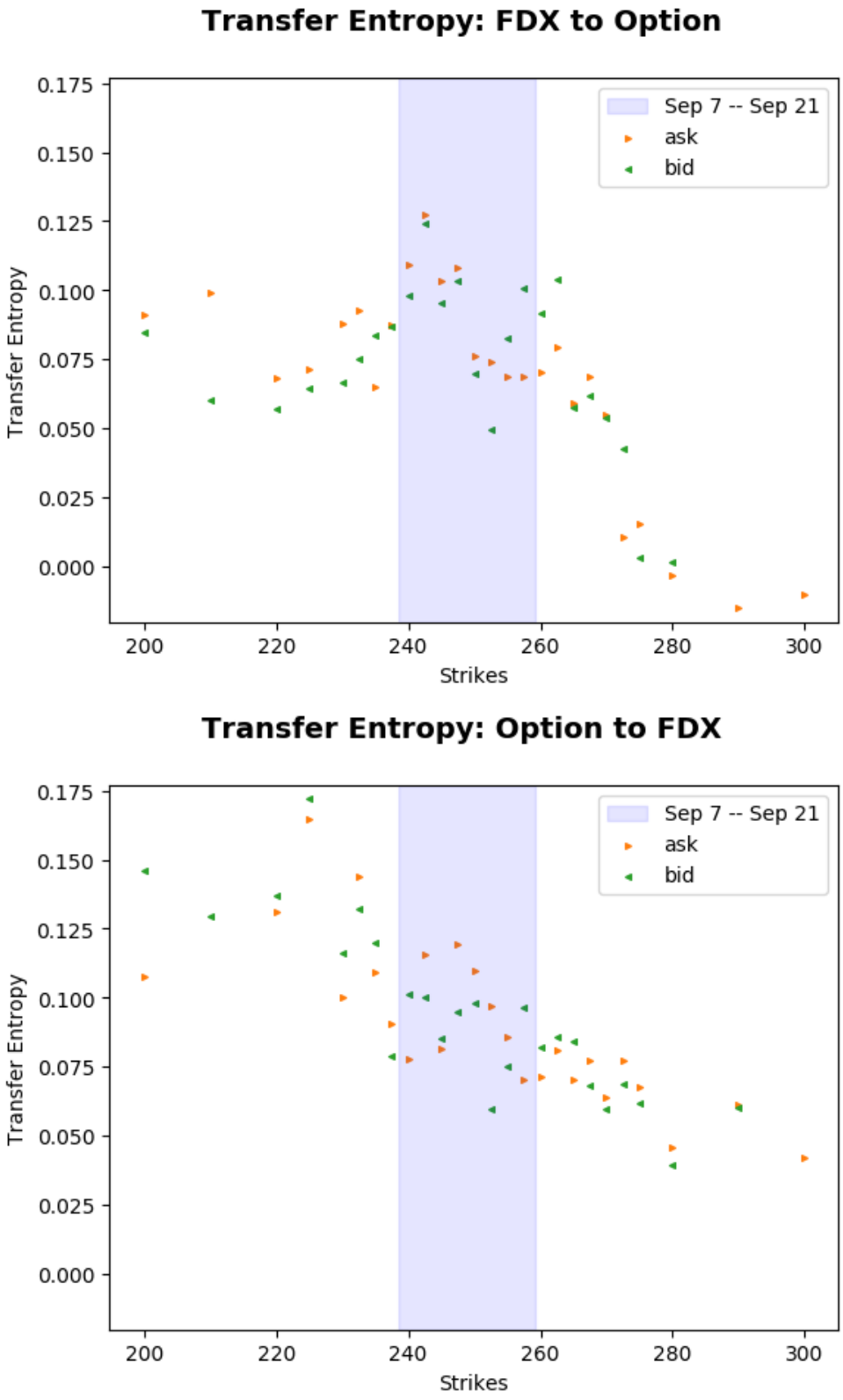

Figure 27 Transfer entropy between FDX and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

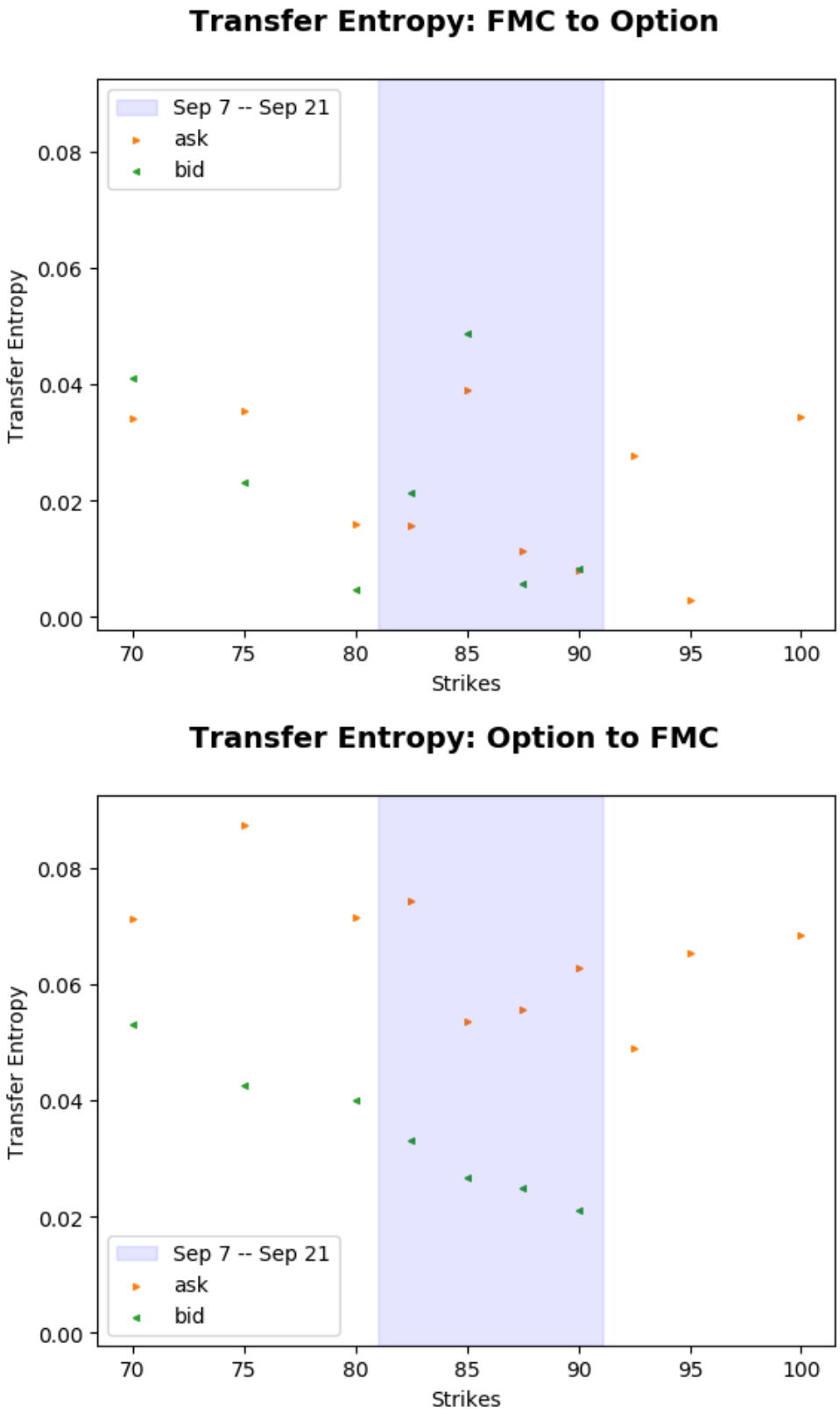

Figure 28 Transfer entropy between FMC and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 10 seconds.

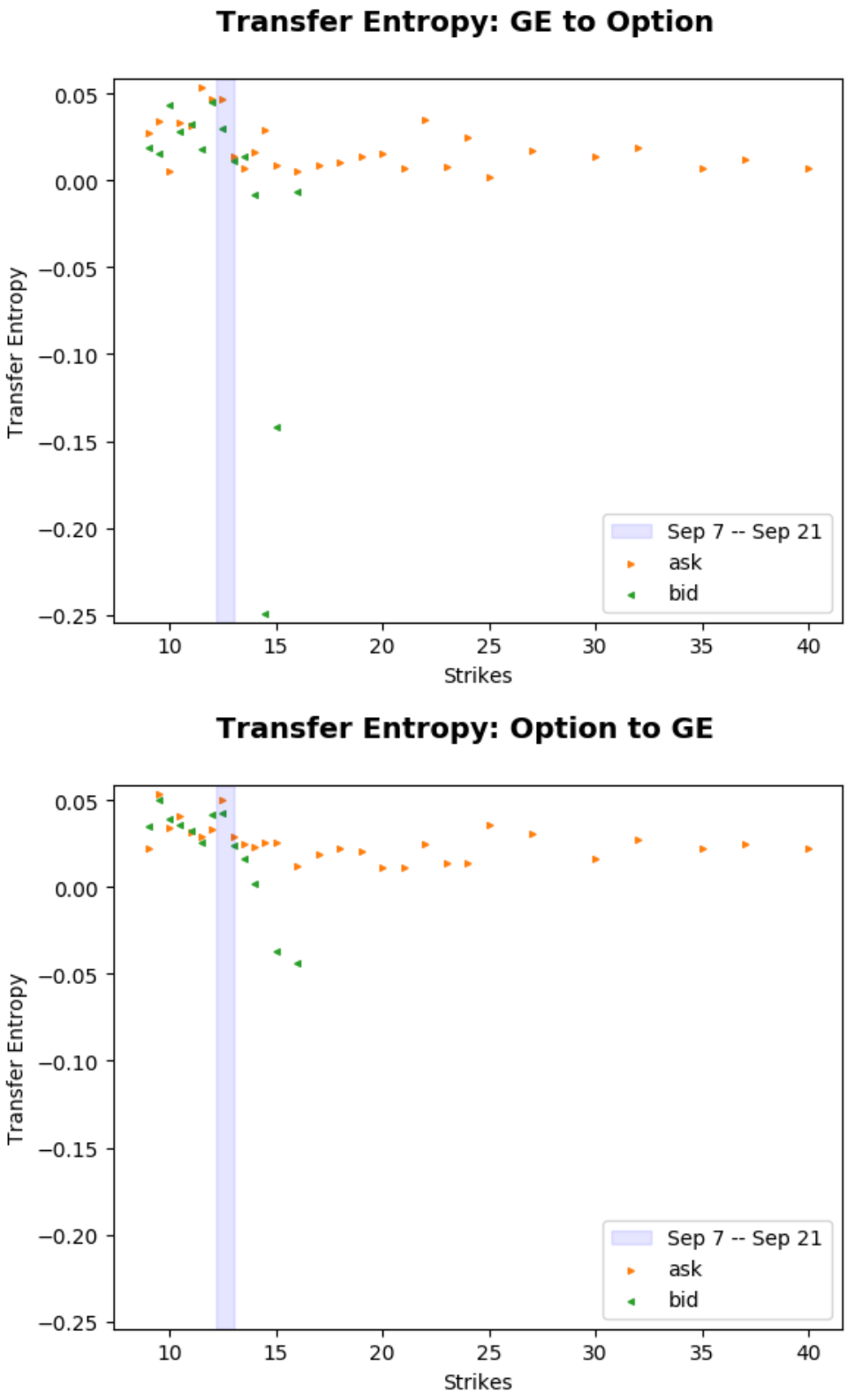

Figure 29 Transfer entropy between GE and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 10 seconds.

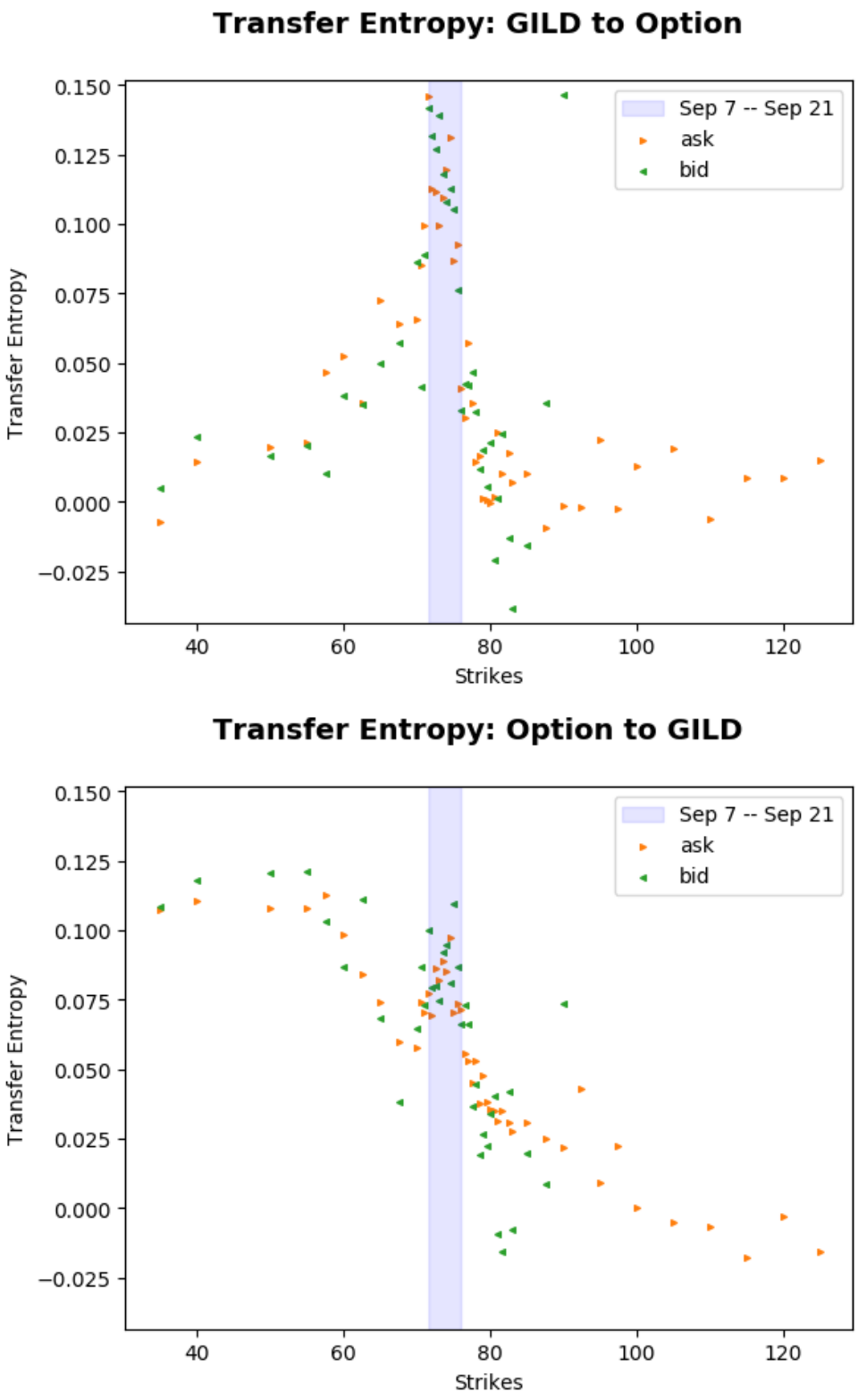

Figure 30 Transfer entropy between GILD and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

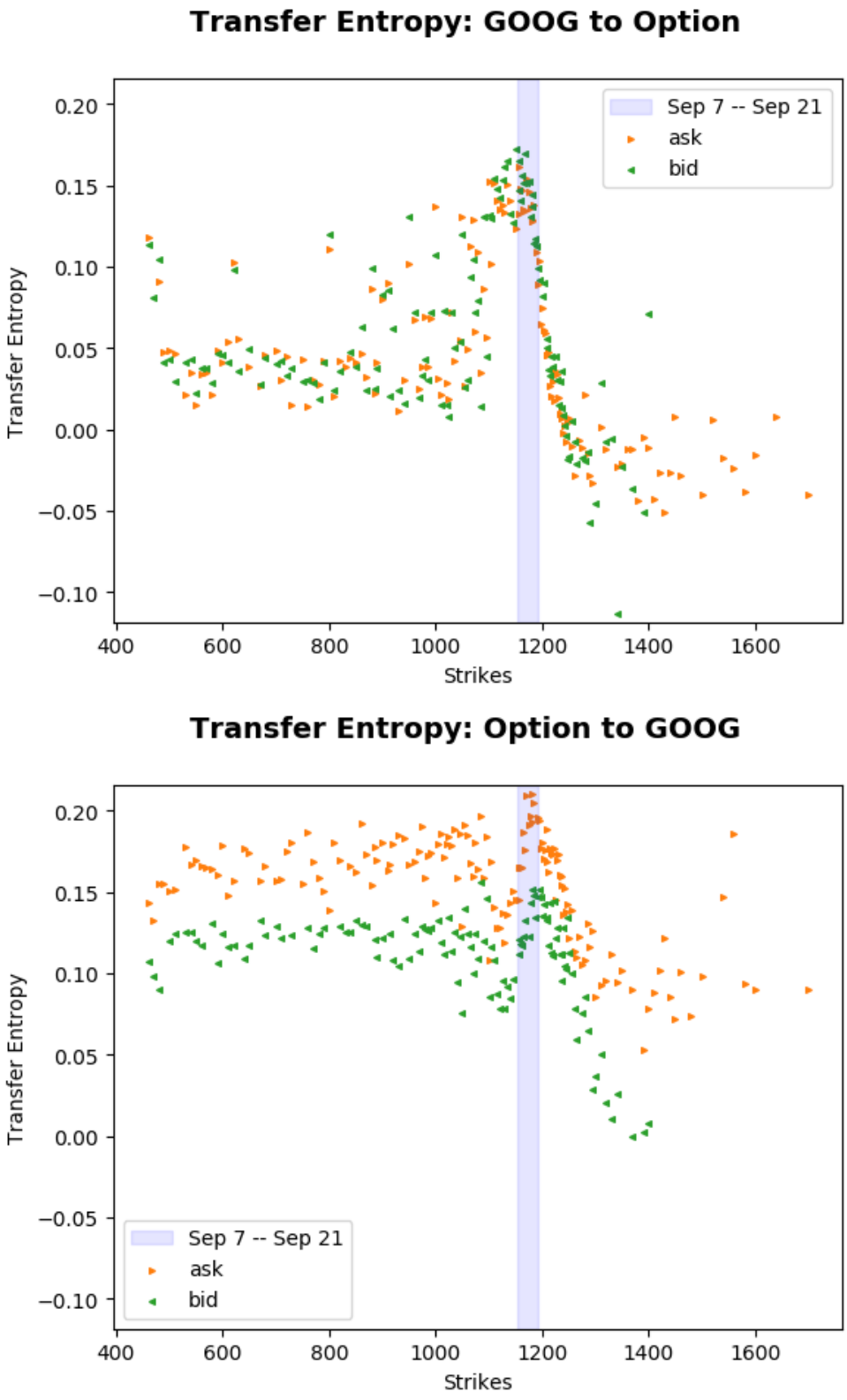

Figure 31 Transfer entropy between GOOG and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 10 seconds.

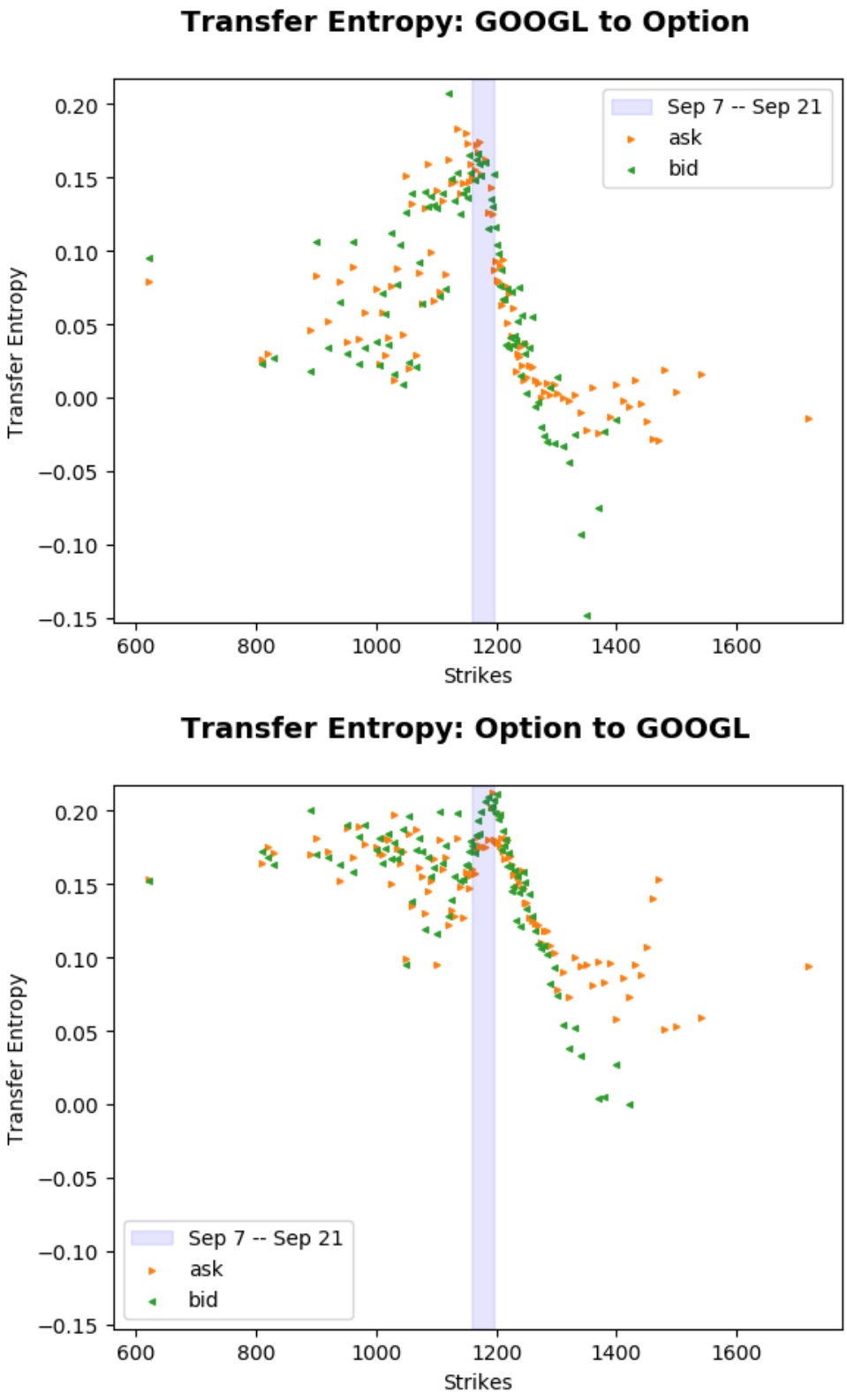

Figure 32 Transfer entropy between GOOGL and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 8 seconds.

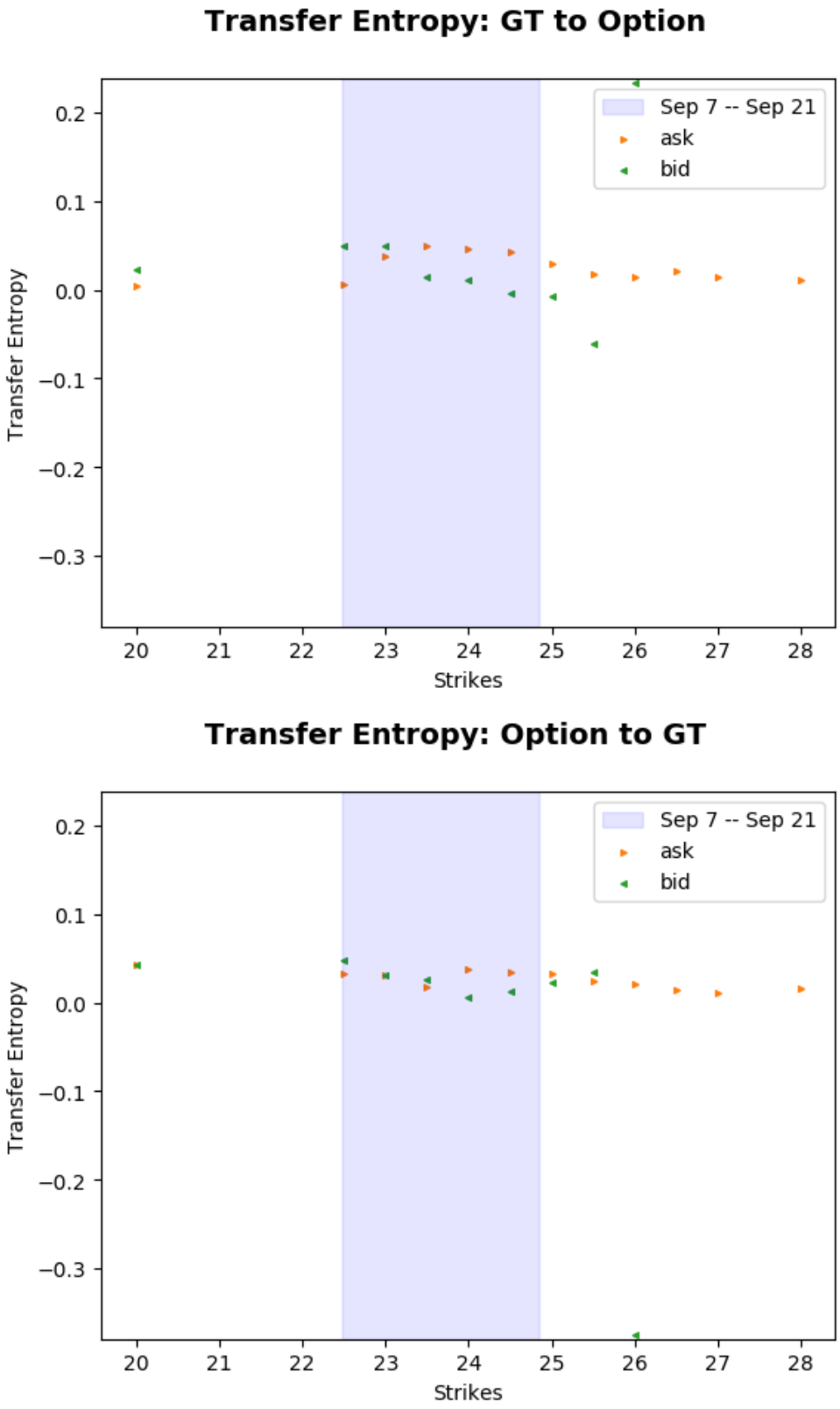

Figure 33 Transfer entropy between GT and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 10 seconds.

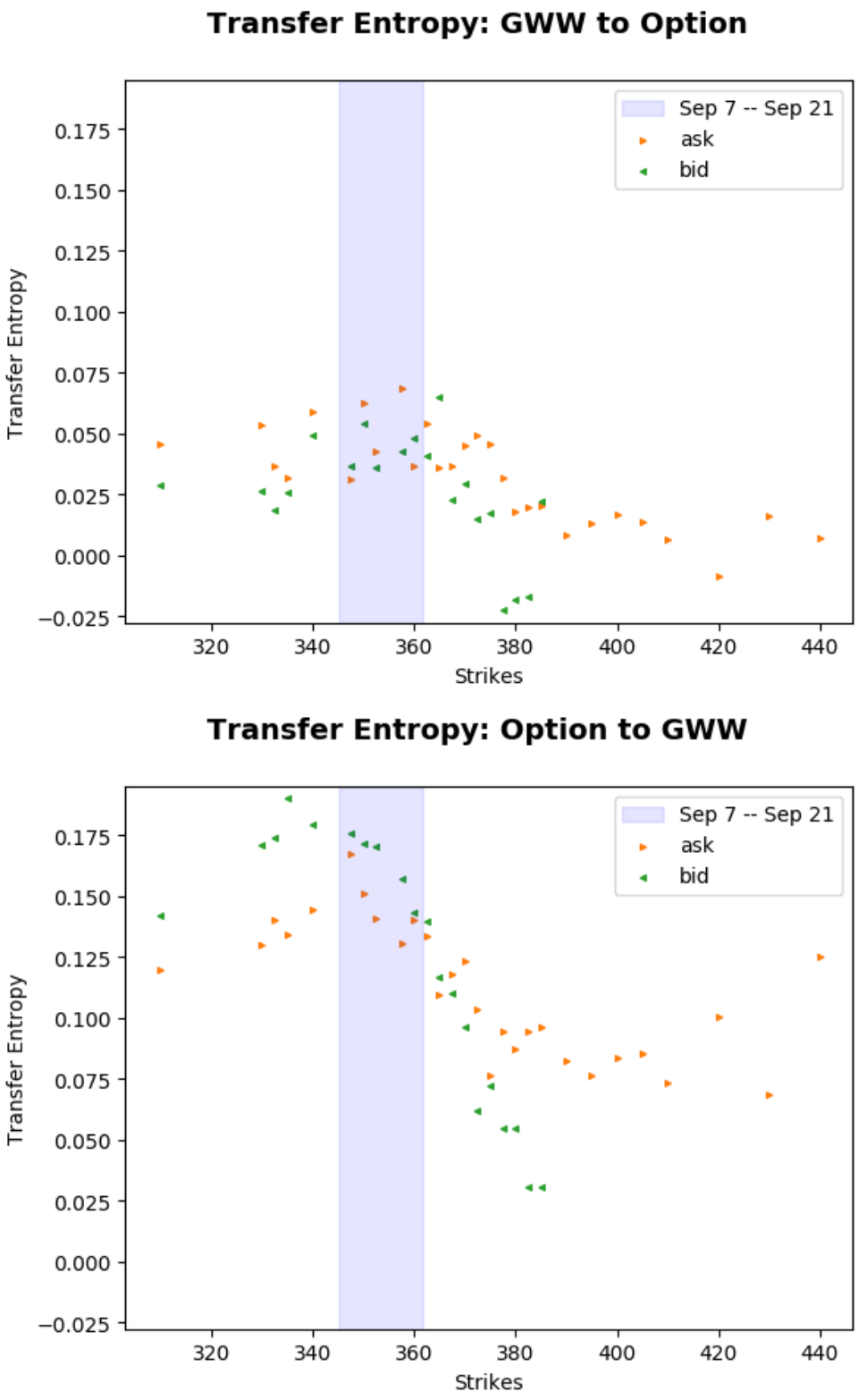

Figure 34 Transfer entropy between GWW and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

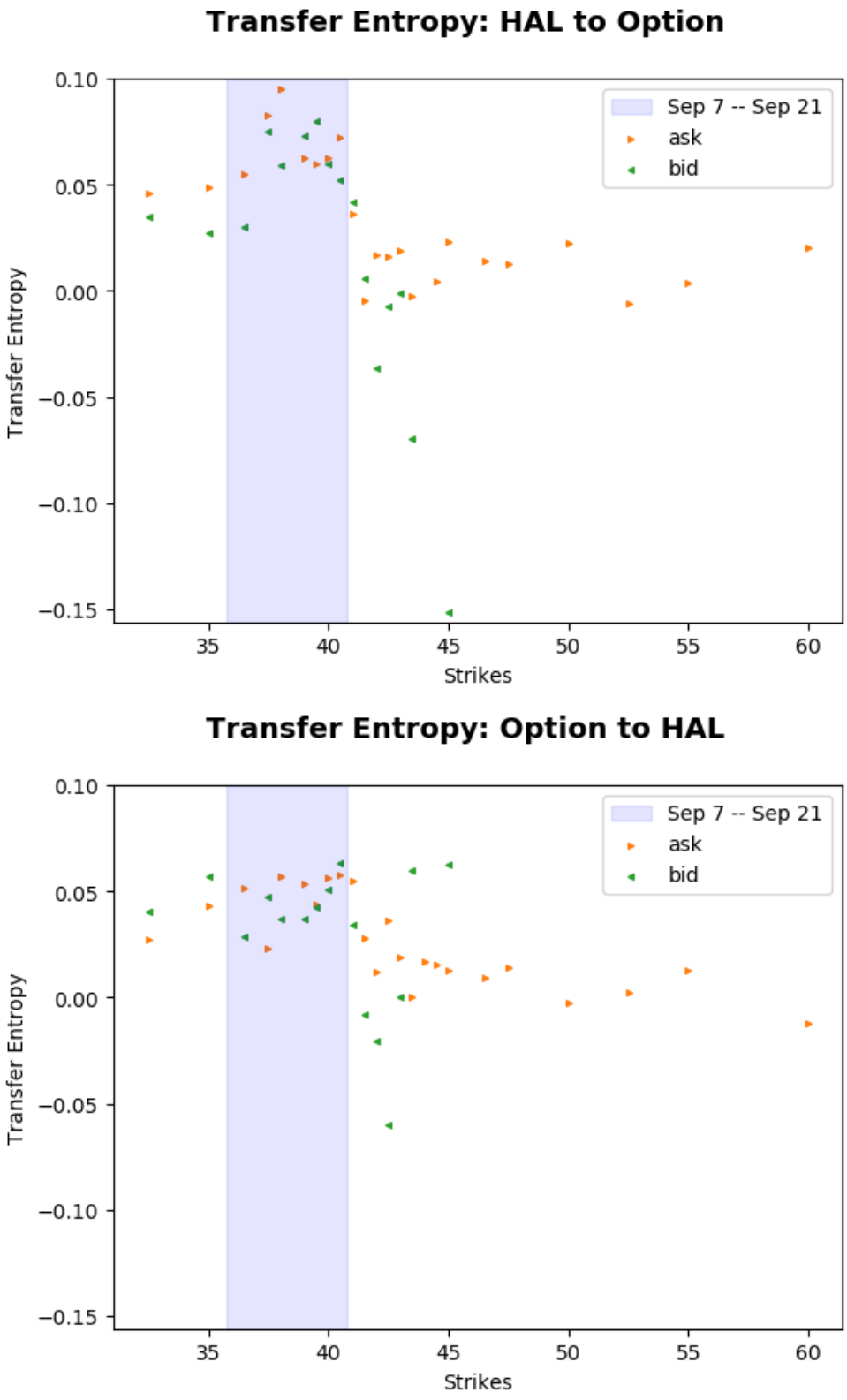

Figure 35 Transfer entropy between HAL and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

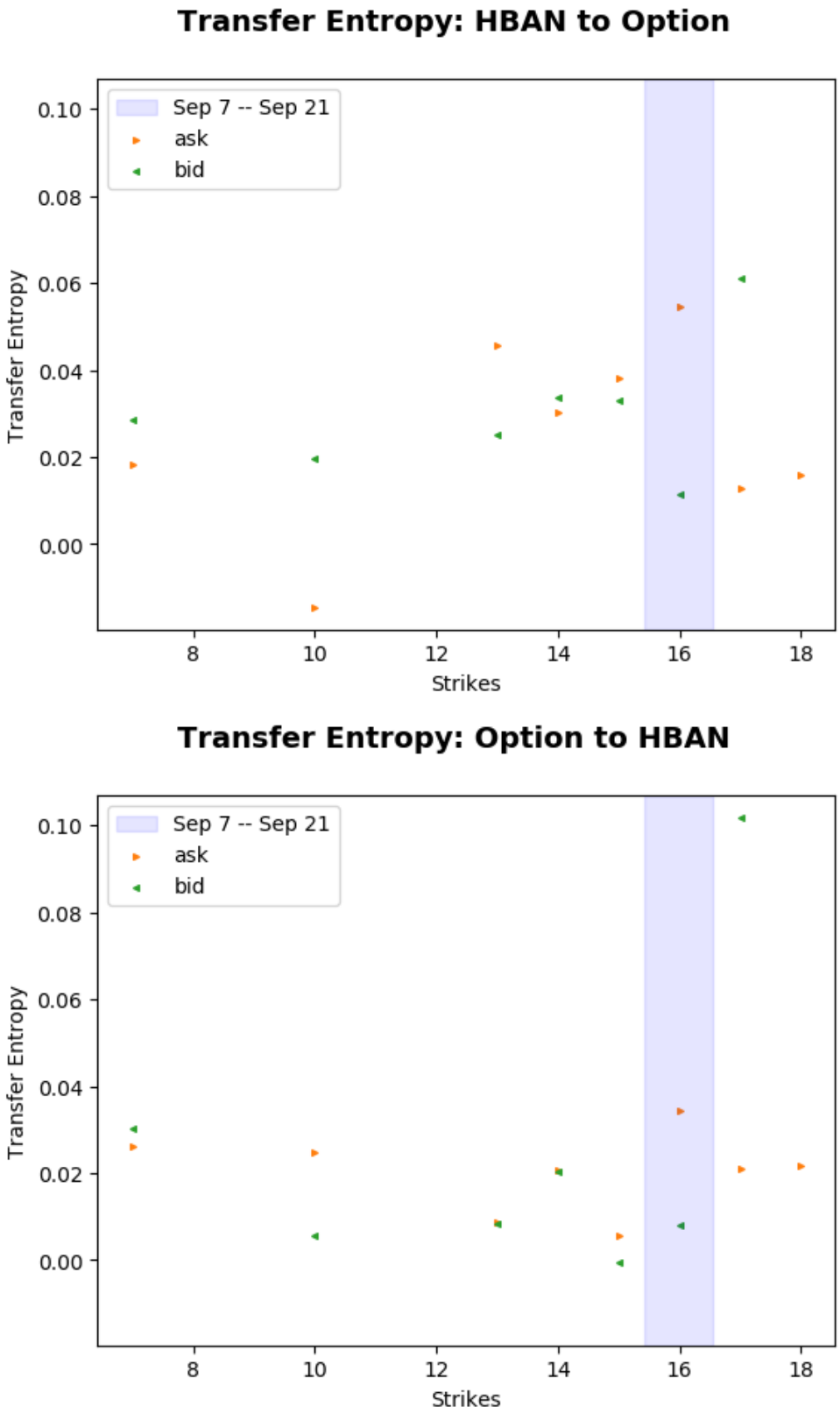

Figure 36 Transfer entropy between HBAN and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

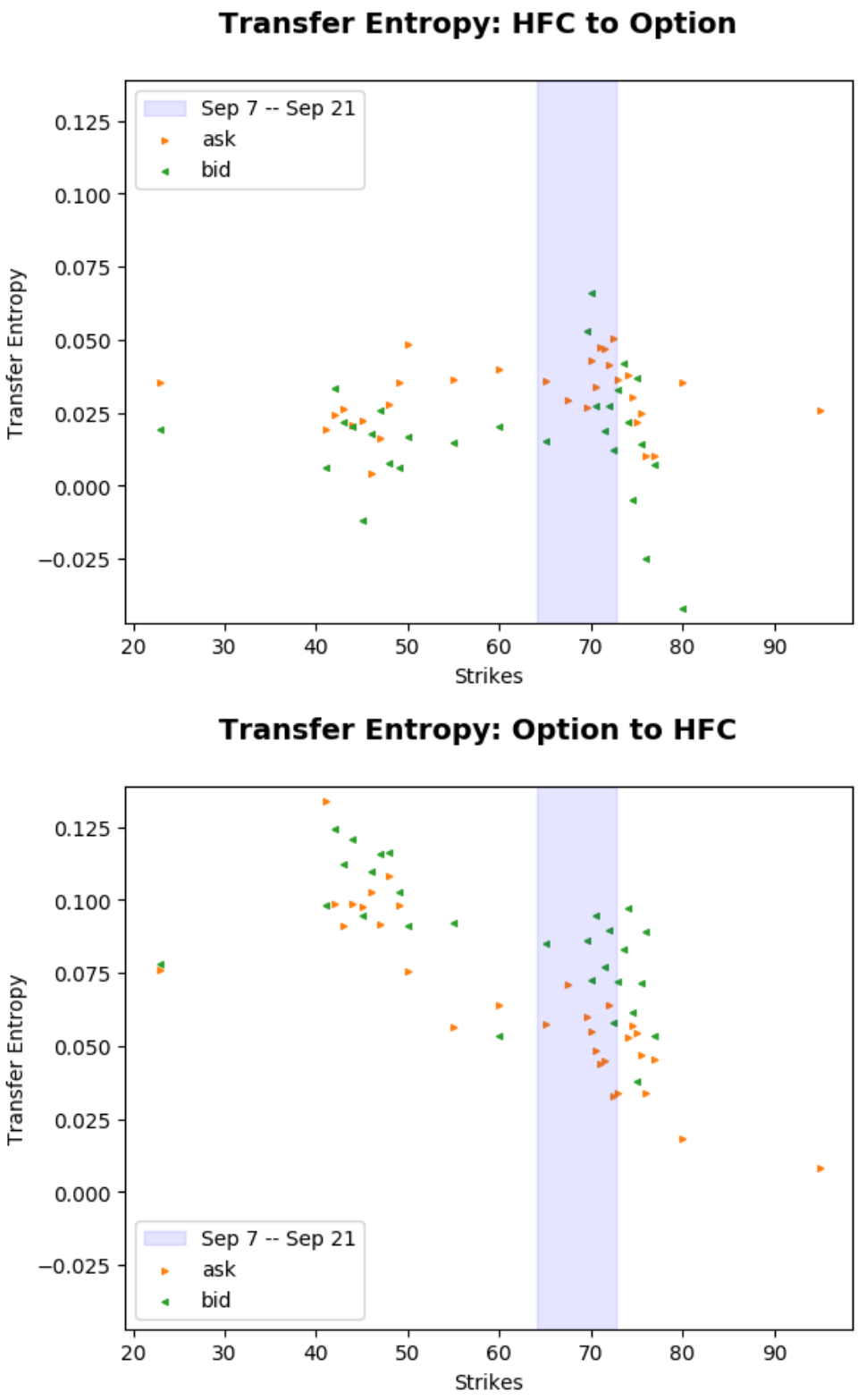

Figure 37 Transfer entropy between HFC and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

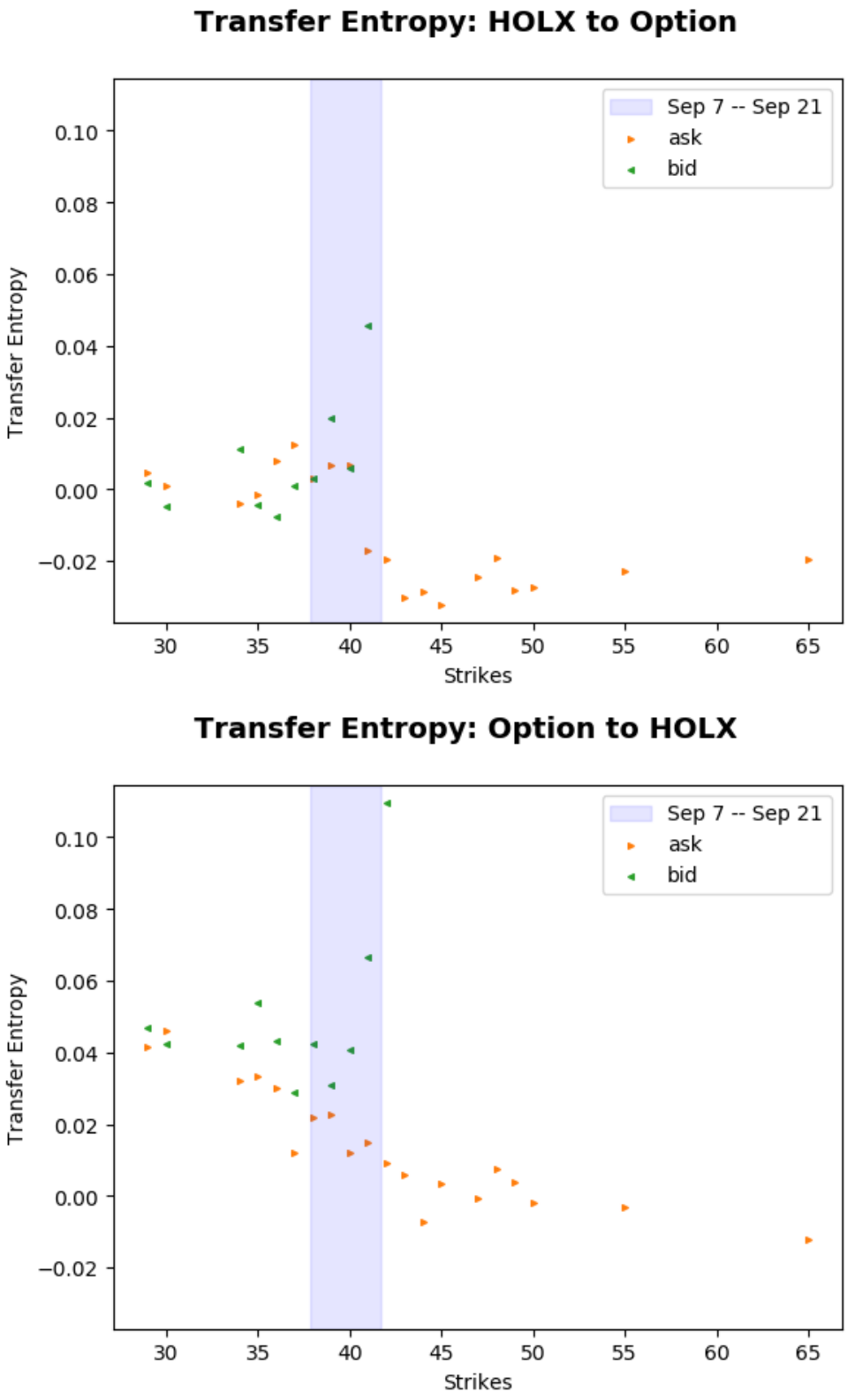

Figure 38 Transfer entropy between HOLX and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 8 seconds. Option to underlying timescale: 12 seconds.

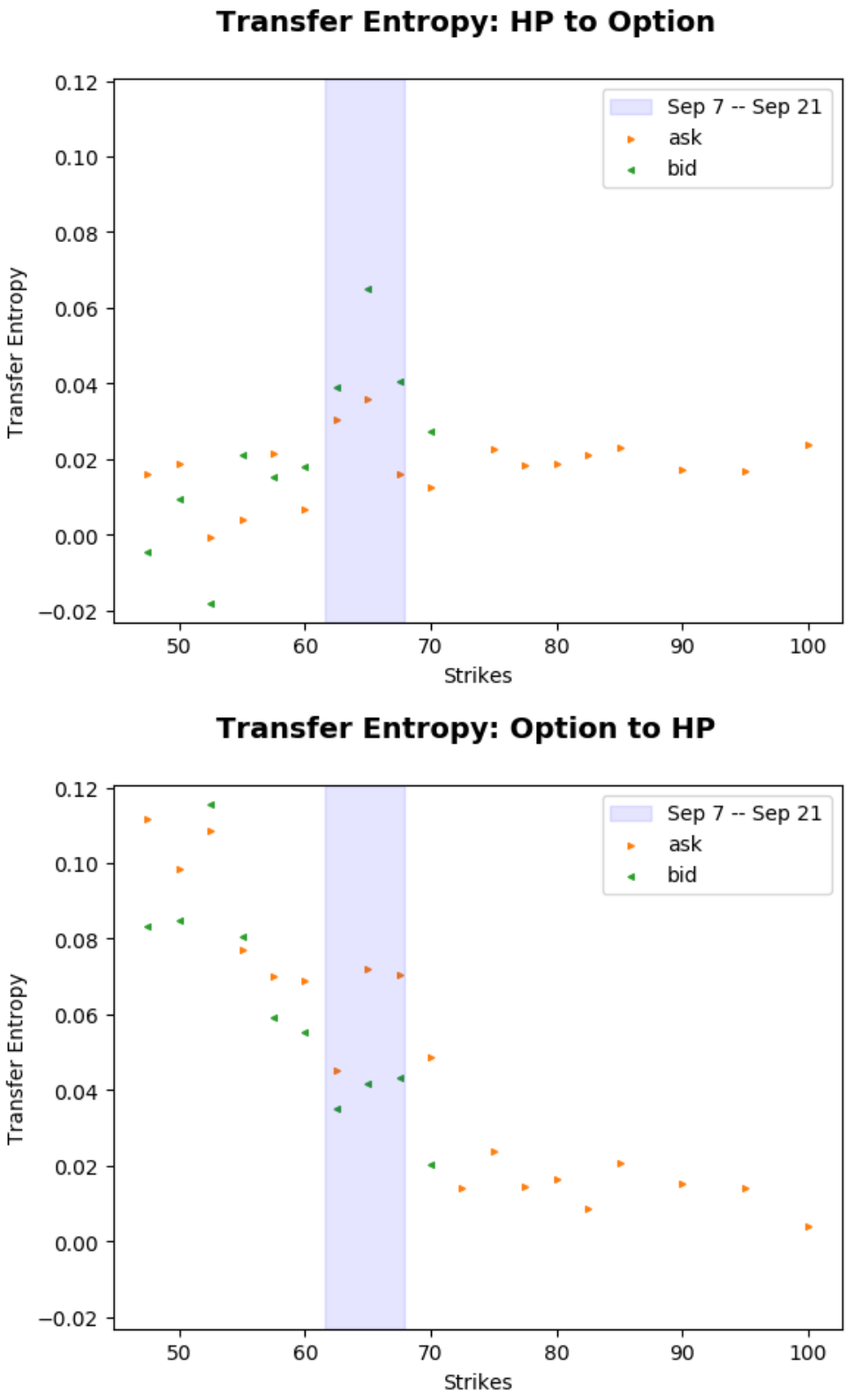

Figure 39 Transfer entropy between HP and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

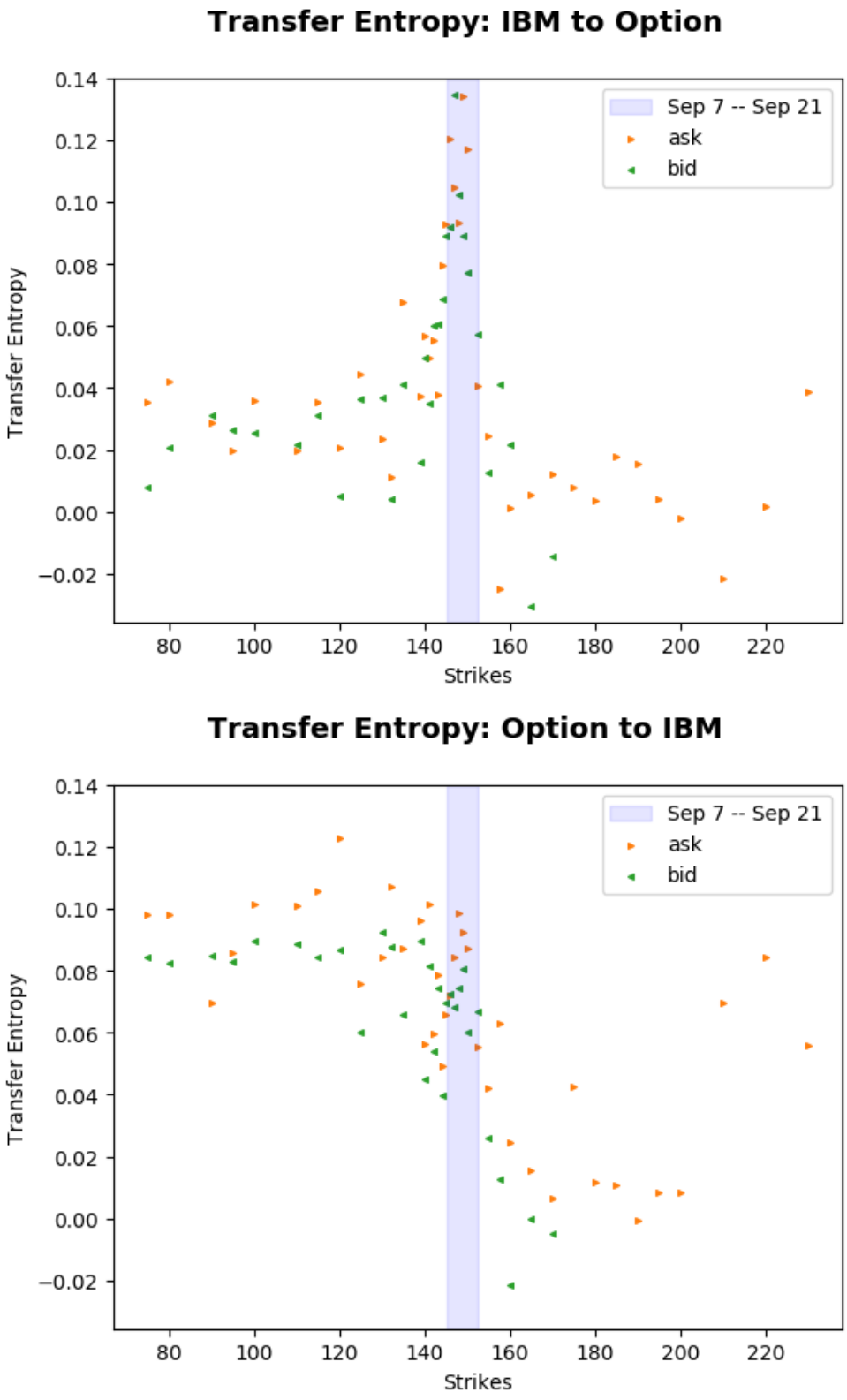

Figure 40 Transfer entropy between IBM and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 8 seconds.

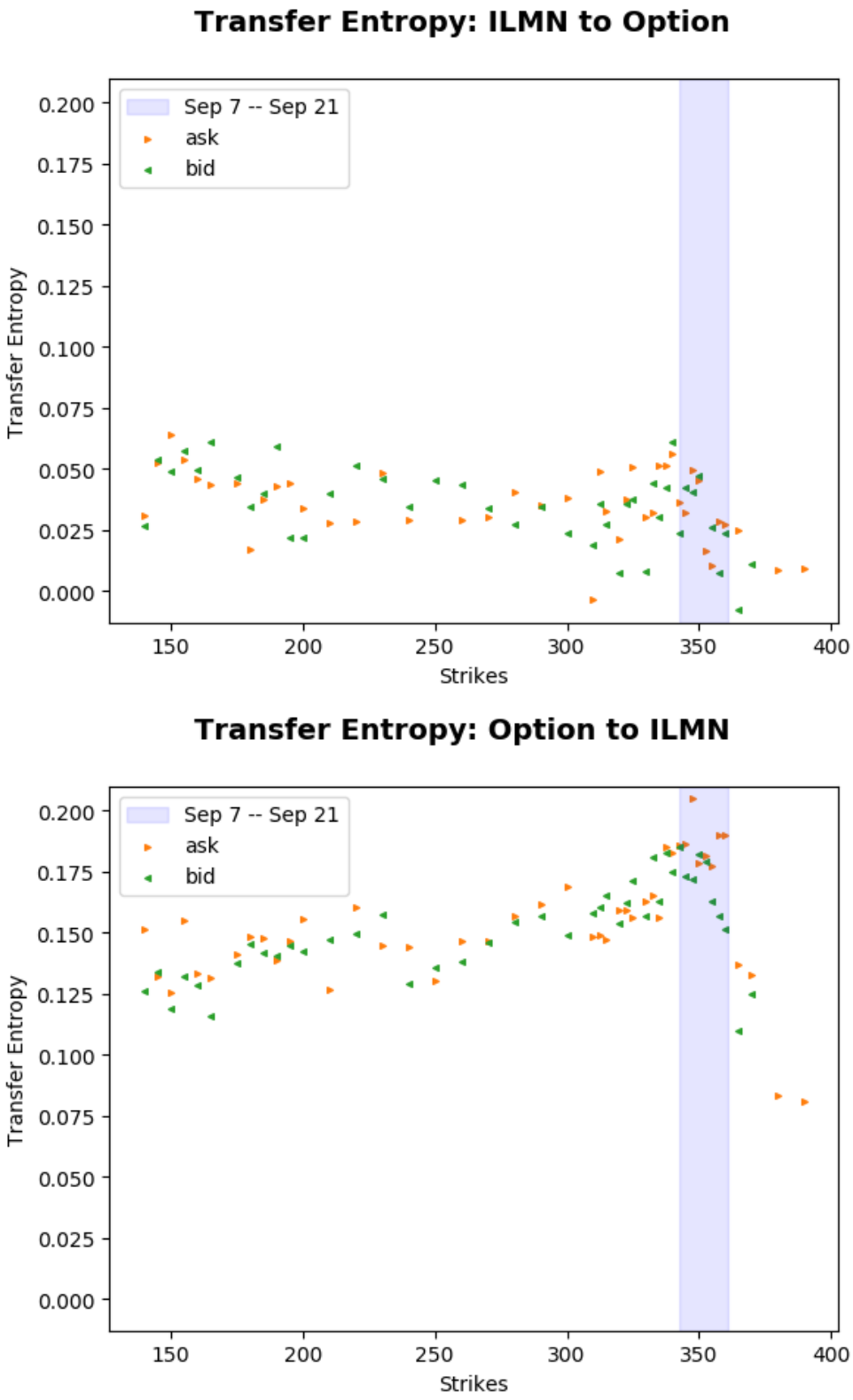

Figure 41 Transfer entropy between ILMN and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

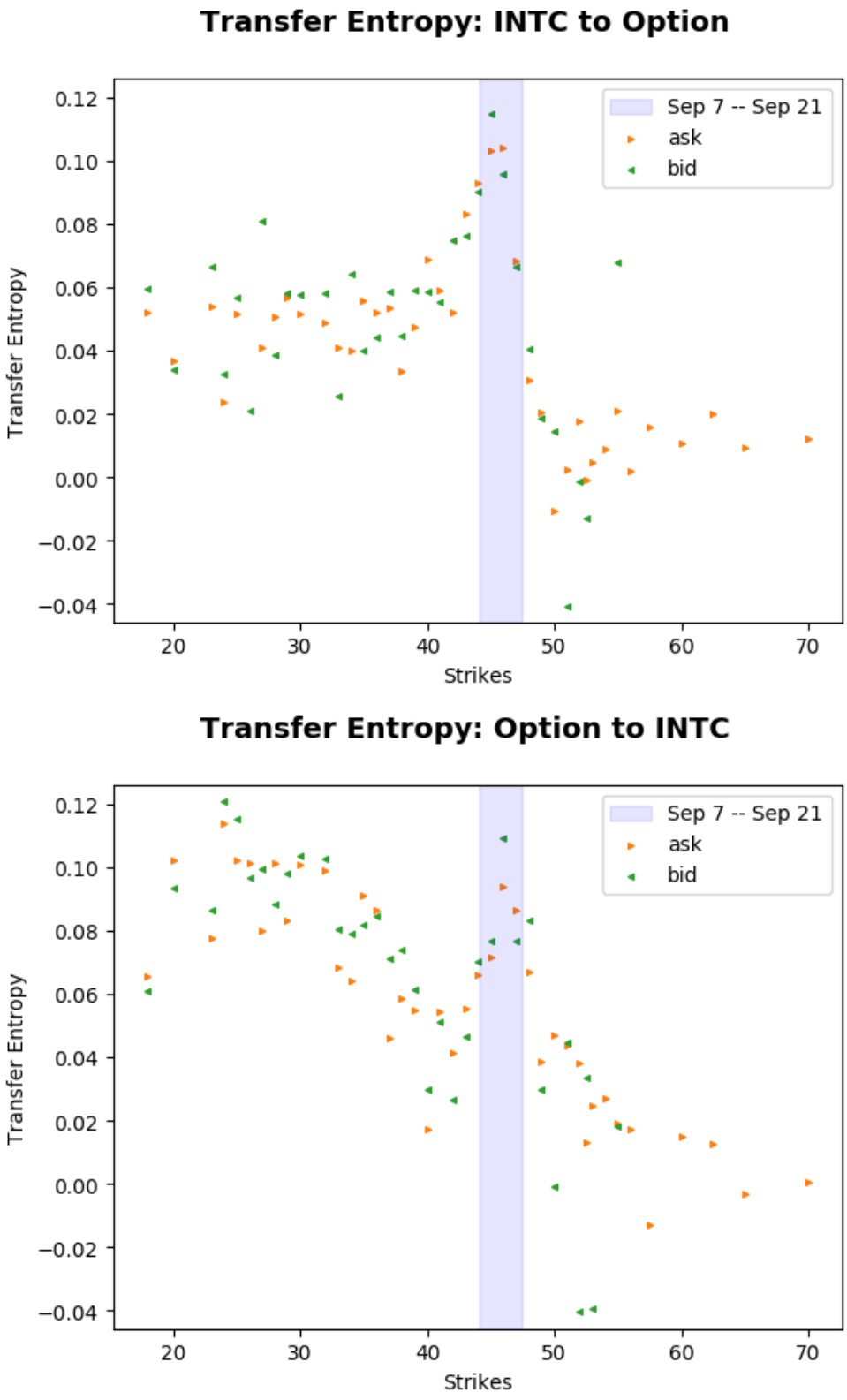

Figure 42 Transfer entropy between INTC and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

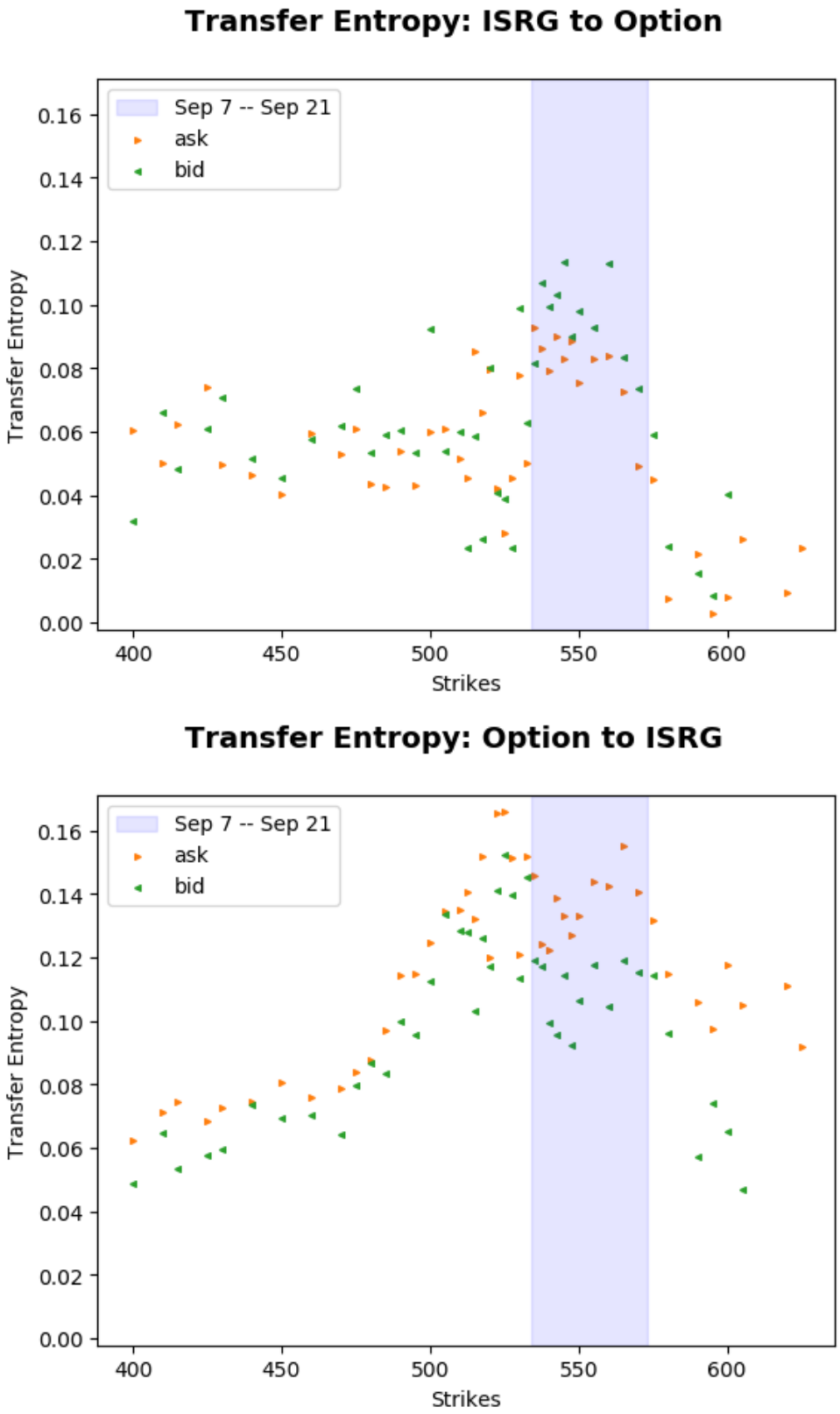

Figure 43 Transfer entropy between ISRG and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 8 seconds.

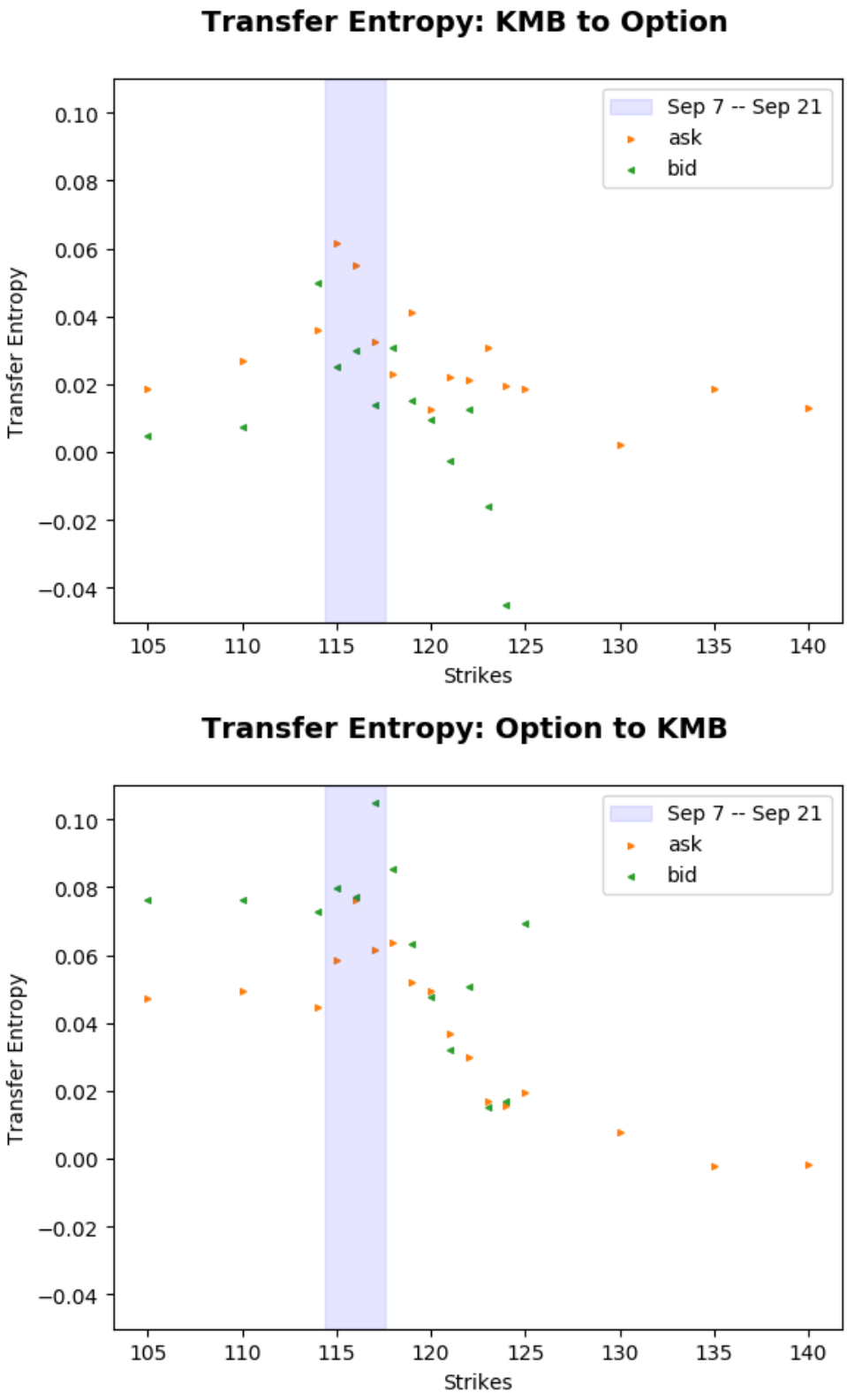

Figure 44 Transfer entropy between KMB and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 12 seconds.

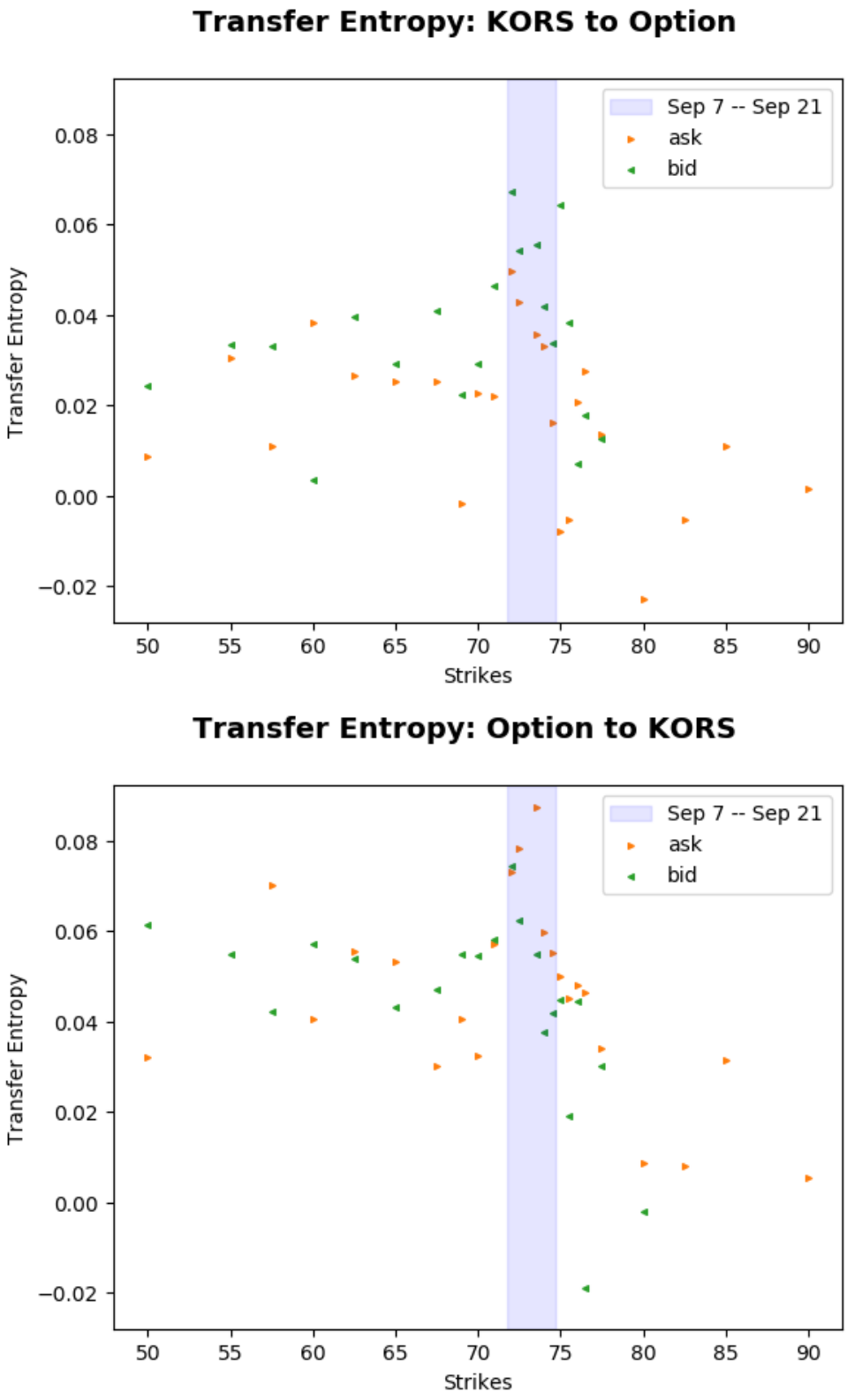

Figure 45 Transfer entropy between KORS and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

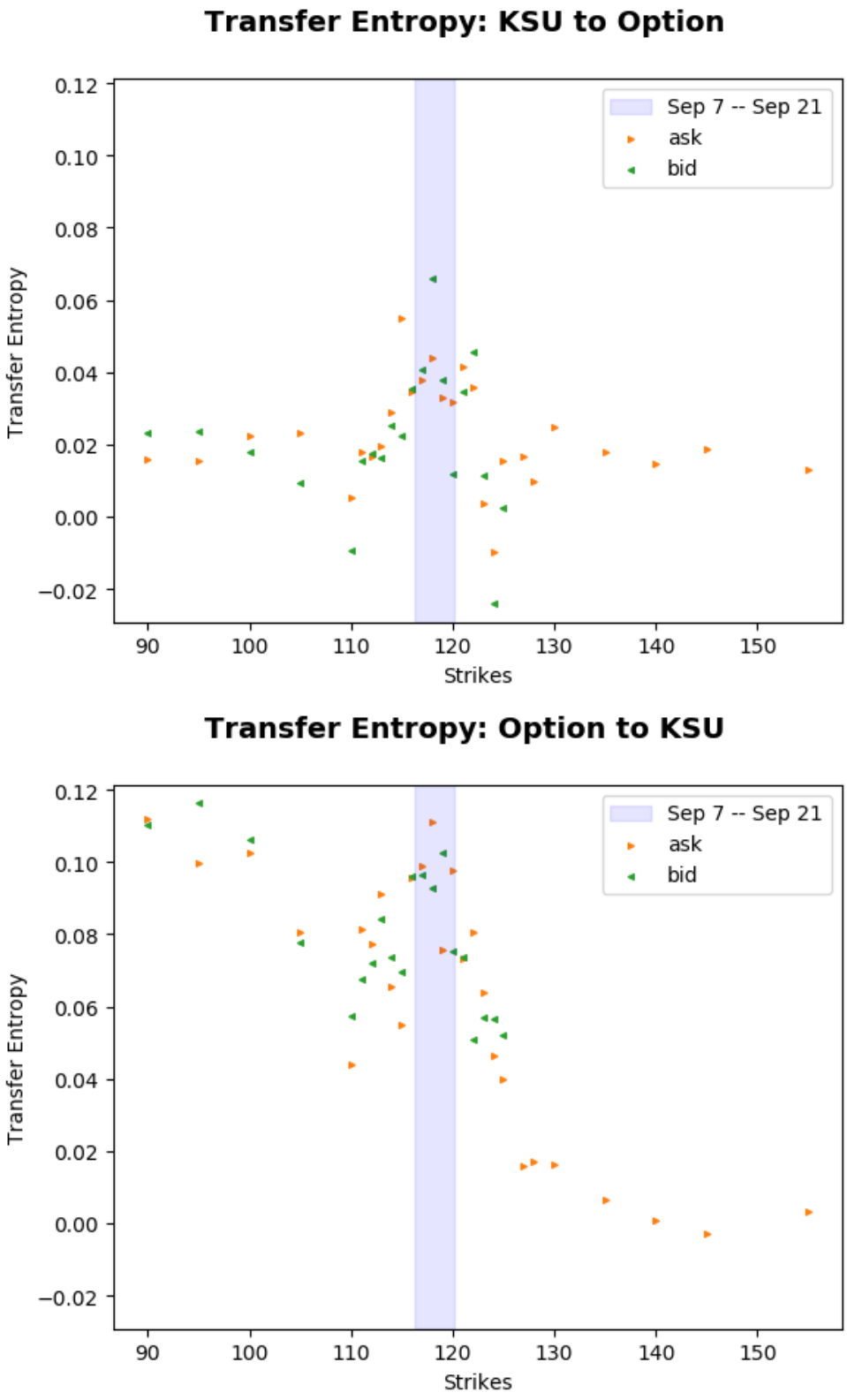

Figure 46 Transfer entropy between KSU and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

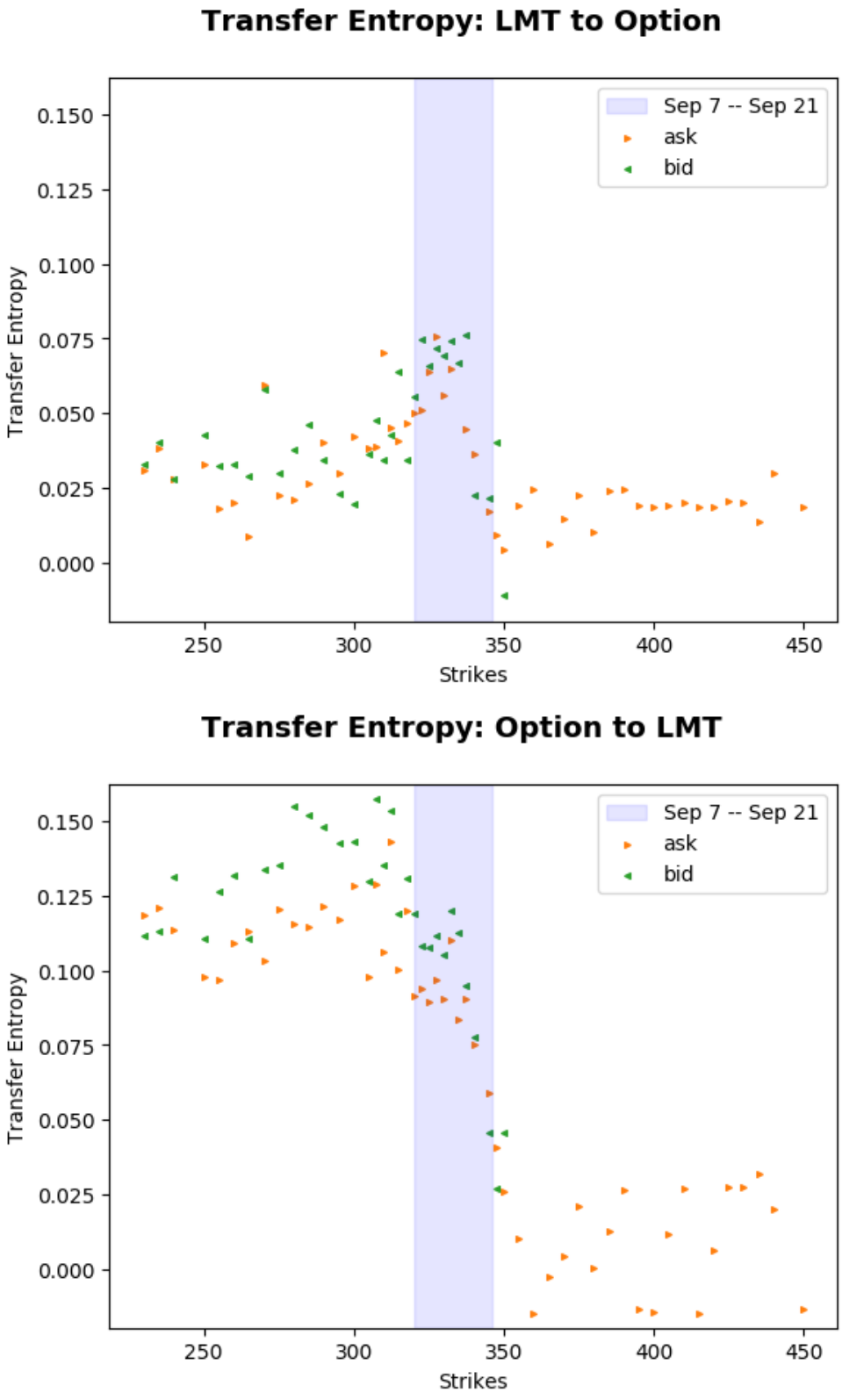

Figure 47 Transfer entropy between LMT and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 12 seconds.

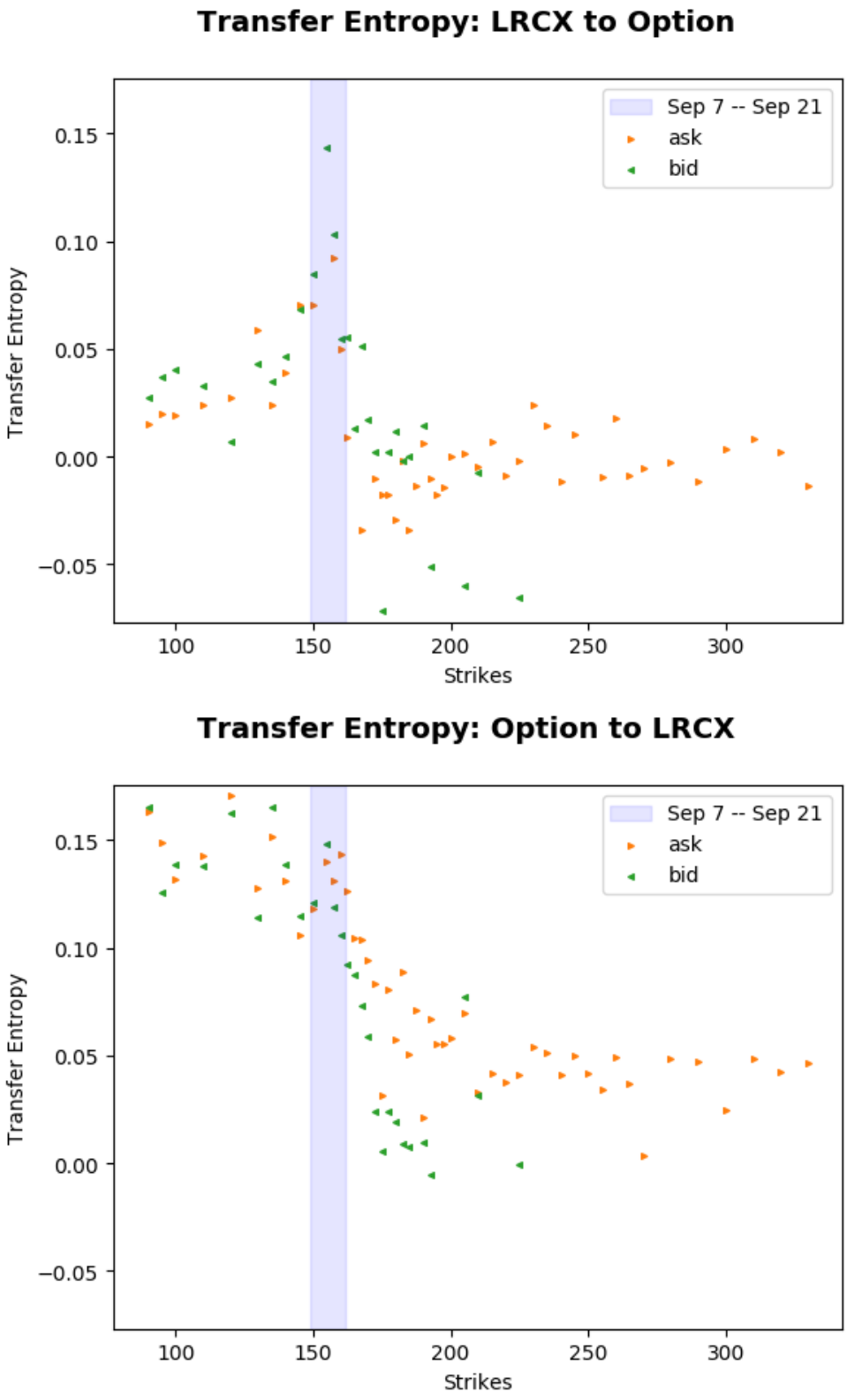

Figure 48 Transfer entropy between LRCX and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 10 seconds. Option to underlying timescale: 12 seconds.

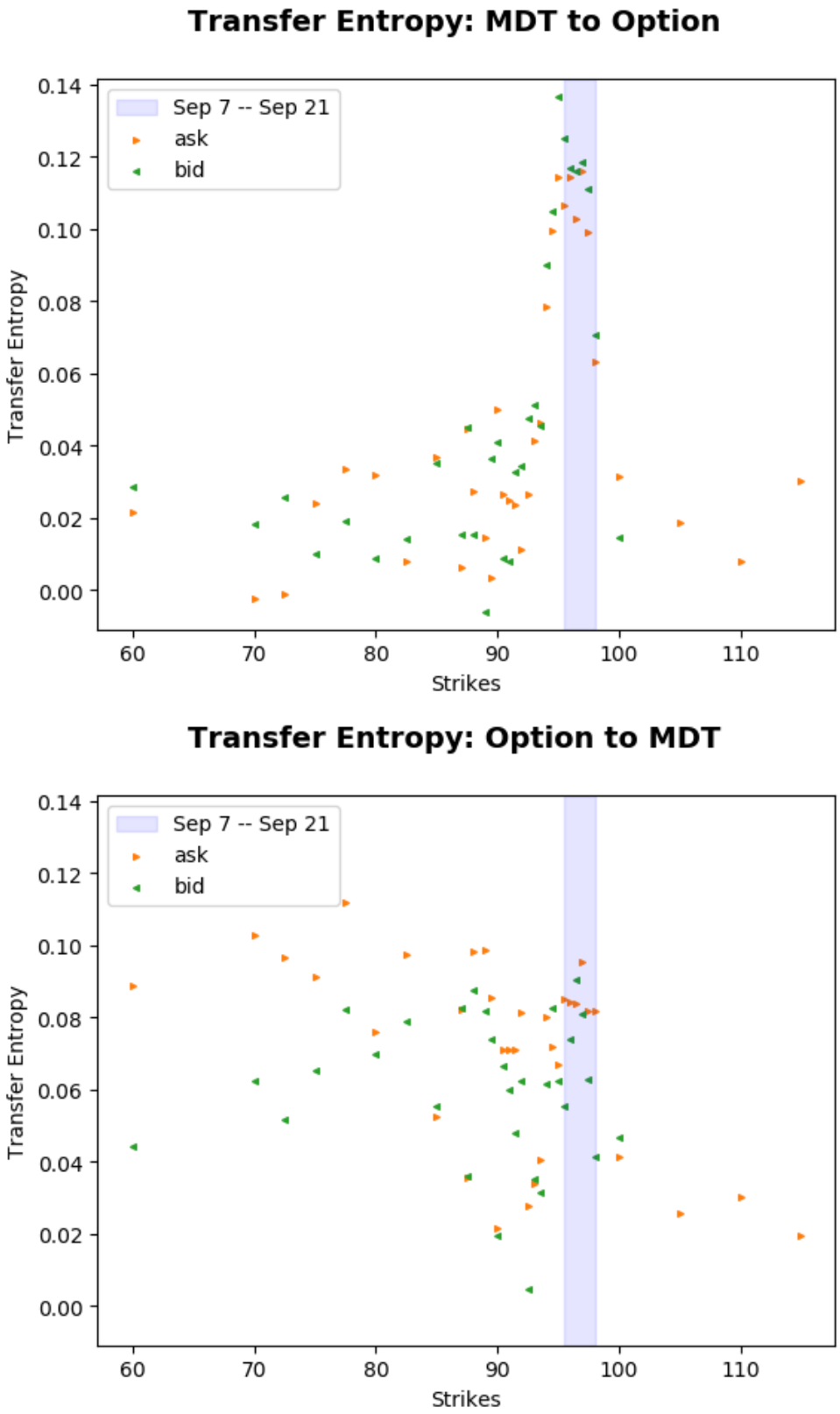

Figure 49 Transfer entropy between MDT and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 12 seconds.

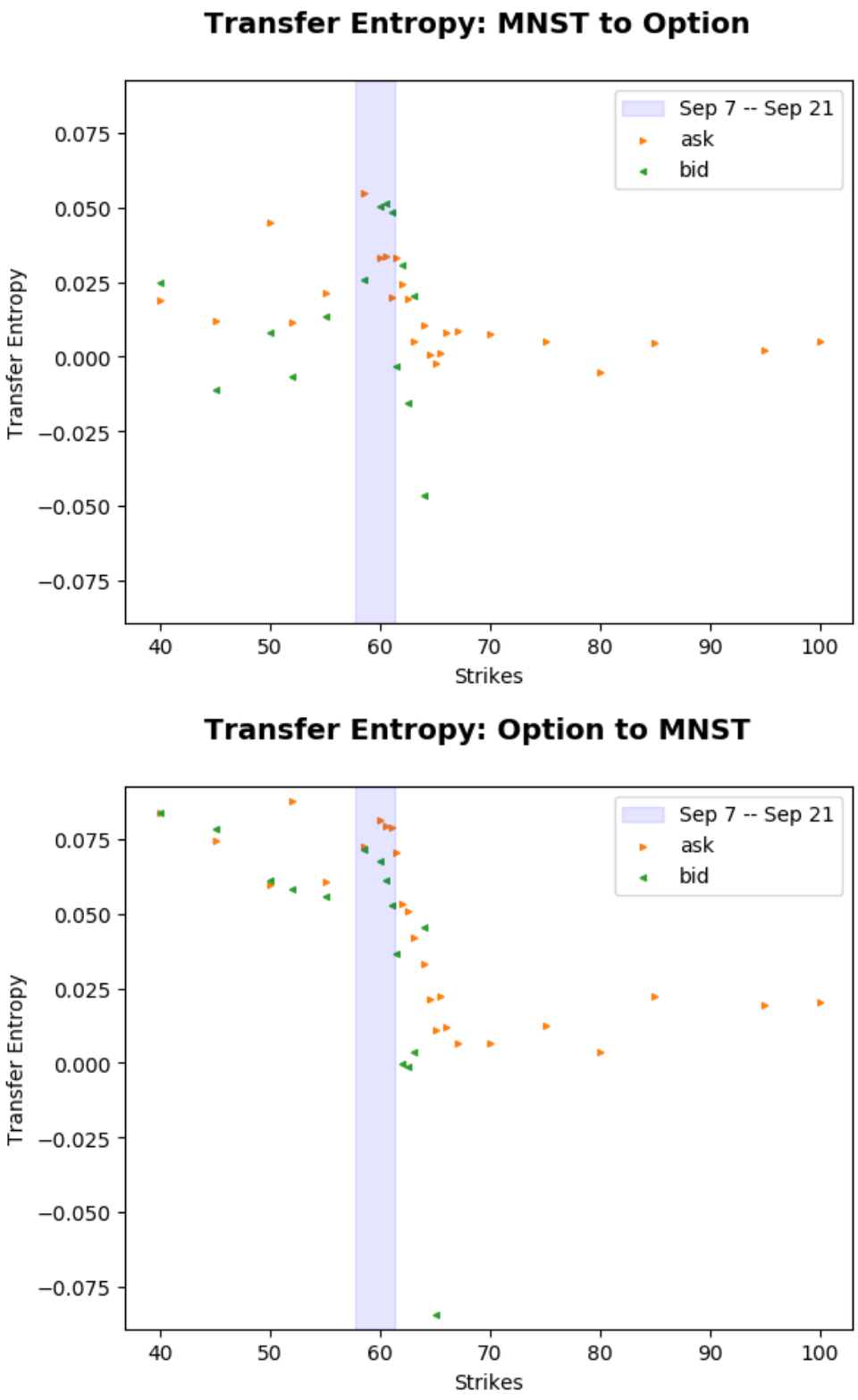

Figure 50 Transfer entropy between MNST and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

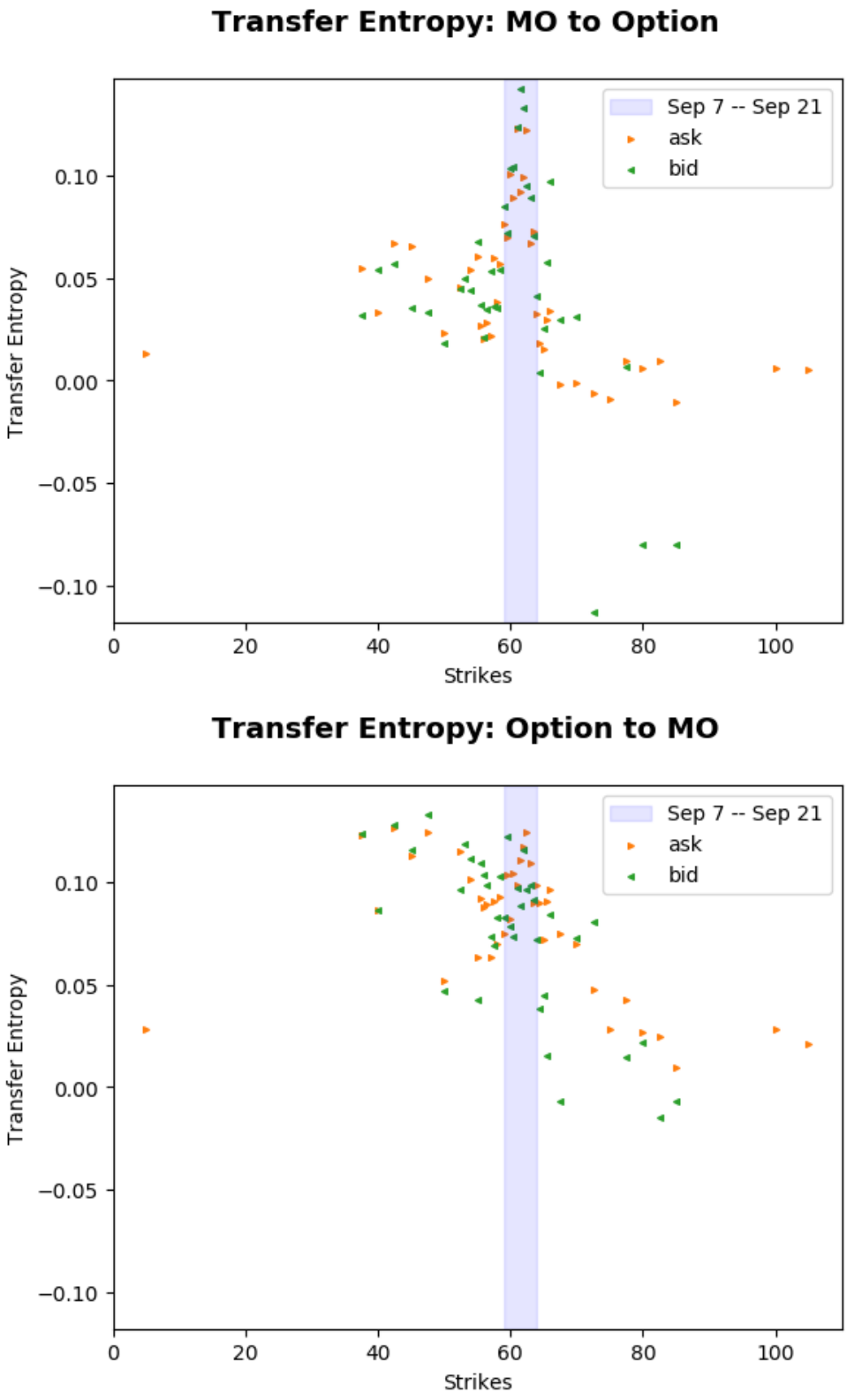

Figure 51 Transfer entropy between MO and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

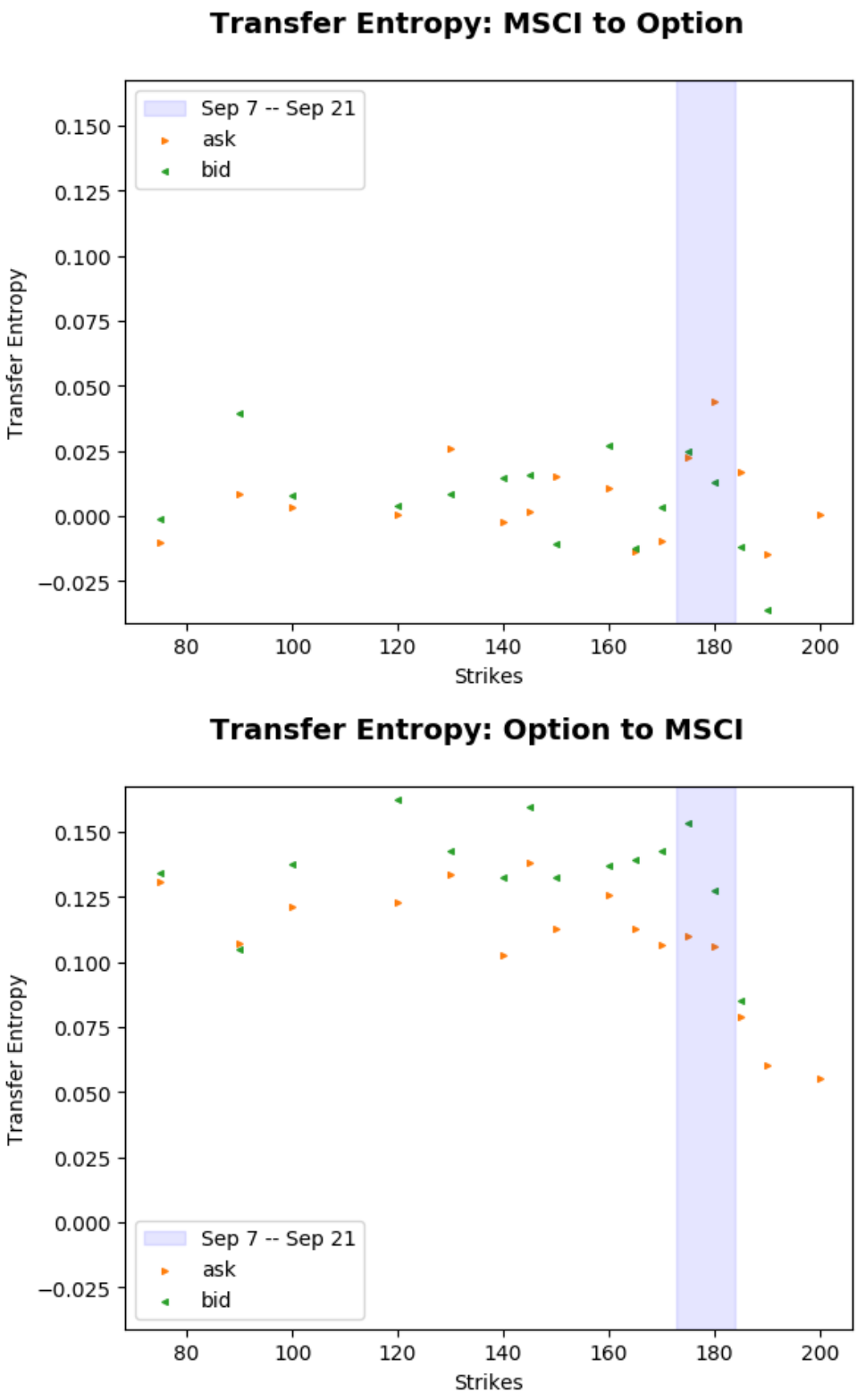

Figure 52 Transfer entropy between MSCI and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 10 seconds. Option to underlying timescale: 14 seconds.

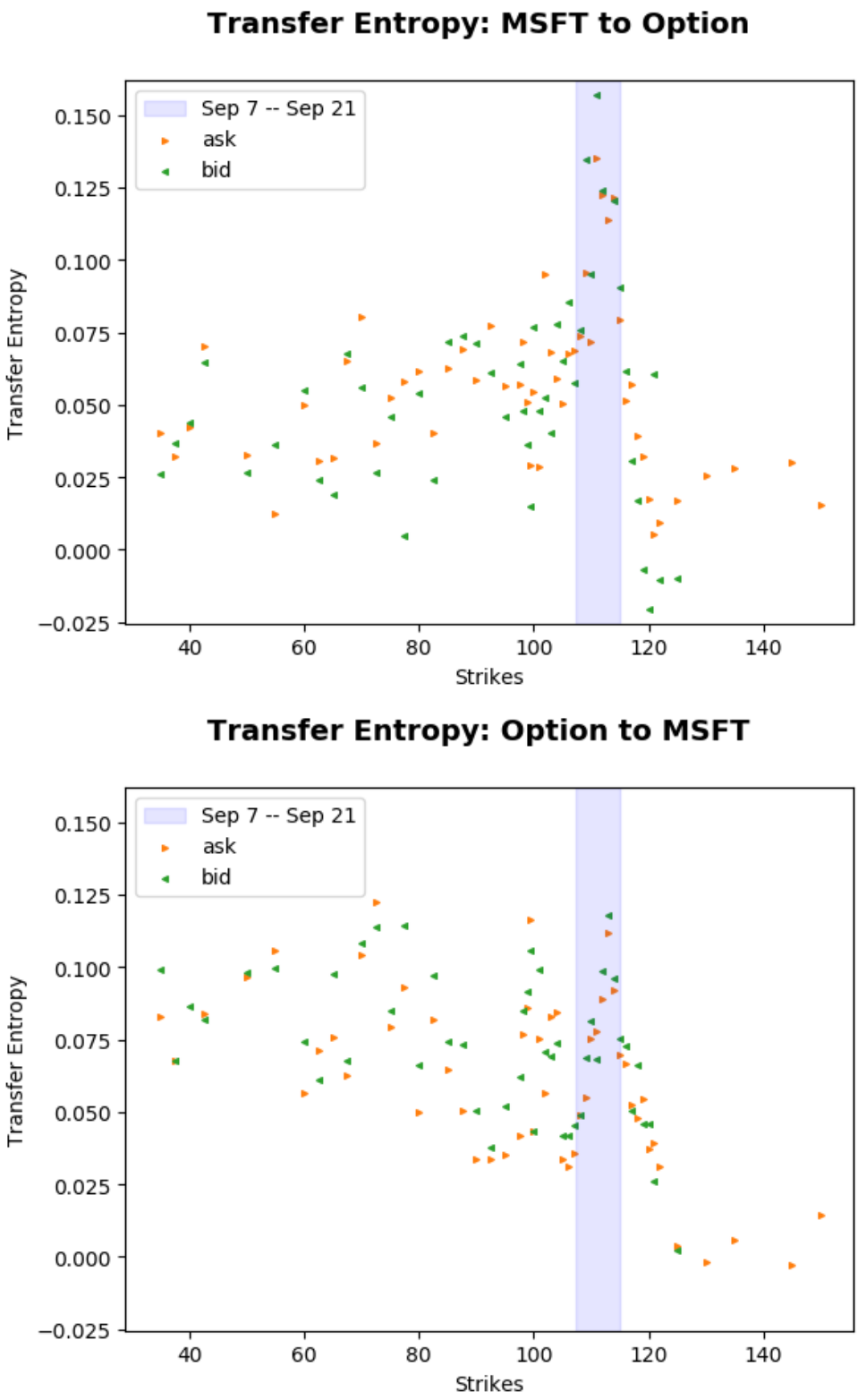

Figure 53 Transfer entropy between MSFT and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 10 seconds.

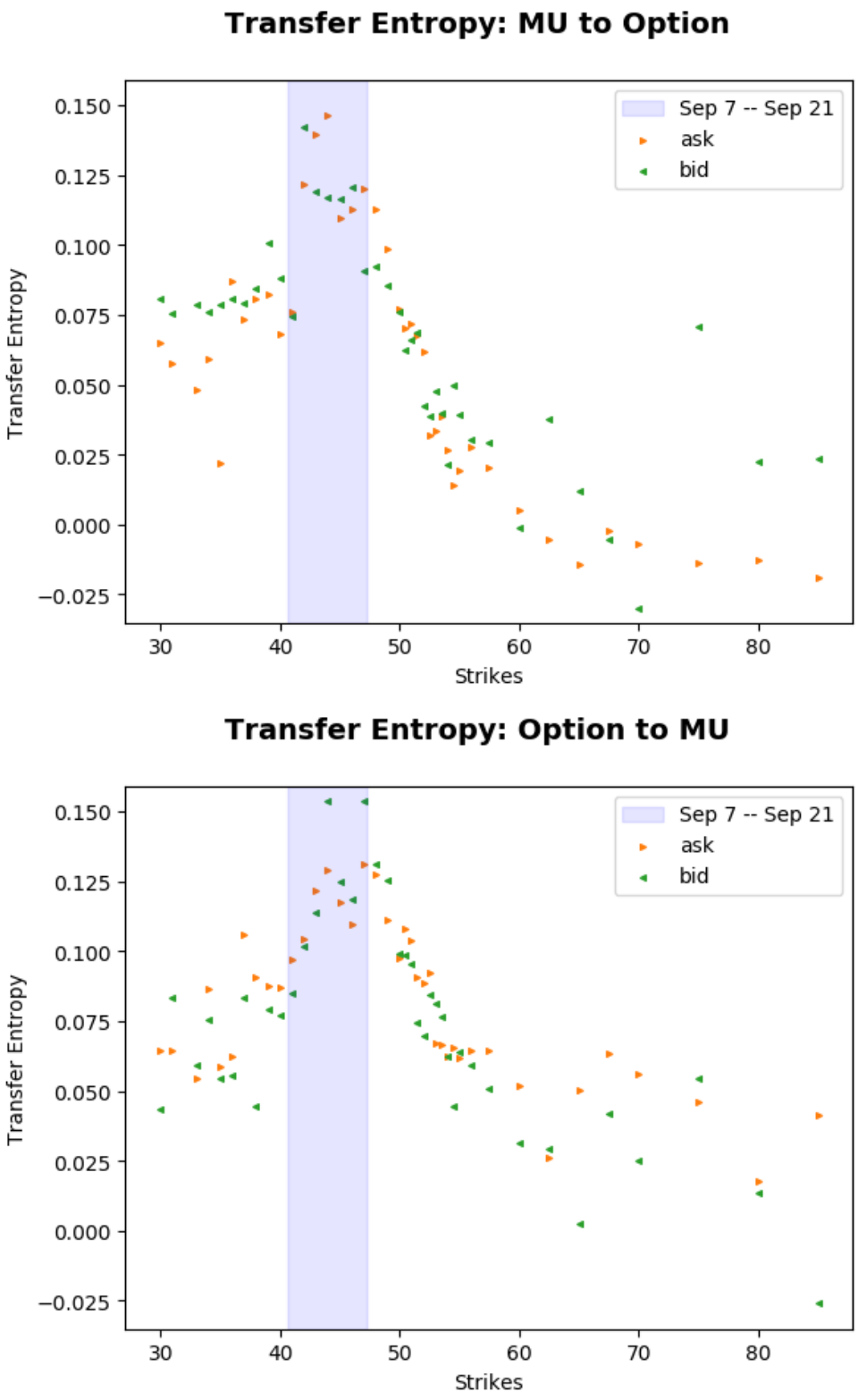

Figure 54 Transfer entropy between MU and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 12 seconds.

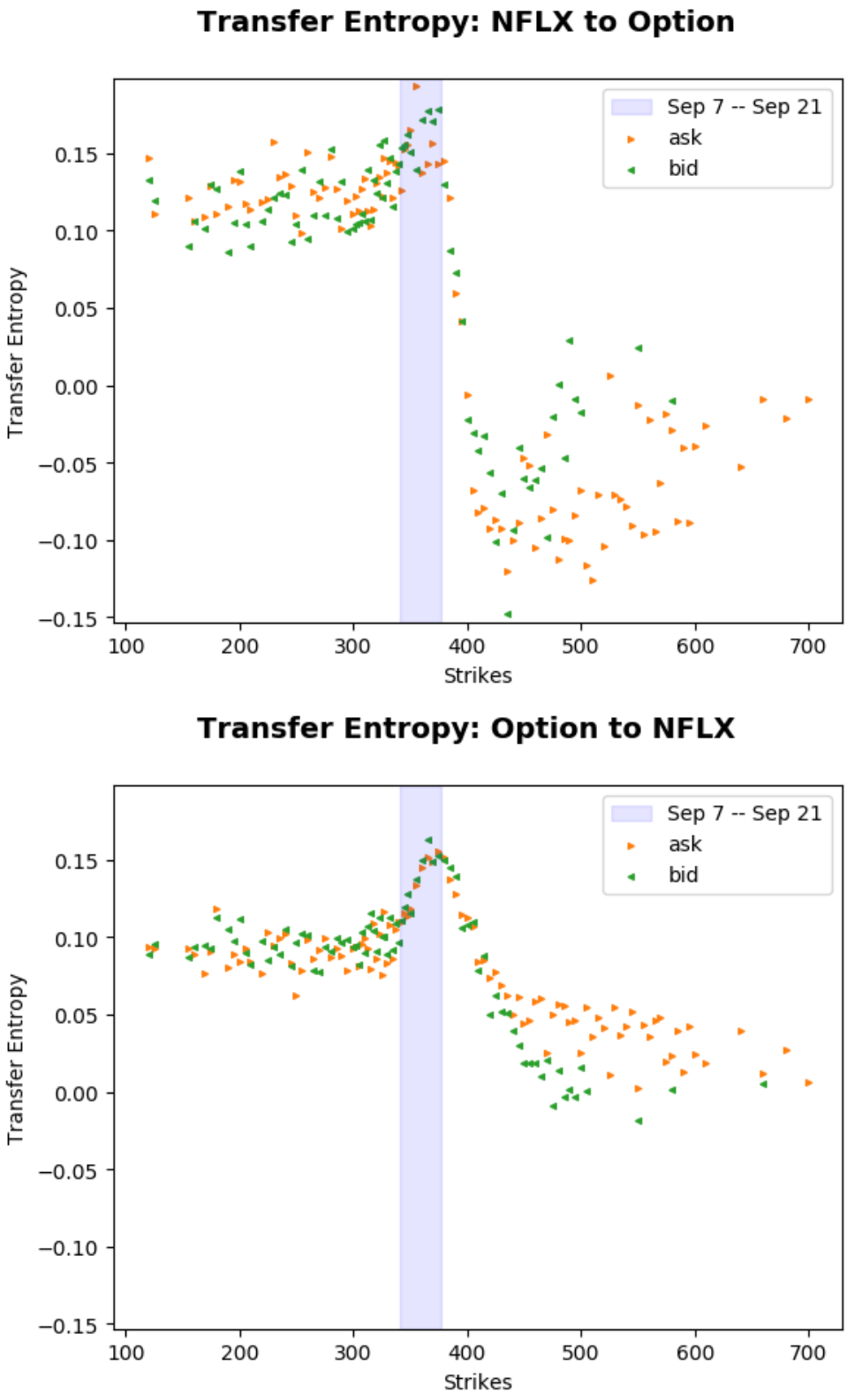

Figure 55 Transfer entropy between NFLX and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 8 seconds.

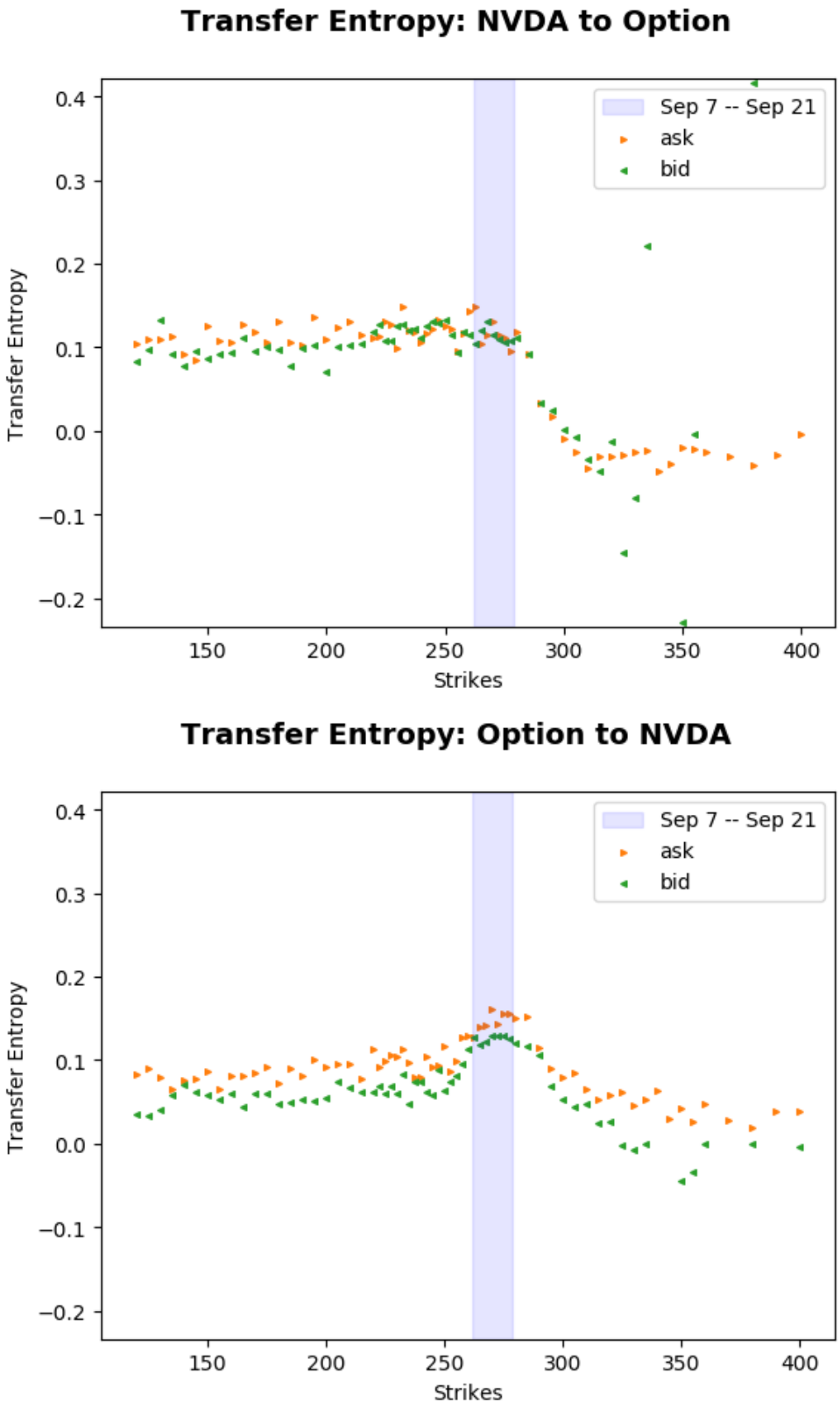

Figure 56 Transfer entropy between NVDA and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 12 seconds.

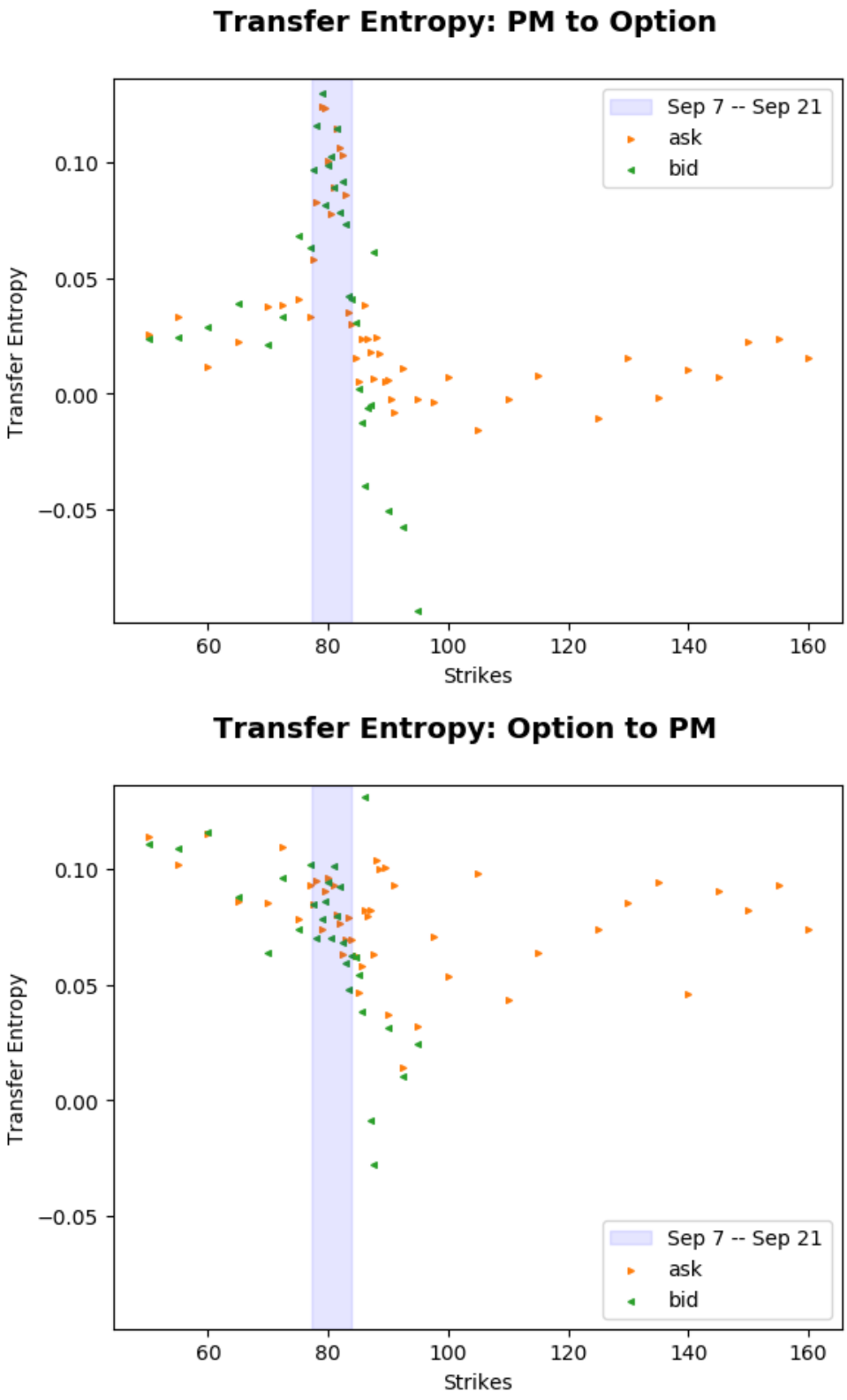

Figure 57 Transfer entropy between PM and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

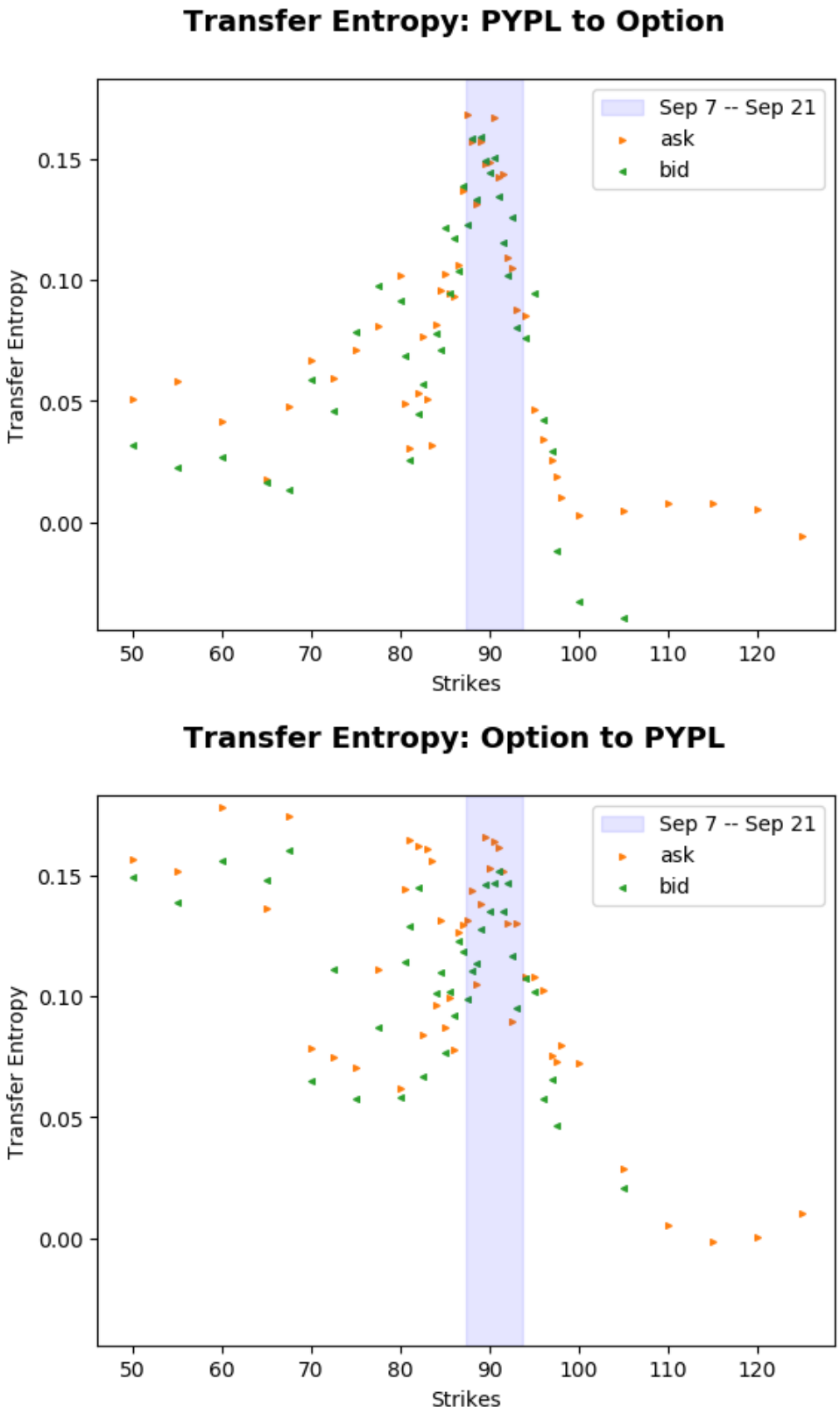

Figure 58 Transfer entropy between PYPL and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

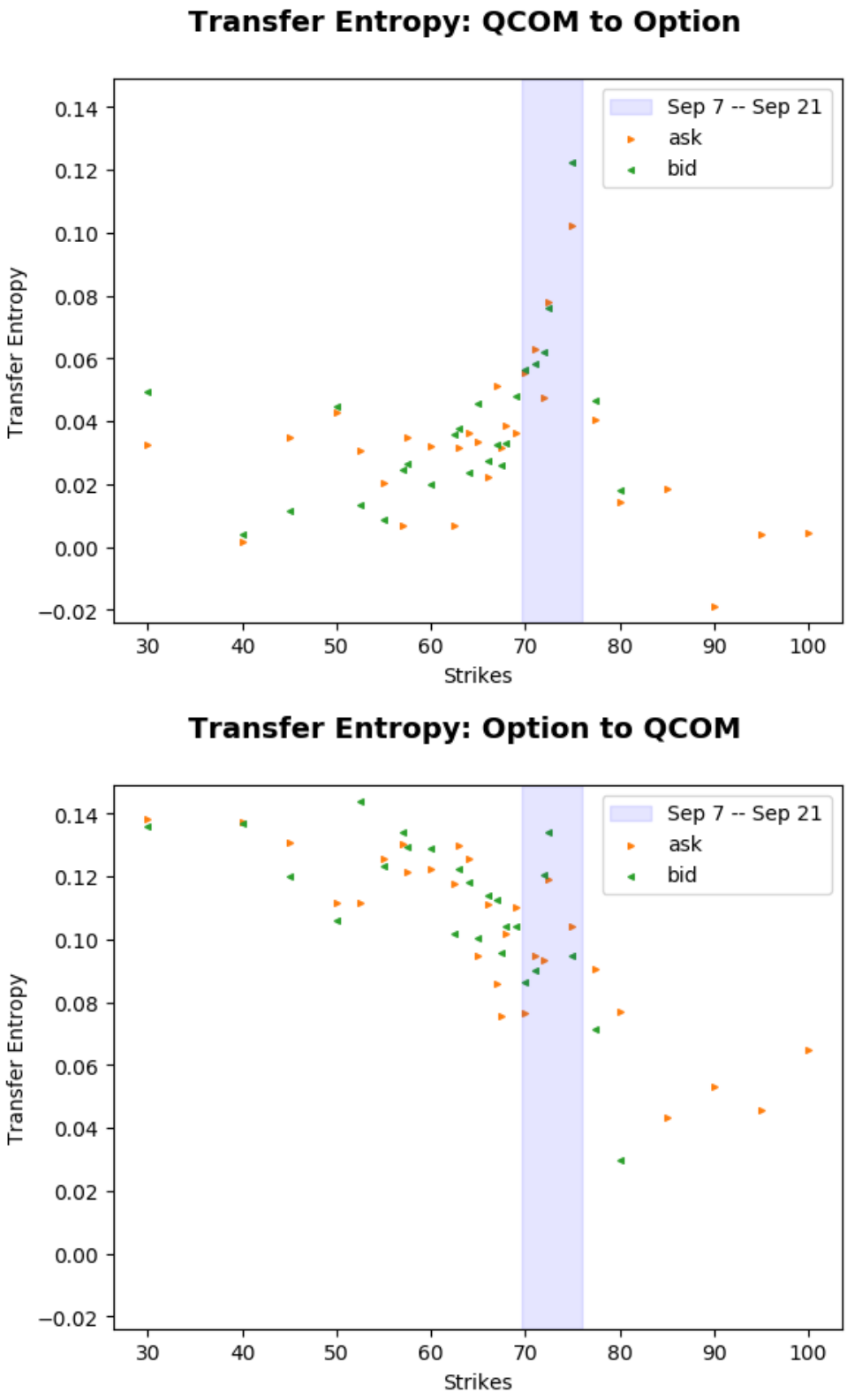

Figure 59 Transfer entropy between QCOM and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

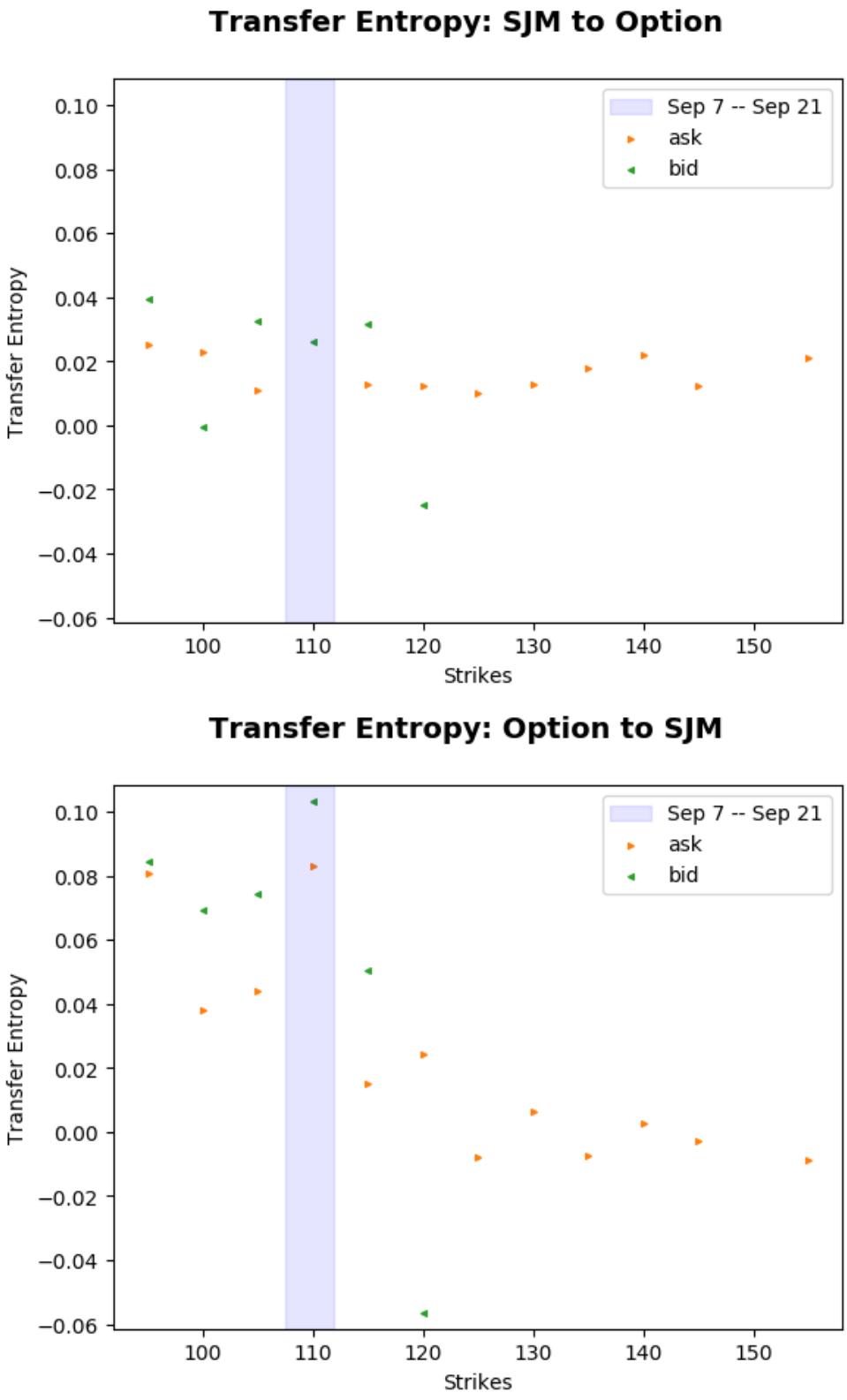

Figure 60 Transfer entropy between SJM and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

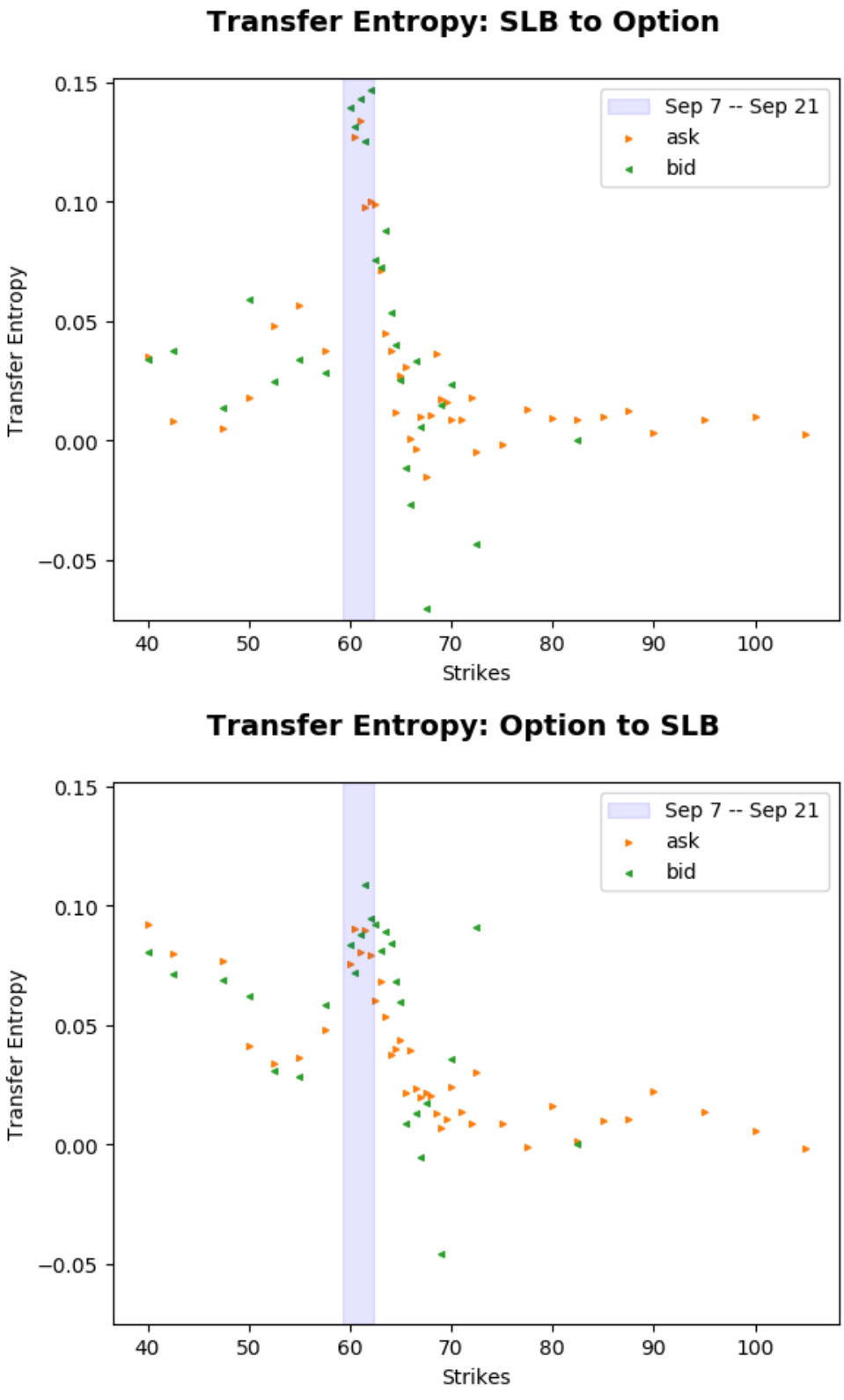

Figure 61 Transfer entropy between SLB and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 10 seconds.

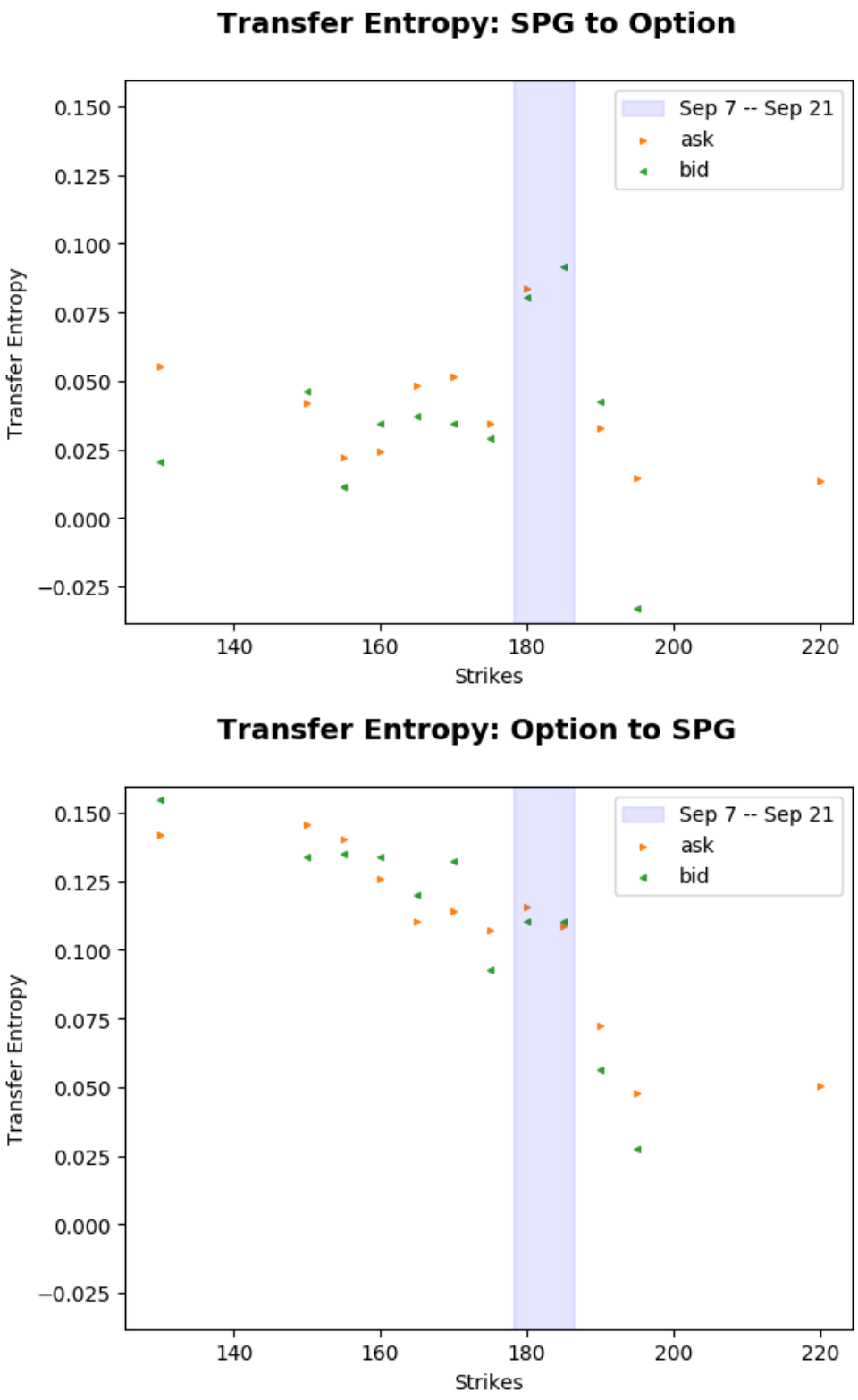

Figure 62 Transfer entropy between SPG and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

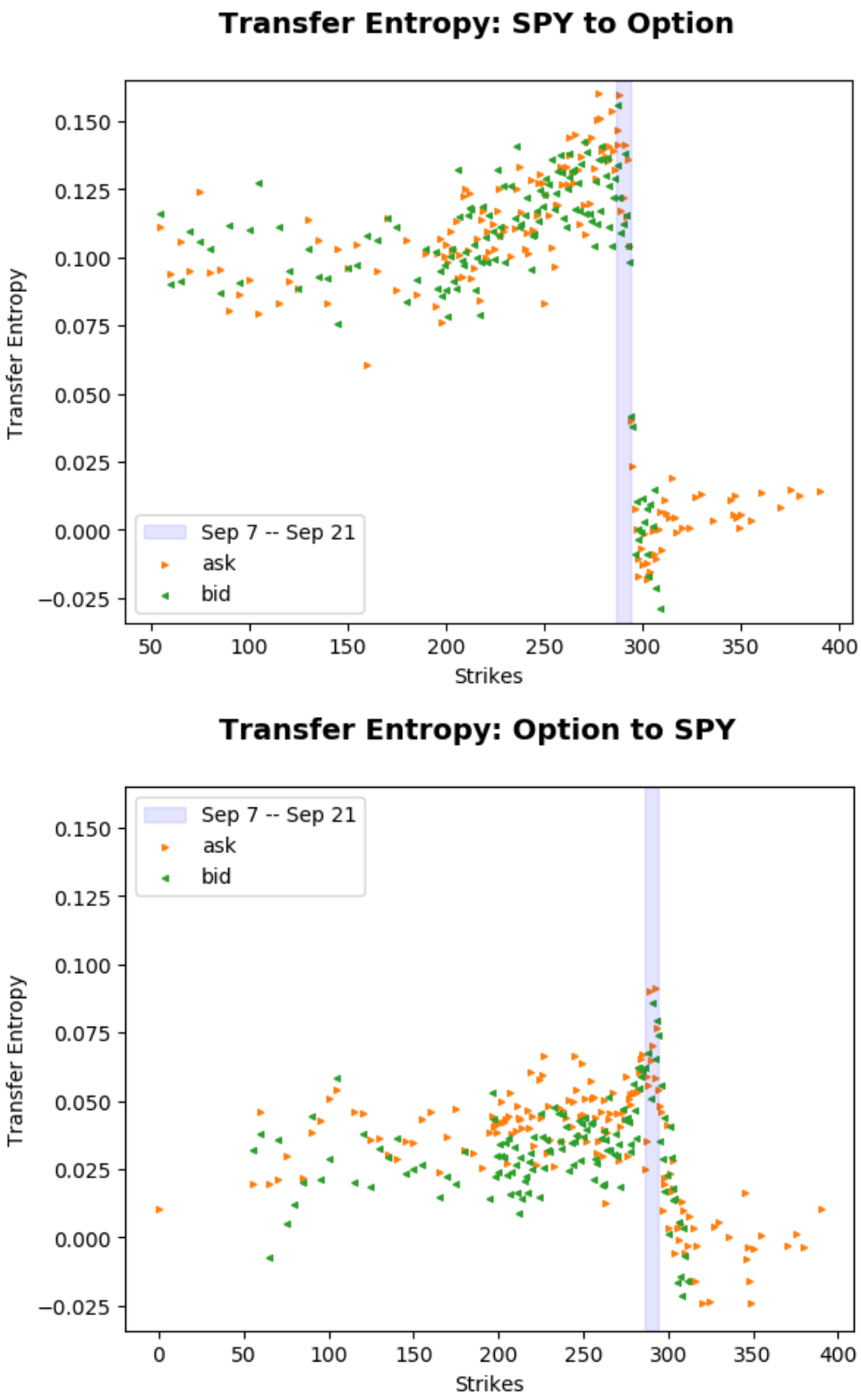

Figure 63 Transfer entropy between SPY and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

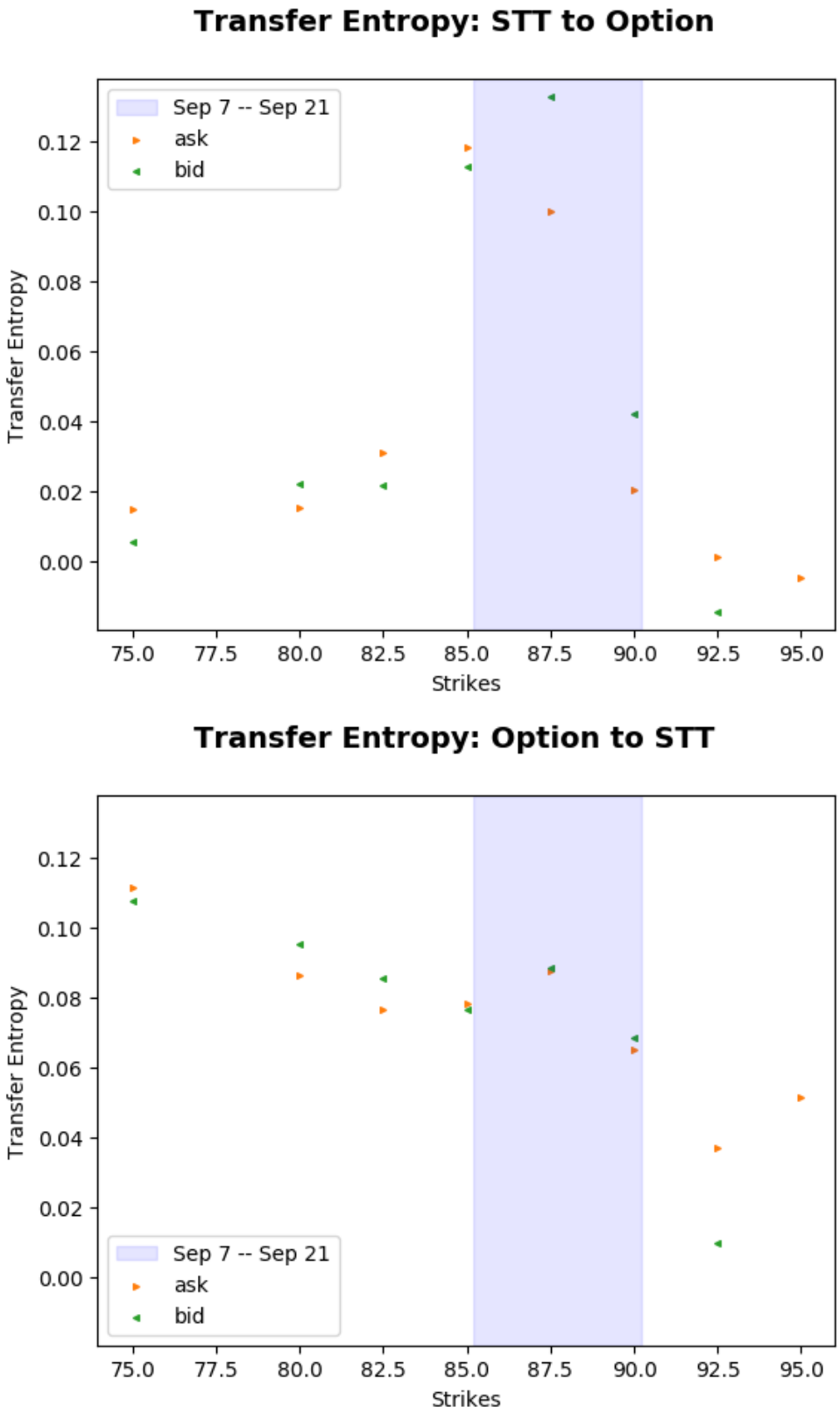

Figure 64 Transfer entropy between STT and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 10 seconds.

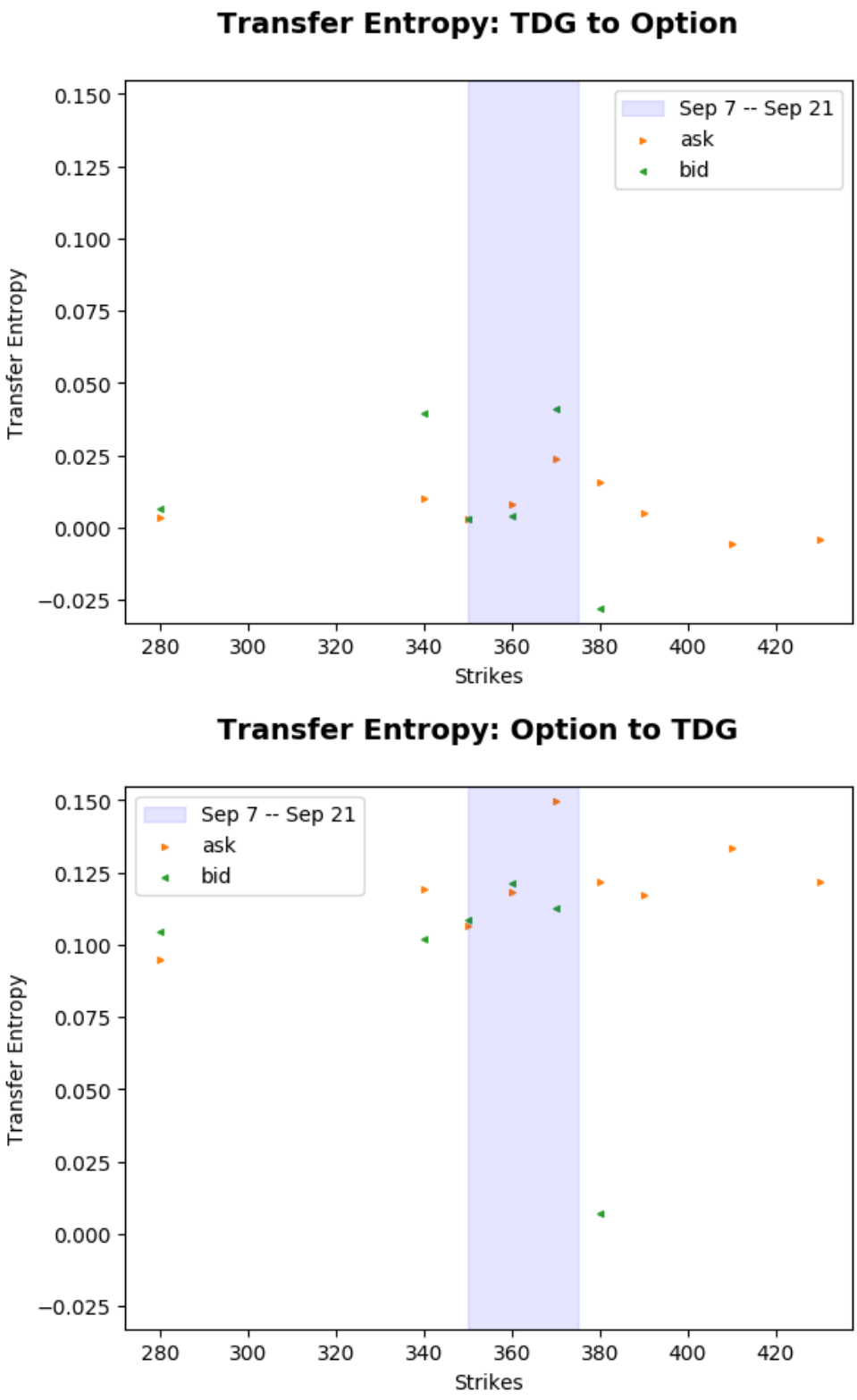

Figure 65 Transfer entropy between TDG and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 10 seconds. Option to underlying timescale: 10 seconds.

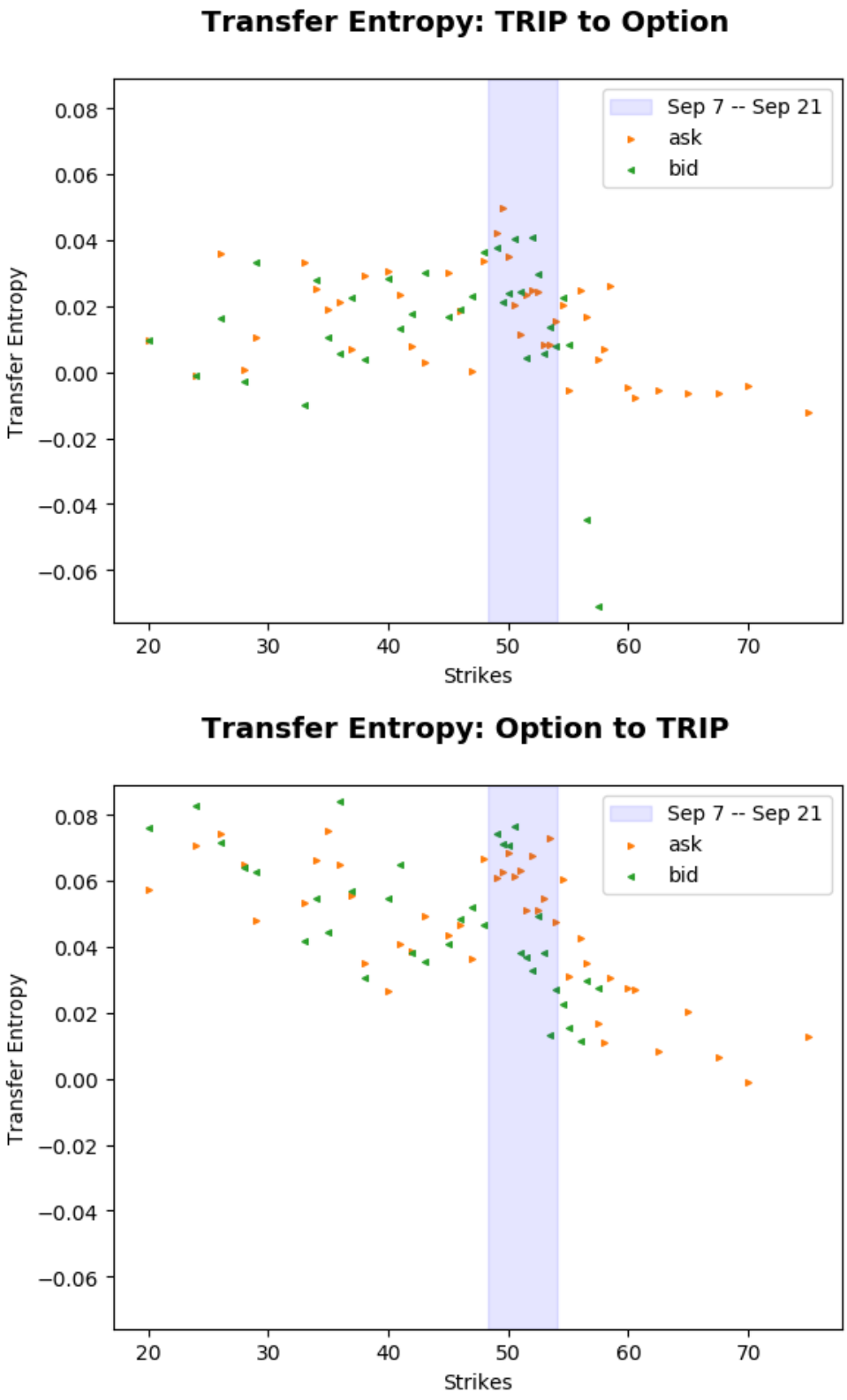

Figure 66 Transfer entropy between TRIP and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 12 seconds.

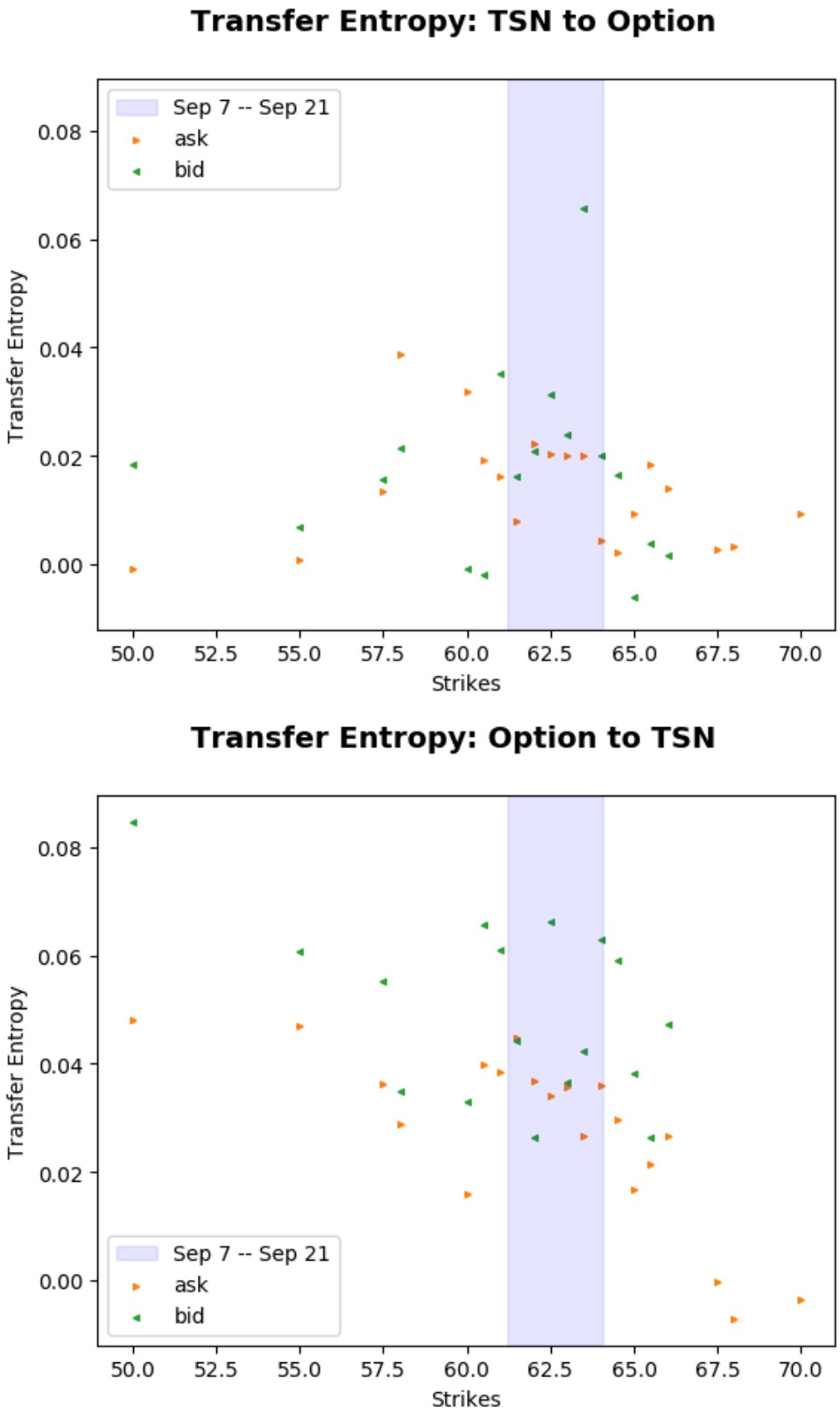

Figure 67 Transfer entropy between TSN and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 12 seconds.

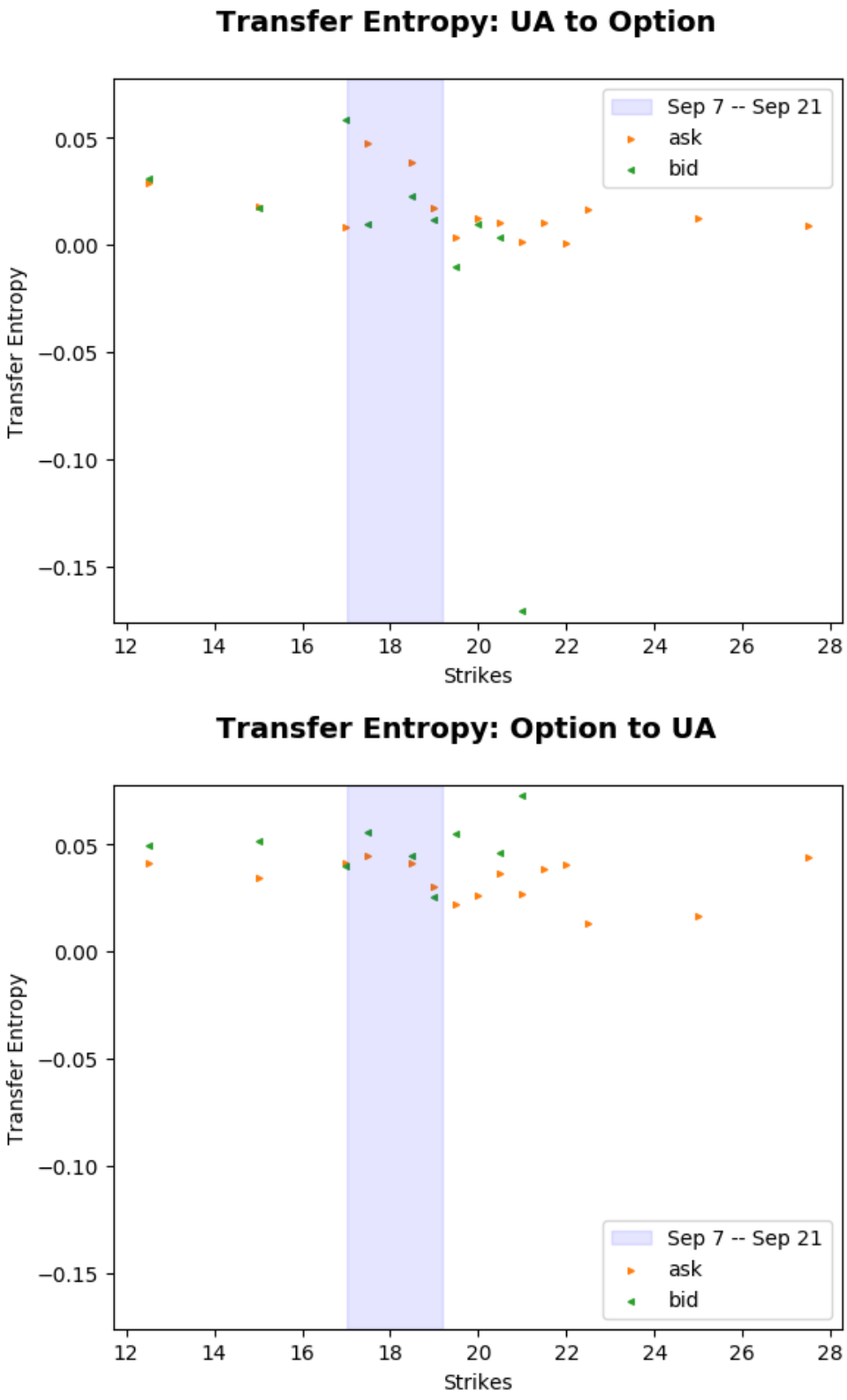

Figure 68 Transfer entropy between UA and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 10 seconds.

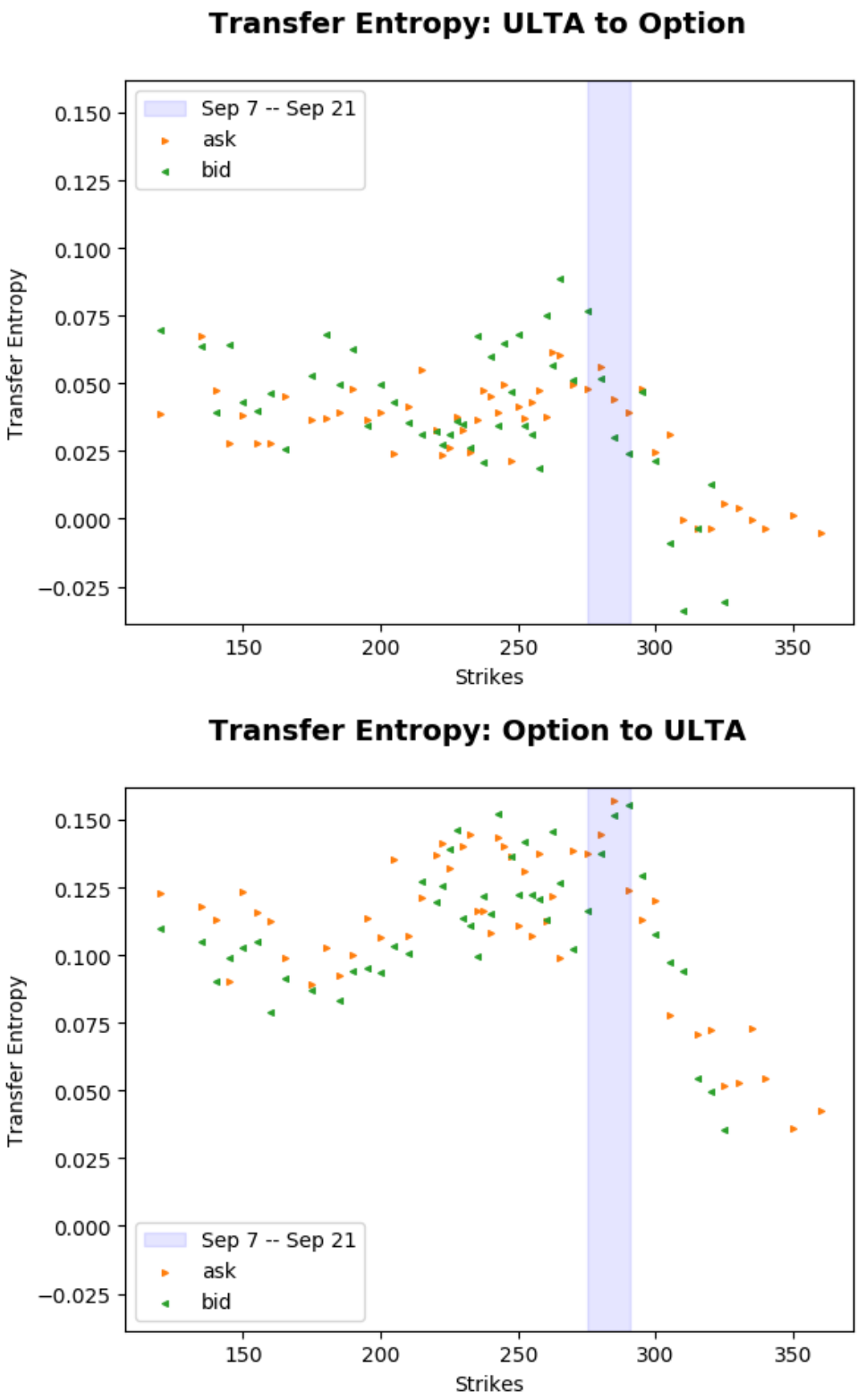

Figure 69 Transfer entropy between ULTA and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

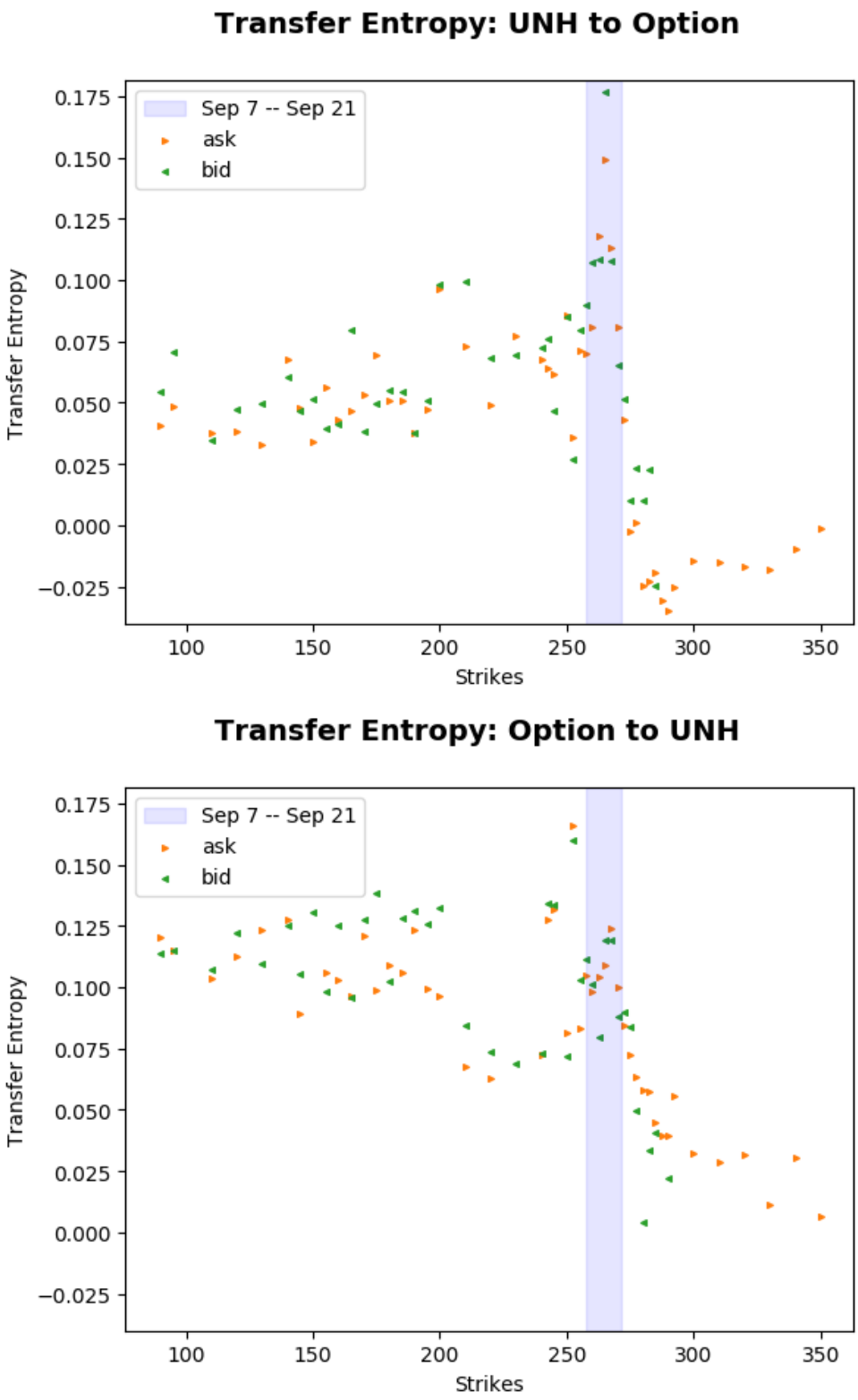

Figure 70 Transfer entropy between UNH and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

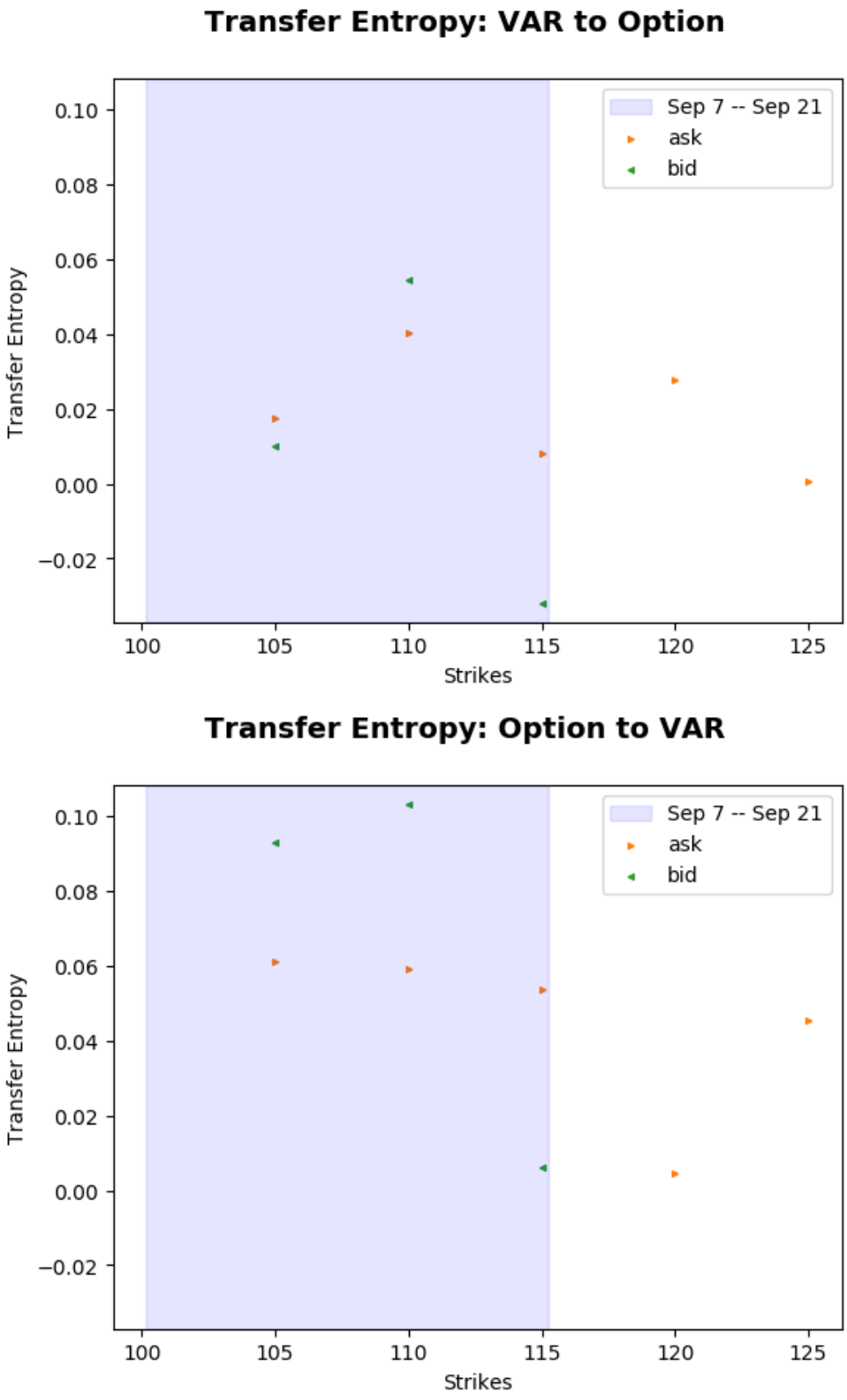

Figure 71 Transfer entropy between VAR and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

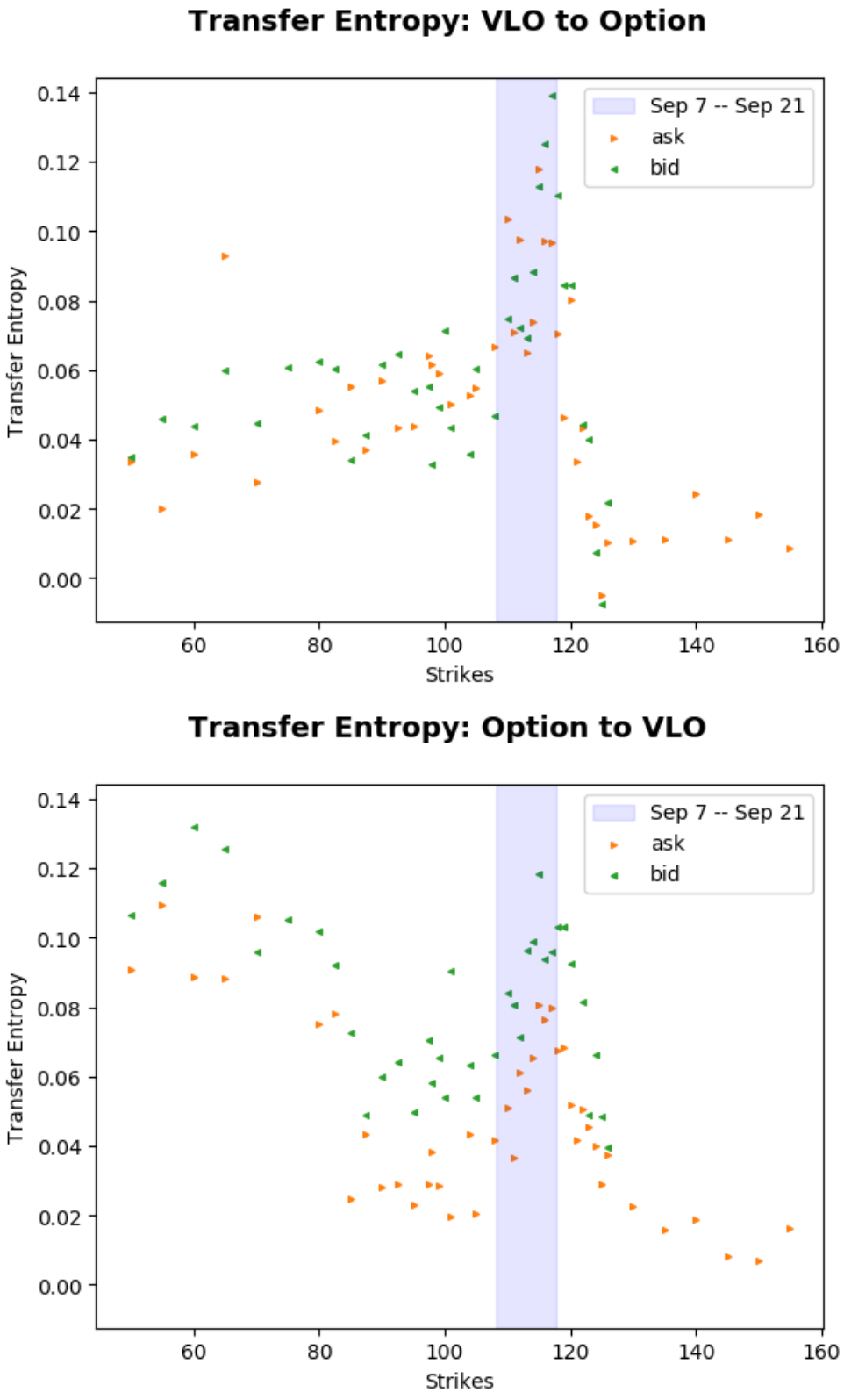

Figure 72 Transfer entropy between VLO and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

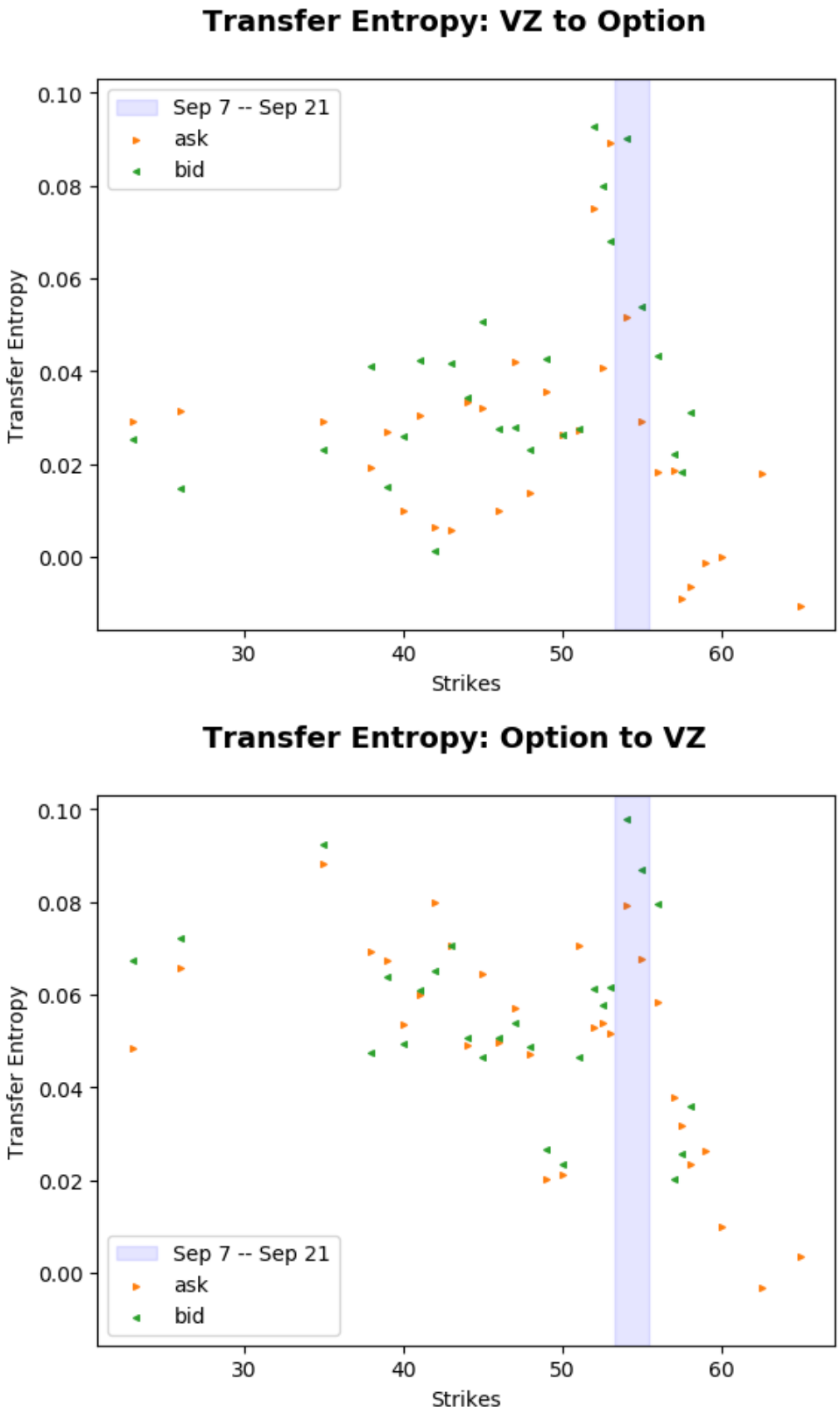

Figure 73 Transfer entropy between VZ and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 10 seconds.

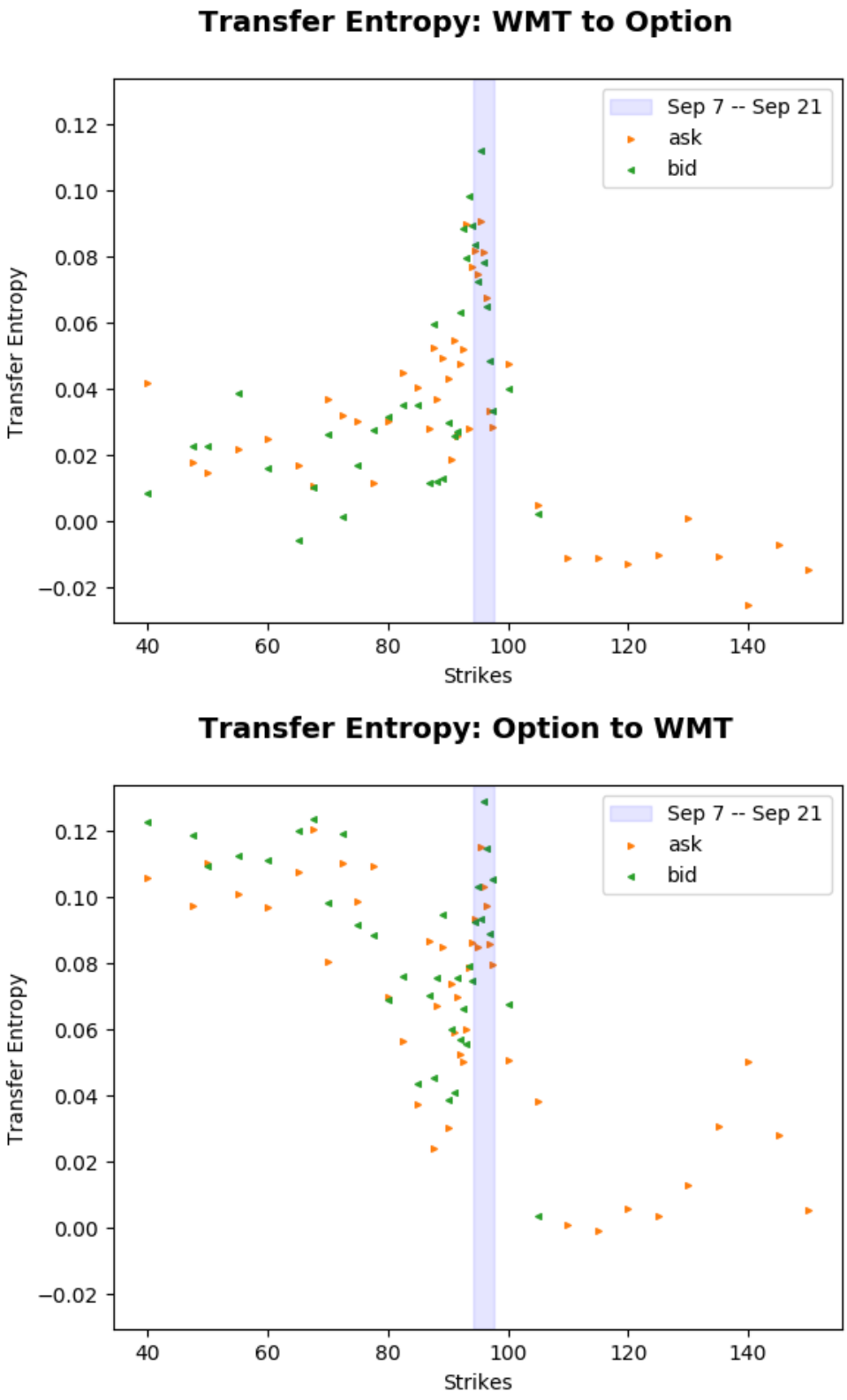

Figure 74 Transfer entropy between WMT and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

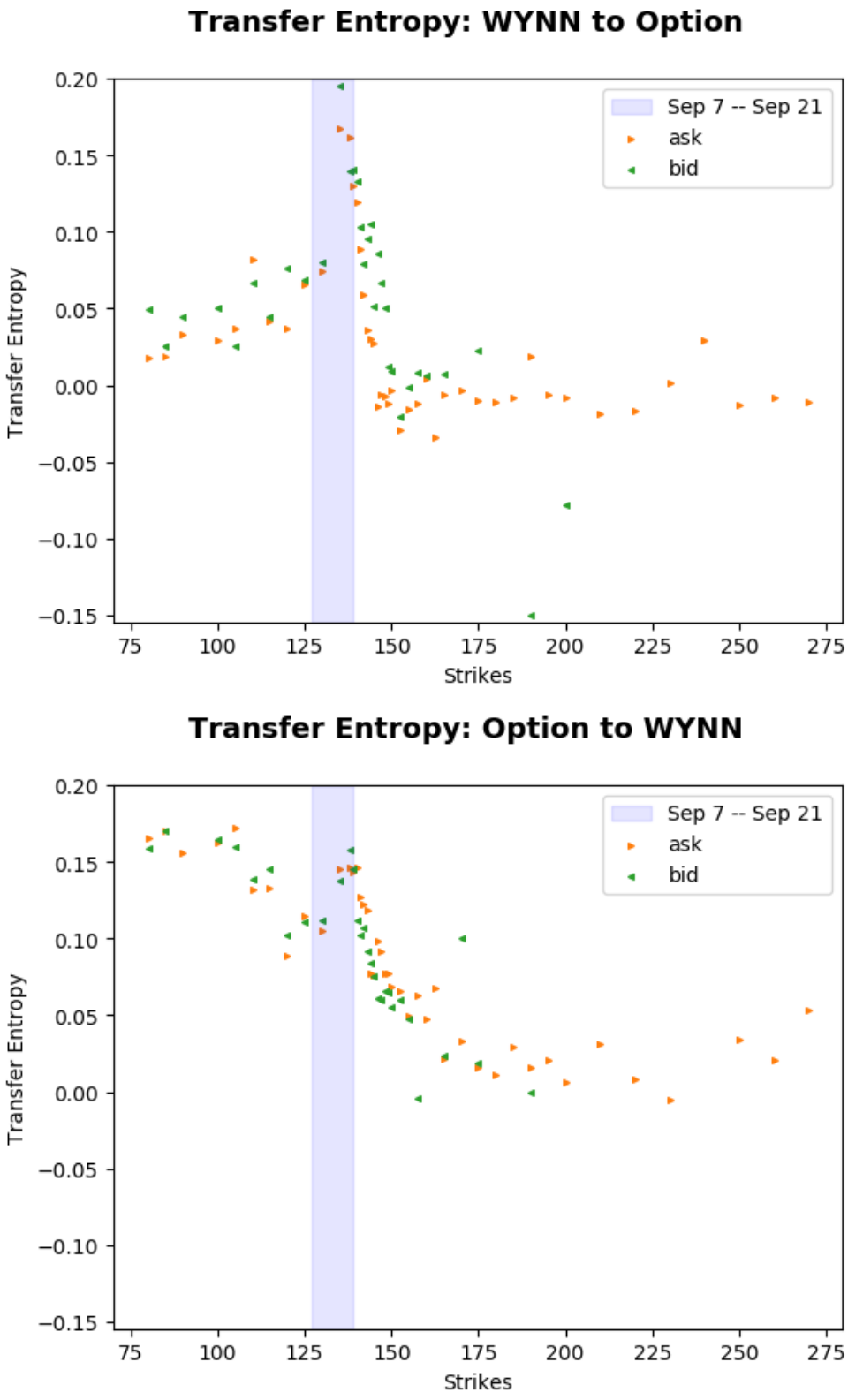

Figure 75 Transfer entropy between WYNN and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 12 seconds. Option to underlying timescale: 14 seconds.

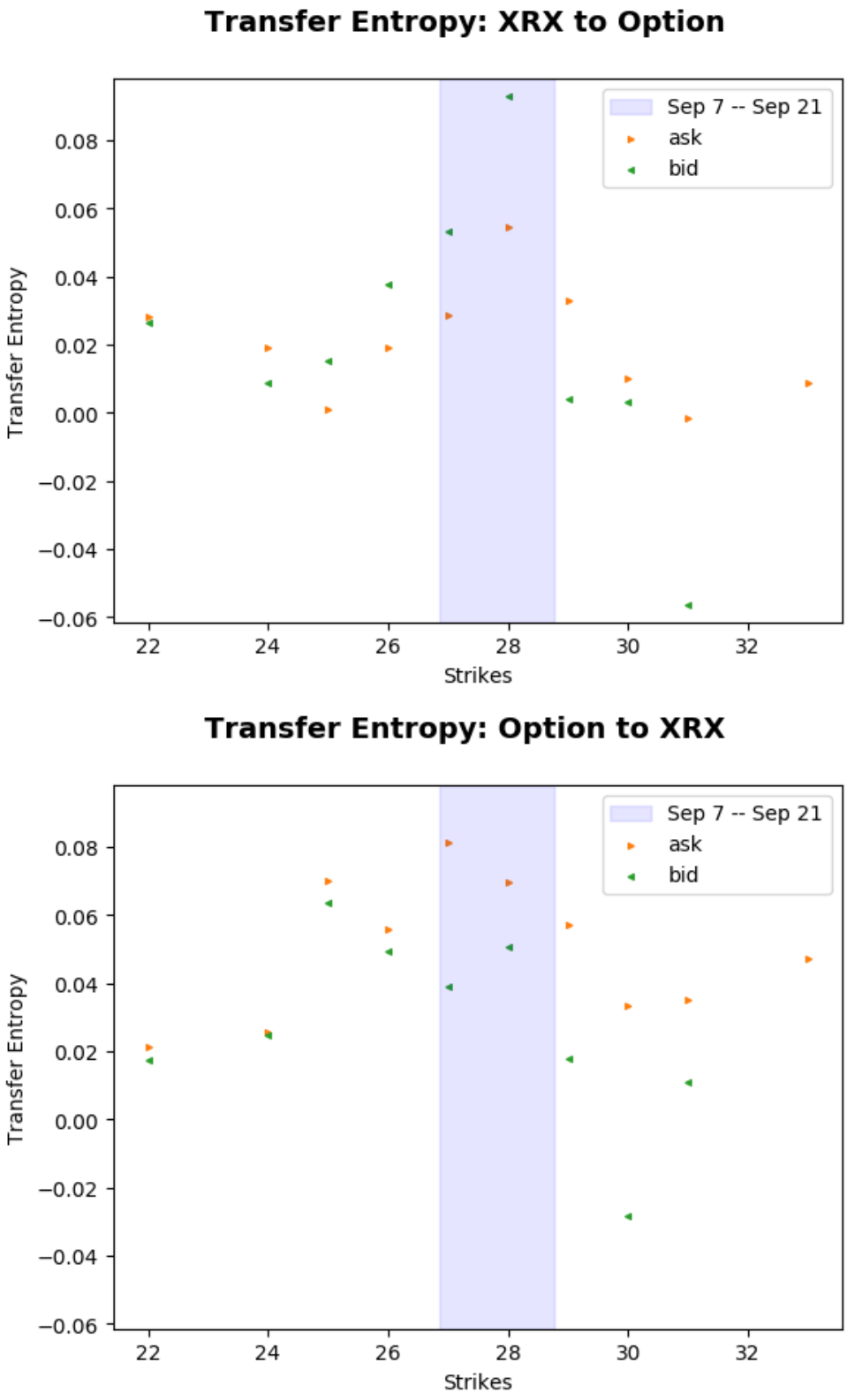

Figure 76 Transfer entropy between XRX and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.

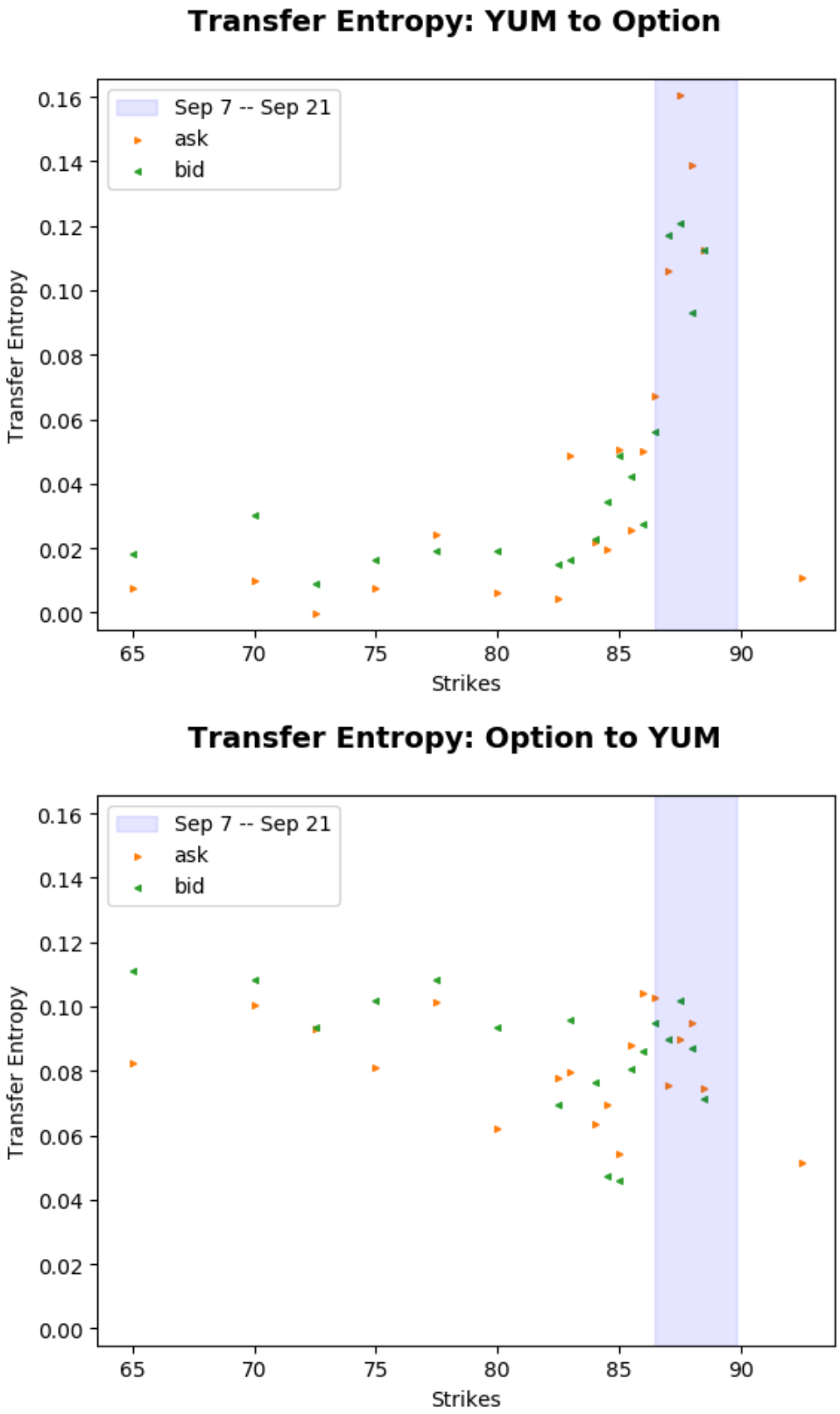

Figure 77 Transfer entropy between YUM and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 12 seconds.

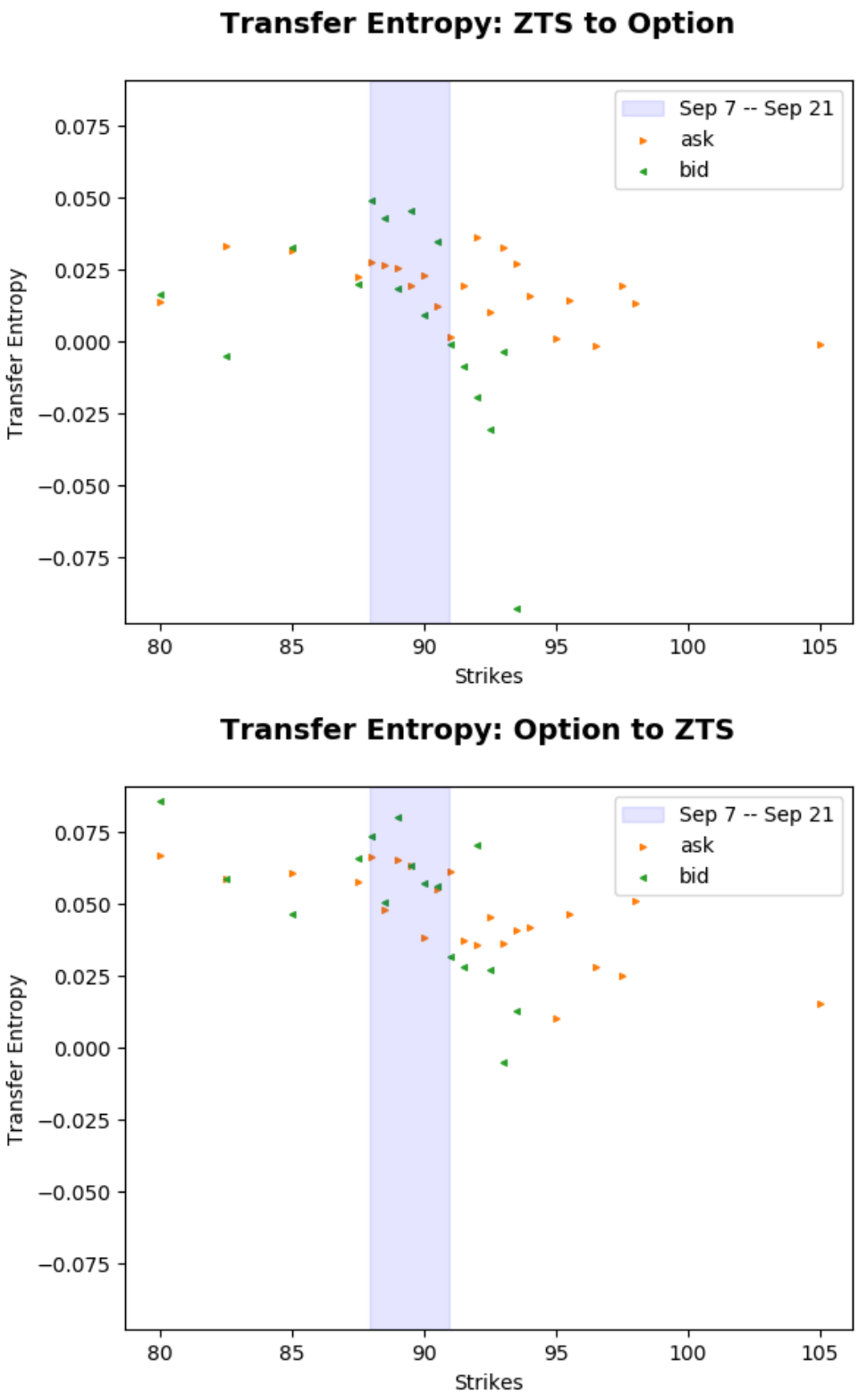

Figure 78 Transfer entropy between ZTS and its options. Kraskov estimator $k = 1$; lookback= 1. Underlying to option timescale: 14 seconds. Option to underlying timescale: 14 seconds.