Cojumps in China's spot and stock index futures markets☆

Hao Wang a, Mengqi Yue b, Hua Zhao b,⁎

a Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI 48824, USA
b School of Economics, Xiamen University, Fujian 361005, China

Abstract

This paper extracts high-frequency cojumps across China’s spot and futures markets to examine the characteristics of cojumps as well as their association with macroeconomic news announcements. The results indicate that there occur significant cojumps and that there is approximately a one-third probability of cojumping when jumps occur in the spot/futures market. The jump covariation attributable to cojump appears more erratic and less persistent than the realized covariance and significantly improves the covariance forecasts. Moreover, we find that electricity consumption, industrial profit, GDP, fixed investment, industrial value-added and retail sales announcements significantly impact cojumps.

© 2015 Elsevier B.V. All rights reserved.

Keywords:
Cojump
Macroeconomic announcement
Jump covariation
Spot
Futures

1. Introduction

Cojumps are joint jumps occurring contemporaneously across different assets or markets. From a risk measurement and portfolio management perspective, describing the characteristics and dynamics of the joint jumps between spot and futures prices and understanding their economic determinants is crucial. Using high-frequency data, this paper attempts to study high-frequency cojumps between China’s spot and futures markets. Specifically, we investigate characteristics of cojumps and explore the role of cojumps in covariance forecasting by extracting jump covariation from the realized covariance. We also examine how macroeconomic news announcements are associated with cojumps.

The earliest influential work for jump dates back to Merton (1976). Recently there have appeared a number of studies in the literature investigating jumps. For example, Huang (2007), Evans (2011) and Miao et al. (2013) emphasize the relationship of jumps to macro news announcements, and Yan (2011), Bollerslev and Todorov (2011) and Bollerslev et al. (forthcoming) focus on jump risk premiums. In addition, some attention has been paid to cojumps among diverse assets or markets. Dungey et al. (2006) examine contemporaneous bond jumps across different maturities and find that the middle of the curve is more likely

☆ We are grateful for helpful comments and suggestions from the editors Charles Cao, S. Ghon Rhee and anonymous referees. The work was supported by the Humanity and Social Science Foundation of Ministry of Education of China(15YJAZ0089), the Ministry of Education Key Laboratory of Econometrics (Xiamen University), Scientific Research Foundation for Returned Overseas Chinese Scholars, State Education Ministry, and the Program for New Century Excellent Talents in Higher Education of Fujian Province.

⁎ Corresponding author.

E-mail addresses: haowang@msu.edu (H. Wang), mengqiye0115@gmail.com (M. Yue), zhaohua@xmu.edu.cn (H. Zhao).

http://dx.doi.org/10.1016/j.pacfin.2015.10.002
0927-538X/© 2015 Elsevier B.V. All rights reserved.
to cojump and that the ends have greater potential for idiosyncratic jumping. Lahaye et al. (2009) investigate cojumps in exchange rates, stocks, and bonds by assessing the impact of macro surprises on jumps and cojumps. They find that nonfarm payroll and federal funds target announcements are the most important news across asset classes. Dungey and Hvzdyk (2012) apply the bivariate jump test developed by Jacob and Todorov (2009) to analyze the cojumping behavior of spot and futures prices in high frequency U.S. treasury data and find that cojumping occurs most frequently at shorter maturities and higher frequency samples.

Cojumps also have important implications in portfolio allocation. For example, Das and Uppal (2004) examine the impact of systemic jumps (cojumps) across international markets on optimal portfolios. They find the cost of ignoring systemic jumps is substantial for aggressive investors who hold heavily levered portfolios. Aït-Sahalia et al. (2009) investigate the consequences of common jumps among a set of assets in the context of dynamic portfolio choice. Bollerslev et al. (2008) study the cojumps between the forty large-cap U.S. stocks and the corresponding equiweighted index, and point out non-diversifiable cojumps are important for portfolio allocation. Through an extensive Monte Carlo study (Li et al. forthcoming) show that ignoring cojumps might lead to severe underestimation of tail risks.

This paper contributes to the literature in three aspects. First and foremost, it is the first study to explore high-frequency cojumps in China's spot and stock index futures markets. Existing literature on cojump subject is focused mainly on U.S. capital markets. Much research has been devoted to the role of jump in volatility forecasting and jump risk in asset pricing in China's stock market, as in Liao (2013) and Zhou and Zhu (2012), yet little attention is paid to high-frequency cojumps in China's financial markets. This paper enriches the cojump literature by investigating the case of China, which is the most important emerging economy and the second largest economy in the world according to GDP. Secondly, rather than using the cojump tests proposed by Jacob and Todorov (2009) we extract jumps from different asset classes with the nonparametric method recently developed by Bollerslev et al. (2013) and then define one cojump occurrence if the detected jumps in two markets appear contemporaneously. This method can identify cojumps and cojump timing directly, thus enabling us to detect high-frequency cojumps between China's stock market and futures market. A Monte Carlo simulation will be made to examine the effectiveness of the identification approach by considering two simulation schemes: a parametric jump diffusion model and a non-parametric model by resampling the real financial data. Last but not least, we extract jump covariation (JCov) from realized covariance and then construct the heterogeneous autoregressive (HAR) model proposed by Corsi (2009) and extended by Andersen et al. (2007b) to examine the benefits of cojumps in covariance and variance forecasting.

Based on the five-minute high-frequency data in China's spot and futures markets from April 16, 2010 to June 30, 2014, we find that there occur significant cojumps. The probabilities of cojump conditional on jumps in the two markets are 38.912% and 29.384%, respectively. Cojumps are more frequent from 11:00 am to 11:30 am than other trading hours. Moreover, the jump frequency in China's capital markets is far higher than in the U.S. and there exists an asymmetry between the positive and negative jumps inherent in the spot and futures markets in China. Specifically, China's spot market experienced 184 negative jumps and 294 positive jumps and the futures market experienced 280 negative and 353 positive ones. Thus, it can be observed that China's financial markets tend to jump upwards.

To demonstrate important gains of cojumps in covariance and variance forecasting, the jump covariation is extracted from the realized covariance. The realized covariance displays strong dependence, though the jump covariation attributable to cojump appears more erratic and less persistent than the realized covariance. When the jump covariation as a regressor is added into HAR model, the jump covariation is statistically significant and improves covariance and variance forecasts over the daily, weekly, and monthly forecasting horizons.

Employing logistic regression on rare events (the relogit model) to analyze how the 12 news surprises affect cojumps, we find that electricity consumption, GDP, fixed investment, retail sales, industrial value-added and industrial profit announcements significantly impact the probability of cojumps among the 12 macro news announcements. In addition, from the sign of the five significant macro news coefficients it can be seen that cojumps are more likely to occur when GDP, fixed investment and industrial value-added announcements contain positive surprises while electricity consumption, industrial profit and retail sales announcements contain negative surprises. The result also suggests that GDP, retail sales, and industrial value-added are more likely to produce cojumps than fixed investment, industrial profit and electricity consumption.

The remainder of this paper proceeds as follows. Section 2 introduces the theoretical framework of the Bollerslev et al. (2013) jump detection method and then defines one cojump occurrence according to the jump timing of diverse assets. In Section 3, the cojump detection method is investigated through Monte Carlo simulation. Section 4 illustrates the evidence of cojumps between the China Securities Index (CSI) 300 and the CSI 300 futures. In Section 5, we implement a HAR model specified on realized covariance and investigate the impact of cojumps on realized covariance. Empirical analysis of the relationship between macroeconomic news announcements and cojumps is presented in Section 6. Section 7 concludes the paper.

2. Cojump identification

We use the Bollerslev et al. (2013) method to identify jumps in asset prices. The method provides a flexible non-parametric estimation procedure for jumps, allows for very general dynamic dependencies in tails, and imposes essentially no restrictions on jump timing.

---


Please cite this article as: Wang, H., et al., Cojumps in China's spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002
on the continuous part of the price process. Assume that the log price process evolves with the generic semimartingale process,

\[ dp_t = \alpha_t dt + \sigma_t dW_t + \int_{\mathbb{R}} \mu(dt, dx), \]

(1)

where \( p_t \) is the log price, \( \alpha_t \) represents a continuous, locally bounded variation process, \( \sigma_t \) is a strictly positive stochastic volatility process, \( W_t \) denotes a standard Brownian motion, and \( \mu(dt, dx) \) is integer-valued random measurements that capture the jumps in \( p_t \) over time \( dt \) and size \( dx \).

Let \( n + 1 \) and \( T \) denote the number of equidistant price observations each day and the number of trading days, respectively. Then the discrete time grids for the whole sample are \( 0, \frac{1}{T}, \frac{2}{T}, \ldots, T \). The within-day high-frequency return over the corresponding discrete time-intervals \([\frac{i-1}{T}, \frac{i}{T}]\) is \( p_{n+i} - p_{n+i} \), and then we obtain \( nT \) high-frequency intraday returns \( r^n_i (s=1,2,\ldots,nT) \).

The realized variation (RV) for day \( t \),

\[ RV_t = \sum_{s=0}^{T} |r^n_s|^2, \]

(2)

provides a natural measure of the daily ex post volatility. Under regularity conditions, \( RV_t \) consistently converges to the total variation comprised of the “continuous” and “jump” components,

\[ RV_t \overset{P}{\to} \int_{t}^{t+1} (\sigma_s)^2 ds + \int_{t}^{t+1} \int_{\mathbb{R}} \mu(ds, dx). \]

(3)

The continuous component is the so-called integrated variance (IV), Barndorff-Nielsen and Shephard (2004a, 2006) have proved that under weak regularity conditions and \( n \to \infty \), the bipower variation (BV) consistently converges to IV even in the presence of jumps,

\[ BV_t \overset{P}{\to} \int_{t}^{t+1} (\sigma_s)^2 ds, \]

(4)

where the BV for each day, \( t = 1, \ldots, T \), is given by

\[ BV_t = \frac{n}{2} \sum_{s=0}^{T} |r^n_s||r^n_{s-1}|. \]

(5)

Separating the realized jumps from the continuous price moves is adaptively based on preliminary estimates of the stochastic volatility over the day together with the time-of-day (TOD) volatility pattern, namely as

\[ TOD_i = \frac{n}{nT} \sum_{s=1}^{T} I \left( \frac{n}{s} I \left( |r^n_s|^2 \leq \tau \right) \frac{BV_t \wedge RV_t}{n^{-\omega}} \right), \quad i = (t-1)n + i. \]

(6)

where \( i = 1, \ldots, n, I(\cdot) \) denotes the indicator function, \( \omega = 0.49 \), and \( \tau > 0 \) is the threshold parameter. In the empirical data analysis of Sections 5 and 6, we use \( \tau = 3.0 \), implying that the price increments beyond three standard deviations of a local estimator of the corresponding stochastic volatility are classified as jumps.\(^2\)

The TOD displays a U-shaped pattern in general. The \( i \)-th high-frequency jump for the \( t \)-th trading day is identified by the truncation approach,

\[ Jump_{t,i} = I \left( \frac{n}{s} I \left( |r^n_s|^2 \leq \tau \right) \frac{BV_t \wedge RV_t}{n^{-\omega}} \right) \frac{BV_{[s/n]} \wedge RV_{[s/n]} + TOD_{t-s[n/n]} n^{-\omega}}{n^{-\omega}}, \quad s = 1, \ldots, nT, \]

(7)

where \( t = [s/n] \) and \( s \in [n] \), with \( I(\cdot) \), \( \omega \) and \( \tau \) set to same values as discussed above.

Testing for contemporaneous jumps across different markets with high-frequency data is a research field which deserves further elaboration. Barndorff-Nielsen and Shephard (2004b) extend the concepts of bipower variation and realized variance to multivariate equivalent to define cojumps of daily frequency, but the corresponding multivariate jump test is not fully operational. Jacod and Todorov (2009) extend BNS work to test cojumps and their procedure is comprised of two cojump tests, though the two cojump tests may conflict in certain areas. Motivated by the need to reduce the impact of a large amount of idiosyncratic

---

\(^2\) Bollerslev et al. (2013) use a slightly smaller value of \( \tau = 2.5 \). We present the results based on a more conservative value of \( \tau = 3.0 \) which shows excellent size and power in our Monte Carlo simulation study of Section 3. However, we have repeated the analyses in Sections 5 and 6 under \( \tau = 2.5 \). The results are qualitatively the same.

---

Please cite this article as: Wang, H., et al., Cojumps in China’s spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002
jumps, Bollerslev et al. (2008) propose a test for common jumps that explicitly utilizes cross-covariance to identify non-diversifiable jumps of a large ensemble of individual stocks. However, the test lacks a reliable asymptotic null distribution and thus relies on simulation-based bootstrap procedures.

Our study uses the univariate analysis of asset-by-asset jump tests of Bollerslev et al. (2013) based on Eq. (7), and then defines one cojump occurrence according to the jump timing of the spot and futures. This method enables us not only to identify cojumps across spot and futures markets but also to ascertain cojump times. Specifically, after jumps in spot and futures markets are identified by Eq. (7), the occurrence of a cojump is determined when the detected jumps in the two markets appear exactly at the same time. Therefore, the cojump at time \( (t, i) \) between the spot and futures markets is defined as

\[
\text{Cojump}_{ij} = I(\text{Jump}_{t,i}^S \cap \text{Jump}_{t,i}^F), \quad t = 1, 2, ..., T, \quad i = 1, 2, ..., n,
\]

where \( I(\cdot) \) is the indicator function for a non zero argument, and \( \text{Jump}_{t,i}^S \) and \( \text{Jump}_{t,i}^F \) denote the events of jumps in spot and futures prices, respectively. As for the effectiveness of this procedure, the extensive Monte Carlo experiments conducted by Bollerslev et al. (2013) and Li et al. (forthcoming) demonstrate its good performance in terms of the estimation of the jump tail risks. In the subsequent section, we perform further experiments confirming its validity based on a different set of evaluation criteria.

3. Monte Carlo simulation

Before applying the jump detection procedure outlined in Section 2 to the real data, we examine its effectiveness via a Monte Carlo simulation that generates the bivariate price processes mimicking the joint dynamics of the real China’s spot and futures index markets. Let \( P_t^S \) and \( P_t^F \) be the prices for the spot and futures index, respectively. We consider two simulation schemes: the first scheme assumes a parametric jump diffusion model and the second scheme generates from a non-parametric model by resampling the real return data.

In the first simulation scheme, \( P_t^S \) and \( P_t^F \) follow the affine jump-diffusion models where the volatility processes is assumed to be the stochastic volatility (SV) models of Heston (1993),

\[
\begin{align*}
\ln P_t^S - \ln P_0^S &= \int_0^t \sqrt{V_t^S} dW_t^S + \sum_{s \in N_{t}^S} Z_{t,s}^{\text{idio.S}} + \sum_{s \in N_{t}^S} Z_{t,s}^{\text{co.S}}, \\
\ln P_t^F - \ln P_0^F &= \int_0^t \sqrt{V_t^F} dW_t^F + \sum_{s \in N_{t}^F} Z_{t,s}^{\text{co.F}} + \sum_{s \in N_{t}^F} Z_{t,s}^{\text{idio.F}}, \\
V_t^S - V_0^S &= \kappa^S \int_0^t (\theta^S - V_s^S) ds + \alpha^S \int_0^t \sqrt{V_s^S} d\beta_s^S, \\
V_t^F - V_0^F &= \kappa^F \int_0^t (\theta^F - V_s^F) ds + \alpha^F \int_0^t \sqrt{V_s^F} d\beta_s^F,
\end{align*}
\]

where \( W_t^S, W_t^F, \beta_t^S, \) and \( \beta_t^F \) are standard Brownian motions, \( N_{t}^{\text{co.S}}, N_{t}^{\text{co.F}} \) and \( N_{t}^{\text{idio.S}}, N_{t}^{\text{idio.F}} \) are Poisson processes with constant intensities \( \lambda_{\text{co.S}}, \lambda_{\text{co.F}} \) and \( \lambda_{\text{idio.S}}, \lambda_{\text{idio.F}} \), and \( Z_{t,s}^{\text{co.S}}, Z_{t,s}^{\text{co.F}} \) and \( Z_{t,s}^{\text{idio.S}}, Z_{t,s}^{\text{idio.F}} \) are the jump sizes. We fix the hyperparameters in Eq. (9) at values that reflect the behavior of the real return processes. Specifically, \( \theta^S \) and \( \theta^F \) control the unconditional mean of \( V_t^S \) and \( V_t^F \), respectively. We set \( \theta^S = 1.0871 \) and \( \theta^F = 1.0984 \) so that the annual volatilities are \( \sqrt{252} \times 1.0871 = 16.55\% \) for the spot and \( \sqrt{252} \times 1.0984 = 16.64\% \) for the futures, both of which match the average continuous variation as computed from the real data. The parameters \( \kappa^S, \kappa^F, \alpha^S \) and \( \alpha^F \) are calibrated as follows. Using a time-discretization of the SV processes in Eq. (9) with an increase of \( \Delta = 1 \) day, we have

\[
V_{t+1} - V_t \approx \kappa^S (\theta^S - V_t) + \alpha^S \sqrt{V_t} \xi_{t+1}, \quad V_0 \sim N(0, 1),
\]

which implies that \( \kappa^S \) and \( \alpha^S \) might be estimated via fitting a linear regression model to \( V_{t+1} \) that is approximated by the daily realized continuous variation computed from the real data. Applying a regression analysis, we estimate the persistent parameters as \( \kappa^S = 0.1209 \) and \( \kappa^F = 0.1756 \) and the volatility-of-volatility parameters as \( \alpha^S = 0.3439 \) and \( \alpha^F = 0.4443 \).

As for the compound Poisson jumps, our empirical analysis on the real data in Section 4 provides some guidance. From the jump counts reported in Tables 3 and 4, we fix the cojump intensity at \( \lambda_{\text{co}} = 186/1018 = 0.1827 \) and the two idiosyncratic jump intensities at \( \lambda^S = (478 - 186)/1018 = 0.2826 \) and \( \lambda^F = (633 - 186)/1018 = 0.4391 \), implying an average unconditional cojump intensity of approximately one cojump every week and an overall jump intensity of approximately one jump every other day. The distribution of the cojump sizes, \( (Z_{t,s}^{\text{co.S}}, Z_{t,s}^{\text{co.F}}), s = 1, 2, ..., \) are assumed to be i.i.d. with a cumulative distribution function (CDF),

\[
G(x^S, x^F) = C_F(x^S) \cdot F(x^F),
\]
where $F(\cdot)$ refers to the CDF of a truncated at 0.2 zero-mean normal variable with standard deviation equal to 0.65 and the Gumbel–Hougaard (logistic) copula,

$$
C_\theta(u_0, u_1) = \exp\left(-\left[(-\log u_0)^{1/\theta} + (-\log u_1)^{1/\theta}\right]^\theta\right), \quad u_0, u_1 \in [0, 1],
$$

introduces tail dependence among jump tails (Bollerslev et al., 2013; Li et al., forthcoming). We set $\theta = 14.2$ so that the tail-dependence coefficient equals to a moderate amount of 0.5 between the two jumps. The distributions of the two idiosyncratic jumps are set as two independent zero-mean normal variables with standard deviations equal to 0.49 and 0.51 and truncated at 0.2 respectively. Finally, we let $\text{corr}(B_i^u, B_i^v) = 0.8050$ and $\text{corr}(W_i^u, W_i^v) = 0.9614$ which are the same as the realized volatility and return correlations computed from the real data. We assume $(W_i^u, W_i^v)$ and $(B_i^u, B_i^v)$ are independent. We use the Euler discretization scheme with an increment of 1 min per tick on the Euler clock to generate prices from Eq. (9). After discarding the burn-in period, we use the 5-minute returns to detect cojumps.

In the second simulations scheme, instead of specifying the parametric SV model in Eq. (9), we use the realized continuous variation computed from the real data as $V_i^u$ and $V_i^v$ for the price processes in Eq. (9). This assumes that the volatility is locally constant within each day. The sequence of $(V_i^u, V_i^v)$ then enables us to generate the diffusion part of the price process by sampling from a bivariate Brownian motions $(W_i^u, W_i^v)$ with the correlation $\text{corr}(W_i^u, W_i^v) = 0.9614$ as before. For the jump part, we randomly assign the observed cojumps and idiosyncratic jumps into the rest of columns in Panel A of Table 1. Under either simulation scheme, the detected cojump correlation always agrees with the true cojump correlation to the second decimal places; the relative contributions of detected cojumps to the total variation among all true cojumps. Thus PSD and PTD might be interpreted as the size and power in a traditional statistical hypothesis-test framework. Additionally, we compare the realized correlation and variances of detected cojumps with those from the true cojumps to assess if these important realized measures are distorted by the fact that cojumps are detected rather than directly observed. The realized correlation and variance of cojumps are relevant because they are the two components determining the realized jump covariation that is found useful in predicting of futures volatility and covariances in Section 5.1. We simulate 1000 replications of 1018 “days” of bivariate price processes and keep the prices for every 5-minute. The performance measures are then computed for each replication and the results are summarized in Table 1.

The PSD column suggests that the test generally guards well against the spurious detection of cojumps with a size of around 0.1%. The PTD column suggests that the test also enjoys a reasonable power in detecting true cojumps. Note that our average jump size adopted in the simulation is approximately 0.6%, which is substantially smaller than the choice of 0.91% in Bollerslev et al. (2013) and is indeed comparable to the small jump scenario considered in Andersen et al. (2007c). Even for such small jump sizes, the cojump test procedure is able to deliver a power of around 56% to 62%, which are in line with the power reported by Andersen et al. (2007c).3 Furthermore, the realized measures are remarkably close between the detected and the true cojumps as is evident in the rest of columns in Panel A of Table 1. Under either simulation scheme, the detected cojump correlation always agrees with the true cojump correlation to the second decimal places; the relative contributions of detected cojumps to the total variation agree with those of the true cojumps to the second decimal places as well. Overall, Panel A shows the cojump test procedure can reliably identify cojumps and does not compromise the important properties of the cojumps. The procedure might allow us to treat cojumps as observable and analyze them directly in the subsequent sections.

4. Evidence of intraday cojumps

4.1. Data

This paper applies the intraday jump detection procedures to the CSI 300 and CSI 300 futures contracts over a three-year period from April 16, 2010 to June 30, 2014. The CSI 300 was created on April 8, 2005 and includes 300 large-cap and actively-traded A-share stocks listed on the Shanghai or Shenzhen Stock Exchanges whose total capitalization accounts for about 60 percent of the market capitalization of each of those exchanges. The CSI 300 futures contract was launched on April 16, 2010 by the China Financial Futures Exchange. The contract value is determined by multiplying the CSI 300 index point value by 300 RMB. There are four contract expirations — current-month, next-month, next-quarter and next-two-quarter — in China’s stock index futures market. We use the current-month futures contract prices to form continuous futures price series for its active trade and substantial trade volume by comparison with other futures contracts. The data were obtained from the China Stock Market and Accounting Research (CSMAR) high-frequency database. Following Andersen et al. (2003), Huang and Tauchen (2005), and Bollerslev et al. (2013), we chose a five-minute sampling frequency. The study utilizes trading hours from 9:30 am to 11:30 am and from 1:00 pm to 3:00 pm4 and excludes overnight returns and returns during the one-hour-and-a-

---

3 The small jump scenario in Andersen et al. (2007c) assumes that the jump size is normally distributed with mean zero and standard deviation 0.5%. Their Table 1 reports that jump tests have a power in the range of 44% and 65% under scenario, which is labeled as scenario #3 there. Although the exact specification differs between their simulation and ours, we take their power numbers as a general magnitude of powers we might expect for jump tests for small jumps.

4 The time is represented by China Standard Time, namely GMT +8 h.
half lunch break in our analysis for both assets, although the futures contract is traded from 9:15 am to 11:30 am and from 1:00 pm to 3:15 pm, 15 min prior to the stock market opening and 15 min after the stock market closing. The study uses data from 1018 trading days, each consisting of 48 five-minute returns. Thus there are 48864 data entries.

Fig. 1 presents five-minute returns of the whole sample in the stock and futures markets respectively. Some surges are distinctly illustrated in the price moving process of the two assets. The majority of five-minute returns of the two assets are between −1% and 1% and the tendencies of five-minute returns in the two markets are greatly similar which reveals close correlation in price behavior of the two assets.

Table 2 summarizes the statistics of five-minute returns for both assets. The mean of the spot returns is positive while futures returns are, on average, negative. Meanwhile, both spot and futures returns are similarly volatile with the standard error of CSI 300 stock index futures returns a bit larger than that of the CSI 300 stock index. Both asset return distributions are positively skewed and exhibit excessive kurtosis, though the futures kurtosis is about twice the size of the spot kurtosis.

4.2. Jump characteristics

Fig. 2 provides a bird’s eye view of the two time series of identified jumps in the two assets, showing that the two markets have certain jumping behaviors in common. Firstly, there exists no significant difference in the jumping frequency of the two assets, which means that the stock market seems to jump almost as frequently as the futures market does. Secondly, the tendencies of jump sizes of the two assets are very similar. For example, comparing the tendency of jump size in the two markets from October 2012 to February 2013 it is easily seen that the shapes of the curves of jump sizes in the two plots are greatly similar and that extremely large jumps occur nearly simultaneously. From observing the above-two common features we conclude that there is close comovement in jumping behavior between the two markets.

Fig. 3 displays the jump frequency of the two markets at different times during trading hours. The upper plot represents the total number of jumps in the CSI 300 index in five-minute intervals from 9:30 am to 3:00 pm during a period of about three and a half years. It also shows that most jumps of 32 occur in the first 5 mins compared with other periods because investors respond rapidly and intensively to information they have collected from 3:00 pm the previous day to 9:30 am during which time the exchange is closed. However, over the same period, the futures market experienced only ten jumps during the first 5 mins, as represented in the lower plot of Fig. 3. This is not surprising because overnight information has already been accounted for in the price at 9:15 am when the stock index futures exchange opens following the overnight closing. For both markets, on average, jump intensity in the morning from 9:30 am to 11:30 am is higher than in the afternoon from 1:00 pm to 3:00 pm, probably due to the greater-frequency of morning news announcements by the Chinese government. Additionally, the two graphs show slightly higher jump intensity around 11:30 am, just prior to the noon closing of both markets. This is probably the result of the cumulative information effect in the morning and of investors’ expectations for afternoon trading.

Table 3 presents a statistical description of significant jumps in the spot and futures markets. In Panel A, compared with the results obtained in the developed U.S. capital markets (see Lahaye et al., 2009), jump frequency in China’s capital markets is far higher than that found in the U.S. This finding is not surprising given that investors in emerging markets such as China’s are less experienced and have less knowledge of investment, thus leading to diminished rationality of investors. The statistics, $P(jump)$, $E(\lvert jump size\rvert | jump)$ and $\sqrt{Var(\lvert jump size\rvert | jump)}$ which are greater in the futures market than in the stock market suggest that the CSI 300 stock index futures jumps a bit more frequently and in larger size than the CSI 300 stock index does. This is probably due to the fact that China’s stock index futures market is less mature and more volatile, which makes futures prices more sensitive to surprises caused by macro news shocks.

Panel B and panel C consider the statistical characteristics of positive and negative jumps. As presented in panel B, the number of positive jumps in the spot market is modestly more than that in the futures market while the mean and volatility of the jump sizes in the spot market are significantly lower than those in the futures market. The statistics of negative jumps presented in Panel C exhibit characteristics that are dissimilar to those of positive jumps. For instance, the mean and volatility of the negative jump sizes in the two markets are similar, while the number of the jumps in the futures market is much higher than in the spot market. Additionally, there is an asymmetry between the positive and negative jumps. Both markets experience positive jumps much more frequently than they do negative ones with the number of positive jumps being 294 and the number of negative

Please cite this article as: Wang, H., et al., Cojumps in China’s spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002
jumps being 184 for the spot market, while 353 positive and 280 negative for the futures market. Consequently, China's financial markets tend to jump upwards.

4.3. Market comovement: analysis of cojumps

This subsection characterizes cojumps across spot and futures markets. Cojumps in the two markets are identified by Eq. (8). Table 4 summarizes the descriptive statistics of cojumps in the two markets. We denote the probability of a cojump simply as \( P(\text{cojump}) \) which shows that the observed proportion of cojumps in the spot and futures markets is 0.381%. To reveal the comovement structure of jumps in the two markets, we calculate \( P(\text{cojump}|\text{jump}) \), the probability of cojumps conditional on jumps, which is defined as the probability of a cojump occurring in both of the spot and futures markets given that a jump occurs in one of them. The last two columns of Table 4 show that the probability of cojumps conditional on jumps in the CSI 300 stock index is 38.912% and 29.384% for the futures, which indicates that there is approximately a one-third probability of cojumping when jumps occur in the spot/futures market. The large probability of cojumps conditional on jumps demonstrates the close comovement in jumping behavior of the two markets, which suggests that the two markets may react to news surprises contemporaneously when news arrives.

Fig. 4 illustrates cojump intensity of every 5 mins on trading days over the sample period. On average, cojumps are more frequent from 11:00 am to 11:30 am than other trading hours, which is consistent with the characteristics of the jumps described in the previous subsection. The reason for the phenomenon is that macro news is announced more frequently in the morning than in the afternoon. Certainly, in addition to the macro news in our sample, factors such as liquidity shocks may be associated with cojumps in the two asset classes.

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary statistics of five-minute returns in spot and futures markets.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Spot returns</td>
</tr>
<tr>
<td>Futures returns</td>
</tr>
</tbody>
</table>

Please cite this article as: Wang, H., et al., Cojumps in China’s spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002
5. The benefit of cojumps in forecasting realized covariance and variance

5.1. HAR-RCov-JCov and HAR-RV-JCov models

Covariance and variance play important roles in Beta pricing, portfolio optimization, and risk management, and we investigate whether cojumps result in covariance and variance forecast improvement. With high-frequency data, Barndorff-Nielsen and Shephard (2004b) obtain an ex post measure of daily covariation without knowing the underlying data generating process, that is realized covariance. Recently Jin and Maheu (2013) have provided a joint-return realized-covariance model for density forecasting. Empirical evidence of strong dependence of realized covariance and variance has, inter alia, already been discussed in Andersen et al. (2003, 2007b) and Chiriac and Voev (2011). Much evidence, together with our empirical results below, points to the idea that realized covariance should be described by models allowing for slowly decaying autocorrelation and possibly long memory. The heterogeneous autoregressive (HAR) model of Corsi (2009), a basically simple regression compared with more complicated models such as the ARFIMA model, can capture the strong dependence of volatility. Andersen et al. (2007b) extend the HAR model by including jump component in model. Similarly, we incorporate cojump component into HAR model to examine the role and impact of cojumps in forecasting realized covariance and variance.

Realized covariance for day $t$ between spot returns $r_{S}^{n}$ and futures returns $r_{F}^{n}$ is defined as

$$ RCov_t = \sum_{t-i-1}^{t} r_{S}^{n} r_{F}^{n}, \quad t = 1, 2, ..., T, $$

where $r_{j}^{n}(j=S,F)$ is defined in Section 2. The HAR-RCov model for realized covariance and HAR-RV model for realized variance are represented as

$$ RCov_{t,h} = \alpha_0 + \alpha_d RCov_t + \alpha_w RCov_{t-5} + \alpha_m RCov_{t-22} + \varepsilon_{t,h}, $$

$$ RV_{t,h} = \alpha_0 + \alpha_d RV_t + \alpha_w RV_{t-5} + \alpha_m RV_{t-22} + \varepsilon_{t,h}. $$

Fig. 2. Jumps in spot and futures markets.
Fig. 3. Jump intensity in spot and futures markets. Note: The trading hours of China’s stock market are from 9:30 am to 11:30 am and from 1:00 pm to 3:00 pm, and the trading hours of China’s futures market are from 9:15 am to 11:30 am and from 1:00 pm to 3:15 pm. We plot jump intensity in five-minute intervals during the common trading hours of the two markets, from 9:30 am to 11:30 am and from 1:00 pm to 3:00 pm. In addition, to make the figure continuous we omit the period 11:30 am through 1:00 pm on the horizon axis, and the next time point after 11:30 am is 1:05 pm.

Table 3
Statistical description of significant jumps.

<table>
<thead>
<tr>
<th></th>
<th>Spot</th>
<th>Futures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: total jumps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#obs</td>
<td>48,864</td>
<td>48,864</td>
</tr>
<tr>
<td>#days</td>
<td>1018</td>
<td>1018</td>
</tr>
<tr>
<td>#jumps</td>
<td>478</td>
<td>633</td>
</tr>
<tr>
<td>( E(#\text{jumps}</td>
<td>\text{jumpday}) )</td>
<td>1.219</td>
</tr>
<tr>
<td>( P(\text{jump})) (%)</td>
<td>0.978</td>
<td>1.295</td>
</tr>
<tr>
<td>( E(</td>
<td>\text{jumpsize}</td>
<td></td>
</tr>
<tr>
<td>( \sqrt{\text{Var}(</td>
<td>\text{jumpsize}</td>
<td></td>
</tr>
<tr>
<td>Panel B: positive jumps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#jumps &gt; 0</td>
<td>294</td>
<td>353</td>
</tr>
<tr>
<td>( P(\text{jump} &gt; 0)) (%)</td>
<td>0.602</td>
<td>0.722</td>
</tr>
<tr>
<td>( E(</td>
<td>\text{jumpsize}</td>
<td></td>
</tr>
<tr>
<td>( \sqrt{\text{Var}(</td>
<td>\text{jumpsize}</td>
<td></td>
</tr>
<tr>
<td>Panel C: negative jumps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#jumps &lt; 0</td>
<td>184</td>
<td>280</td>
</tr>
<tr>
<td>( P(\text{jump} &lt; 0)) (%)</td>
<td>0.377</td>
<td>0.573</td>
</tr>
<tr>
<td>( E(</td>
<td>\text{jumpsize}</td>
<td></td>
</tr>
<tr>
<td>( \sqrt{\text{Var}(</td>
<td>\text{jumpsize}</td>
<td></td>
</tr>
</tbody>
</table>

Note: Panel A displays the number of observations (#obs), the number of sample days (#days), the total number of jumps (#jumps), the average number of jumps per jump day \( E(\#\text{jumps}|\text{jumpday}) = \#\text{jumps}/\#\text{jump days} \), jump ratio (in%) over the sample observations \( P(\text{jump}) = 100 \times (\#\text{jumps}/\#\text{obs}) \), and absolute mean size and standard deviation of jumps \( E(|\text{jumpsize}||\text{jump}) \) and \( \sqrt{\text{Var}(|\text{jumpsize}||\text{jump})} \). Panel B and Panel C exhibit statistical characteristics of the positive and negative jumps.

Please cite this article as: Wang, H., et al., Cojumps in China’s spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002
where $Y_{t+h} \equiv h^{-1}(Y_{t+1} + Y_{t+2} + \ldots + Y_{t+h})$. $Y$ is $RCov$ or $RV$. When $h = 1, 5, 22$, $RCov_{t+h}$ ($RV_{t+h}$) signifies daily, weekly and monthly realized covariance (variance). The coefficients $\alpha_d$, $\alpha_w$ and $\alpha_m$ capture the impact of short-term, medium-term, and long-term covariation (variation) on future realized covariance (variance).

In order to explore the impact of cojumps on realized covariance and variance, we extract jump covariation on day $t$, the component attributable to cojumps, from the total realized covariance, and define it as follows:

$$JCov_t = \sum_{i=1}^{n} Jump_{i,t}^S Jump_{i,t}^F, \quad t = 1, 2, \ldots, T,$$

where $Jump_{i,t}^S$ and $Jump_{i,t}^F$ denote jumps in spot returns and futures returns, respectively, as in Eq. (7).

As in the ABD model of Andersen et al. (2007b), we define the HAR-RCov-JCov and HAR-RV-JCov models as

$$RCov_{t+h} = \alpha_0 + \alpha_d RCov_t + \alpha_w RCov_{t-5} + \alpha_m RCov_{t-22} + \alpha_J JCov_t + \epsilon_{t+h},$$

$$RV_{t+h} = \alpha_0 + \alpha_d RV_t + \alpha_w RV_{t-5} + \alpha_m RV_{t-22} + \alpha_J JCov_t + \epsilon_{t+h},$$

where $\alpha_J$ signifies the impact of cojumps on realized covariance or variance.

### Table 4

| #cojumps | $P$(cojump)% | $P^d$(cojump|jump)% | $P^f$(cojump|jump)% |
|----------|--------------|----------------|----------------|
| 186      | 0.381        | 38.912         | 29.384         |

Note: The table displays, from left to right, the number of cojumps in the spot and futures markets, the cojump proportion over the sample observations ($P$(cojump) = #cojumps/#obs), the ratio of cojumps to jumps ($P^d$(cojump|jump) = #cojumps/#jumps) in the spot market, and the ratio ($P^f$(cojump|jump) = #cojumps/#jumps) in the futures market.

Fig. 4. Cojump intensity. Note: We plot cojump intensity in five-minute intervals during the common trading hours of the two markets, from 9:30 am to 11:30 am and from 1:00 pm to 3:00 pm. To make the figure continuous we omit the period 11:30 am through 1:00 pm on the horizon axis, and the next time point after 11:30 am is 1:05 pm.

Please cite this article as: Wang, H., et al., Cojumps in China’s spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002
5.2. Forecasting realized covariance and variance

To illustrate the characteristics of realized covariance and cojumps, we plot Fig. 5 including graphs of the realized covariance, jump covariation, and realized variances for spot and futures returns from April 16, 2010 through June 28, 2014. The realized co-covariance and variances displays strong dependency, that is, both high volatility and low volatility are persistent. The jump covariation appears noticeably more erratic and less persistent than the total realized covariance, but when the realized covariance

![Realized covariance](5.2. Forecasting realized covariance and variance)

5.2. Forecasting realized covariance and variance

To illustrate the characteristics of realized covariance and cojumps, we plot Fig. 5 including graphs of the realized covariance, jump covariation, and realized variances for spot and futures returns from April 16, 2010 through June 28, 2014. The realized covariance and variances displays strong dependency, that is, both high volatility and low volatility are persistent. The jump covariation appears noticeably more erratic and less persistent than the total realized covariance, but when the realized covariance

### Table 5

<table>
<thead>
<tr>
<th>HAR–RCov–JCov</th>
<th>One day</th>
<th>One week</th>
<th>One month</th>
<th>HAR–RV–JCov</th>
<th>One day</th>
<th>One week</th>
<th>One month</th>
<th>HAR–RV–JCov</th>
<th>One day</th>
<th>One week</th>
<th>One month</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>0.186**</td>
<td>0.257**</td>
<td>0.422**</td>
<td>0.145**</td>
<td>0.228**</td>
<td>0.406**</td>
<td></td>
<td>0.259**</td>
<td>0.366**</td>
<td>0.638***</td>
<td></td>
</tr>
<tr>
<td>$\alpha_d$</td>
<td>0.304*</td>
<td>0.185**</td>
<td>0.090***</td>
<td>0.199**</td>
<td>0.151**</td>
<td>0.081**</td>
<td></td>
<td>0.282**</td>
<td>0.183**</td>
<td>0.084***</td>
<td></td>
</tr>
<tr>
<td>(1.867)</td>
<td>(3.031)</td>
<td>(3.698)</td>
<td></td>
<td>(2.697)</td>
<td>(3.433)</td>
<td>(3.413)</td>
<td></td>
<td>(2.169)</td>
<td>(2.781)</td>
<td>(3.341)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_w$</td>
<td>0.265***</td>
<td>0.187***</td>
<td>0.157***</td>
<td>0.347***</td>
<td>0.249***</td>
<td>0.183***</td>
<td></td>
<td>0.330***</td>
<td>0.248**</td>
<td>0.206***</td>
<td></td>
</tr>
<tr>
<td>(3.053)</td>
<td>(2.765)</td>
<td>(5.038)</td>
<td></td>
<td>(5.230)</td>
<td>(3.889)</td>
<td>(4.348)</td>
<td></td>
<td>(4.130)</td>
<td>(5.117)</td>
<td>(5.542)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_m$</td>
<td>0.268**</td>
<td>0.355**</td>
<td>0.275***</td>
<td>0.274***</td>
<td>0.349***</td>
<td>0.295***</td>
<td></td>
<td>0.233**</td>
<td>0.305**</td>
<td>0.222***</td>
<td></td>
</tr>
<tr>
<td>$\alpha_J$</td>
<td>-0.498*</td>
<td>-0.267**</td>
<td>-0.151**</td>
<td>-0.115***</td>
<td>-0.068**</td>
<td>-0.044**</td>
<td></td>
<td>-0.638**</td>
<td>-0.334**</td>
<td>-0.163***</td>
<td></td>
</tr>
<tr>
<td>(-1.822)</td>
<td>(-2.696)</td>
<td>(-2.933)</td>
<td></td>
<td>(-2.840)</td>
<td>(-2.650)</td>
<td>(-1.898)</td>
<td></td>
<td>(-1.964)</td>
<td>(-2.605)</td>
<td>(-1.749)</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.211</td>
<td>0.322</td>
<td>0.308</td>
<td>0.277</td>
<td>0.392</td>
<td>0.333</td>
<td></td>
<td>0.269</td>
<td>0.389</td>
<td>0.324</td>
<td></td>
</tr>
<tr>
<td>$R^2_{HAR}$</td>
<td>0.195</td>
<td>0.321</td>
<td>0.302</td>
<td>0.271</td>
<td>0.388</td>
<td>0.330</td>
<td></td>
<td>0.256</td>
<td>0.382</td>
<td>0.322</td>
<td></td>
</tr>
</tbody>
</table>

Note: The models are estimated with $h = 1$ (one day), $h = 5$ (one week) and $h = 22$ (one month). The Newey–West adjusted t-statistics are reported in parentheses. $HAR–RV–JCov$ and $HAR–RV^2–JCov$ signify the $HAR–RV–JCov$ models of spot and futures markets, and reflect the impacts of cojumps on realized variances for spot and futures markets, respectively. The $R^2$ is for the $HAR–RCov–JCov$ and $HAR–RV–JCov$ models defined by Eqs. (11) and (12), and $R^2_{HAR}$ is for the HAR model with no jump covariation as a regressor defined by Eqs. (11) and (12).

* Denote statistical significance at 10% level.
** Denote statistical significance at 5% level.
*** Denote statistical significance at 1% level.
changes sharply the corresponding jump covariation moves the same way, especially in the second half of 2012 and the first half of 2013.

Based on the Newey–West heteroscedasticity consistent covariance matrix estimator for the estimates of the HAR-RCov-JCov model defined by (14) and the HAR-RV-JCov model defined by (15), we present the results of the these models respectively specified for daily, weekly, and monthly covariance and variance forecasts in Table 5. Estimates of $\alpha_d$, $\alpha_w$, and $\alpha_m$ in these models confirm highly persistent covariance and variance dependence. Interestingly, Table 5 shows the relative importance of the daily and weekly covariance and variance components decrease from the daily to the monthly regressions for the HAR-RCov-JCov model and the HAR-RV-JCov models for the spot and futures markets, whereas the monthly covariance and variance components tend to be relatively more important for the weekly and monthly regressions.

When the jump covariation is extracted from the realized covariance, Table 5 shows that the jump covariation component attributable to cojumps significantly impacts not only on daily covariance and variance, but also on weekly and monthly covariance and variance. Furthermore, compared with the daily, weekly and monthly covariance and variance components, the jump covariation component makes a greater impact on the daily forecasting of covariance and futures variance. The adjusted $R^2$ statistics reported in Table 5 indicate that the performance of these models decreases from the weekly, to the monthly to the daily covariance and variance forecasting. Comparing the adjusted $R^2$ of the HAR-RCov-JCov and the HAR-RV-JCov models with the adjusted $R^2_{\text{HAR}}$ of the standard HAR model wherein the cojump component is absent, we find that the gains by the cojump extraction from the realized covariance are significant and consistent over the daily, weekly, and monthly horizons. Thus, we conclude that cojumps between spot and futures markets significantly result in the out-of-sample covariance and variance forecast improvements, which is analogous to the demonstration of Andersen et al. (2007b) who extract jump component from realized variance to explore the impact of jumps on realized variance forecast.

6. Cojumps and macroeconomic announcements

This section describes the characteristics of the 12 macro news announcements which are released during the trading hours of the spot and futures markets, and analyzes the impact of the different announcements on cojumps between China’s CSI 300 futures and the CSI 300 index to determine the extent to which macro surprises cause cojumps.

6.1. Descriptive analysis

6.1.1. Macroeconomic announcements

We collected 12 macroeconomic news announcements from the WIND financial database. Table 6 describes the news announcements data. In the table, in addition to CPI, PPI, GDP, and M2, IP is short for industrial profit, RS for retail sales, FI for fixed investment, TB for trade balance, PMI for purchasing managers’ index, IV for industrial value-added and NL for new loans in CNY. Because the sample periods are from 9:30 am to 11:30 am and from 1:00 pm to 3:00 pm on every trading day, macroeconomic news announced in these two periods was collected while news announced beyond the periods was removed. For the sample data, macroeconomic news is generally released at a scheduled time and most often released at 10:00 am, and few announcements are made at 9:30. All news is released on a monthly basis except for GDP which is announced quarterly, so there are only 12 GDP announcements in our sample. M2 announcement times in the past three and a half years have, for the most part, been unscheduled and often released after 3:00 pm on news days, so many M2 announcements released after 3:00 pm have been removed and the number of M2 announcements is only 17 though M2 is released on a monthly basis.

To evaluate the impact of the various types of announcements on cojumps, we define the standardized surprise in announcements. Let $A_{k,i}$ denote the released value at period $i$, $k$ for announcement $k$, $E_{k,i}$ the mean of forecast survey, and $\sigma^k$ the standard deviation of the forecast error for the $k$th news, we measure the standardized surprise $S^k_{i,t}$ in announcement $i$ as

$$S^k_{i,t} = \frac{A_{k,i} - E_{k,i}}{\sigma^k}.$$  

6.1.2. Matching cojumps and macroeconomic announcements

This subsection considers the relationship between different macroeconomic news announcements and cojumps across the CSI 300 futures and the CSI 300 stock index, and investigates which news is most likely to produce cojumps. Firstly, we consider cojump-news matches. We count one cojump-news match if a cojump occurs within one hour after a news item is announced.

---

5 The adjusted $R^2$ statistics are low in Table 5 due to the limited sample for the futures market. Our sample just spans four-year period for the CSI 300 futures contract was launched on April 16, 2010. When we use the longer-period spot market data to test the HAR model for spot market, e.g. six-year period, the $R^2$ statistics are above 50%.

6 The WIND financial database was created by the Wind Information Co., Ltd, which serves more than 90% of the financial enterprises in China’s financial markets.

7 There are mainly two categories of PMI in China: official PMI and HSBC/Markit PMI. Official PMI is released at 9:00 am, while HSBC/Markit PMI is released at 9:45 am (before October 2012) and at 10:30 am (after that date). Because the announcement time of the HSBC/Markit PMI falls within the trading hours of the spot and futures markets, and the release time is historically earlier than for the official PMI, we use the HSBC/Markit PMI.

---

Please cite this article as: Wang, H., et al., Cojumps in China’s spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002
as Lahaye et al. (2009) do, and present the number of cojump-matches respectively for the 12 macroeconomic news announcements in the fourth column of Table 7. To analyze the associations between cojumps and diverse macroeconomic news, two probabilities are introduced to convey two different categories of information. One is \( P(\text{cojump}|\text{news}) \), describing the likelihood that a news release causes a cojump, and the other is \( P(\text{news}|\text{cojump}) \), suggesting what proportion of a given cojump is associated with a particular type of news. The two different probabilities are reported in the last two columns of Table 7. For \( P(\text{cojump}|\text{news}) \), the greatest four probabilities are 8.333% for RS, 6.250% for GDP, 5.556% for IV, and 5.405% for FI, which means that there is probability of 8.333%, 6.250%, 5.556% and 5.405% respective for RS, GDP, IV, FI to produce cojumps. The largest \( P(\text{news}|\text{cojump}) \) are 1.613% for RS, 1.075% for IV, EC, FI, PPI, and PMI indicating that 1.613% of cojumps are relative to RS and 1.075% of cojumps are associated with IV, EC, FI, PPI and PMI. Compared with \( P(\text{cojump}|\text{news}) \), the \( P(\text{news}|\text{cojump}) \) is much less due to the low frequency of news releases, thus resulting in the small number of cojump-news matches. From the results it can be seen that RS, GDP, IV and FI news announcements are most likely to produce cojumps in stock and futures markets among the 12 macroeconomic news announcements.

M2 announcement is noteworthy news with \( P(\text{cojump}|\text{news}) \) and \( P(\text{news}|\text{cojump}) \) being zero. M2 announcements regarding monetary policy seem not to be important factors in producing cojumps between the spot and the futures markets for the following two reasons. First, as we illustrated above, there is no fixed schedule of M2 announcements and so much of the M2 news announced before 9:30 am in the morning or after 3:00 pm in the afternoon has been removed from the sample. Second, China's

### Table 6
Descriptive analysis of all the 12 scheduled news.

<table>
<thead>
<tr>
<th></th>
<th>CPI</th>
<th>PPI</th>
<th>GDP</th>
<th>IP</th>
<th>EC</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>#news</td>
<td>40</td>
<td>40</td>
<td>16</td>
<td>30</td>
<td>42</td>
<td>36</td>
</tr>
<tr>
<td>Announcements Frequency</td>
<td>Monthly</td>
<td>Monthly</td>
<td>Quarterly</td>
<td>Monthly</td>
<td>Monthly</td>
<td>Monthly</td>
</tr>
<tr>
<td>Time of announcement</td>
<td>9:30/10:00</td>
<td>9:30/10:00</td>
<td>10:00</td>
<td>9:30/10:00</td>
<td>11:00</td>
<td>10:00/1:30*</td>
</tr>
<tr>
<td>Max abs(surprise)</td>
<td>2.283</td>
<td>2.609</td>
<td>2.026</td>
<td>4.058</td>
<td>4.039</td>
<td>5.691</td>
</tr>
<tr>
<td>Median abs(surprise)</td>
<td>0.457</td>
<td>0.580</td>
<td>0.203</td>
<td>0.159</td>
<td>0.473</td>
<td>0.281</td>
</tr>
<tr>
<td>Average abs(surprise)</td>
<td>0.753</td>
<td>0.667</td>
<td>0.439</td>
<td>0.534</td>
<td>0.681</td>
<td>0.476</td>
</tr>
<tr>
<td># of positive surprise</td>
<td>22</td>
<td>23</td>
<td>10</td>
<td>8</td>
<td>28</td>
<td>17</td>
</tr>
<tr>
<td># of negative surprise</td>
<td>13</td>
<td>19</td>
<td>5</td>
<td>22</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td># of zero surprise</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>FI</td>
<td>37</td>
<td>35</td>
<td>42</td>
<td>36</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>TB</td>
<td>35</td>
<td>35</td>
<td>42</td>
<td>36</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>PMI</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>36</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>IV</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>M2</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>NL</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Note: IP is short for industrial profits, EC for electricity consumption, RS for retail sales, FI for fixed investment, TB for trade balance, PMI for purchasing managers' index, IV for industrial value-added, and NL for new loans. In order to save space, we display most announcement times and omit certain announcement times when news is announced in limited frequency. For example, CPI, PPI, M2 and NL are announced only once at 1:30 pm.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 7
Cojumps and news.

|       | #news | P(\text{news})| #cojump-news | P(\text{cojump}|\text{news}) | P(\text{news}|\text{cojump}) |
|-------|-------|----------------|---------------|-------------------------------|-------------------------------|
| CPI   | 40    | 0.062          | 1             | 2.500                         | 0.538                         |
| PPI   | 40    | 0.082          | 2             | 5.000                         | 1.075                         |
| EC    | 42    | 0.086          | 2             | 4.762                         | 1.075                         |
| FI    | 37    | 0.076          | 2             | 5.405                         | 1.075                         |
| IP    | 30    | 0.061          | 0             | 0                             | 0                             |
| M2    | 17    | 0.035          | 0             | 0                             | 0                             |
| NL    | 20    | 0.041          | 0             | 0                             | 0                             |
| PMI   | 42    | 0.086          | 2             | 4.762                         | 1.075                         |
| RS    | 36    | 0.074          | 2             | 8.333                         | 1.613                         |
| TB    | 35    | 0.072          | 0             | 0                             | 0                             |
| IV    | 36    | 0.074          | 2             | 5.566                         | 1.075                         |
| GDP   | 16    | 0.033          | 1             | 6.250                         | 0.538                         |

Note: The table gives, from left to right, the macroeconomic news variables, the number of the announcements (#news), the probability (in %) of an announcement \( P(\text{news}) = \text{#news}/\text{#obs} \), the number of cojumps occurring within one hour after announcement arrival (#cojump-news), the probability (in %) of a cojump given a release \( P(\text{cojump}|\text{news}) = \text{#cojump-news}/\text{#news} \), and the probability (in %) of an announcement given a cojump \( P(\text{news}|\text{cojump}) = \text{#cojump-news}/\text{#cojumps} \) over the sample period April 16, 2010 to June 28, 2014.

Please cite this article as: Wang, H., et al., Cojumps in China's spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002
stock market participants pay more attention to the benchmark interest rate formulated by China’s Central Bank and to deposit-reserve ratio policy than to M2, so M2 is not the most accurate indicator representing monetary policy in China with which to analyze the relationship between cojumps and monetary policy.

6.2. Source of cojumps

The previous subsection described the conditional probabilities of cojumps and macroeconomic announcements and informally revealed how macroeconomic news is associated with cojump occurrence. This subsection provides a more formal context in which to discuss the relationship between cojumps and news surprises using logistic regression on rare events to model the probability of a cojump as a function of the 12 news surprises.

6.2.1. Relogit framework

Forecasting rare events such as cojumps in financial markets presents a challenge. King and Zeng (2001) indicate that classical logistic regression tends to bias the coefficients of independent variables which, in turn, leads to underestimations of the probability of rare events. One method for solving the problem is to increase the data size so as to collect more rare events, which is usually infeasible because of the high cost of doing so. In our case, it was not possible to collect more samples because the stock index futures exchange was not launched in China until April 16, 2010. Consequently, the sample was not of sufficient size to avoid bias in the classic logistic model and so in our estimate framework we have used logistic regression on rare events to get consistent and efficient estimates.

King and Zeng (2001) provide methods to address the bias problem in two different situations: endogenous stratified sampling and rare-event finite sampling. They propose a prior correction and weighting method to solve the bias problem for endogenous stratified sampling and extend the McCullagh and Nelder (1989) bias correction method by using their procedures to derive one which is used with the weighting method to solve the bias problem for finite endogenous stratified sampling. In this paper, considering the small size of our finite random sample, we use bias correction without the weighting method to adjust the estimators.

In the standard logistic regression, a single outcome variable $Y_i$ (1,2,3,...,N) follows a Bernoulli probability function that takes on the value 1 with probability $\pi_i$ and the value 0 with probability $1 - \pi_i$. Relogit regression employs the same framework with which $\pi_i$ varies over the observations as an inverse logistic function of vector $x_i$, which includes a constant and $k - 1$ explanatory variables:

$$
\pi_i(Y_i = 1|x_i) = \frac{e^{x_i\beta}}{1 + e^{x_i\beta}} = \Lambda(x_i\beta),
$$

(17)

where the unknown parameter $\beta = (\beta_0, \beta_1, \ldots, \beta_k)$ is a $k \times 1$ vector. $\beta_0$ is a scalar constant term and $\beta_k$ is a vector with elements corresponding to the explanatory variables. The marginal effect for the logistic model is denoted as

$$
\frac{\partial \pi_i(Y_i = 1|x_i)}{\partial x_i} = \frac{\partial \Lambda(x_i\beta)}{\partial x_i} = \Lambda(x_i\beta)(1 - \Lambda(x_i\beta))\beta_i.
$$

(18)

In general, for computing marginal effects, one can calculate this value at, say, the means of the regressors. The likelihood function of the logistic model is calculated by

$$
L(\beta|Y) = \prod_{i=1}^{N} \pi_i^{Y_i}(1 - \pi_i)^{1-Y_i},
$$

(19)

and the log-likelihood function simplifies to

$$
lnL(\beta|Y) = \sum_{Y_i=1} \ln(\pi_i) + \sum_{Y_i=0} \ln(1 - \pi_i) = -\sum_{i=1}^{N} \ln \left(1 + e^{1-2Y_iX_i\beta}\right).
$$

(20)

The bias-corrected estimate $\hat{\beta}$ is denoted as

$$
\beta = \hat{\beta} - bias(\hat{\beta}).
$$

(21)

where $bias(\hat{\beta}) = (X'WX)^{-1}X'Wc$ is computed for any generalized linear model. The term $X'WX$ is the Fisher information matrix and $c_i = -0.5(\mu_i/\mu_i')Q_n$, where $\mu_i = E(Y_i) = \pi_i = e^{x_i\beta}/(1 + e^{x_i\beta})$, $\mu_i'$ and $\mu_i''$ are the first and second derivatives of $\mu_i$ with respect to $x_i\beta$, and $Q_n$ are the diagonal elements of $X(X'WX)^{-1}X'$. After we estimate the bias of $\beta$ using the weighted least-squares regression with $X$ as the “independent variables”, $c$ as the “dependent variable”, and $W$ as the weight, then bias-corrected estimates of $\beta$ are obtained. The variance of corrected estimation of the $\beta$ is smaller than $\beta$, thus the corrected estimators $\beta$ are more efficient (King and Zeng, 2001).

Please cite this article as: Wang, H., et al., Cojumps in China’s spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002
6.2.2. Modeling cojumps to assess the impact of macroeconomic announcements

The relogit regression is used to analyze the 12 macroeconomic news announcements acting on cojumps between spot and futures markets and the regression model is defined as

\[
CJ_{i,t} = \alpha_0 + \sum_{k=1}^{12} \beta_k S_{i,t}^k + \epsilon_{i,t},
\]

\[
Cojump_{i,t} = I(CJ_{i,t}>0), \quad t = 1, 2, ..., 770, \quad i = 1, 2, ..., 48.
\]

where \(I(CJ_{i,t}>0)\) is an indicator function that transforms the latent variable \(CJ_{i,t}\) into the binary variable \(Cojump_{i,t}\), indicating the occurrence of cojumping and \(S_{i,t}^k(k=1, 2, ..., 12)\) accounts for the surprises of the \(k\)th news released at time \((t, i)\) and is 0 if no news was released that time. Note that the latent variable \(CJ_{i,t}\) can be viewed as the propensity of the occurrence of cojumps at time \((t, i)\) and is thus different from the jump sizes.

The regression coefficients for Eq. (22) and marginal effects calculated by using the sample means of news surprises according to Eq. (18) are presented in Table 8. The second column shows that EC, FI, GDP, IP, RS, and IV are statistically significant, meaning that the six news surprises have a significant impact on cojumps. This is consistent with Section 6.1.2 excepting IP. Among the six significant news surprises, GDP, IV, RS, and FI summarize real economic activity. Specifically, IV, RS and FI are important compositions of GDP in different calculating method. IP, as a reflection of the state of operations in a single industry, reports something closely related to the benefit of investors. EC, on the other hand, reflects production condition in different industries to some degree. Therefore, investors would pay a great deal of attention to these news surprises. The significance of the six macroeconomic news announcements in explaining the probability of cojump conveys the fact that investors in China’s spot and futures markets pay much attention to information regarding macroeconomic and industrial development conditions. The coefficients of EC, IP and RS are negative while the coefficients of GDP, IV, and FI are positive, suggesting that cojumps are more likely to occur when GDP, IV, and FI have positive surprises and when EC, IP and RS have negative surprises. In addition, the absolute values or marginal effects of GDP, RS and IV appear greater than those of EC, FI and IP, which indicates that GDP, RS, and IV surprises are more likely to produce cojumps than EC, FI and IP. Worth noting is the insignificance of both CPI and PPI in the model which may be attributed to goods and service price shocks not producing cojumps or to price shocks only impacting jumps occurring in one market. In a nutshell, macroeconomic news announcements are closely associated with cojump occurrence between China’s spot and futures markets and RS, GDP, IV, FI, IP and EC are the most important macroeconomic announcements among the 12 macroeconomic news categories.

7. Conclusions

Despite the existence of a number of studies addressing jumps, little attention has been paid to the occurrence of cojumps across China’s spot and stock index futures markets. This paper extracts jumps by the nonparametric method developed by Bollerslev et al. (2013), identifies cojumps with jump timing of both markets sampled from April 16, 2010 to June 28, 2014, and characterizes the dynamic of these cojumps.

The results lead us to the conclusion that China’s stock and stock index futures markets jump much more frequently with 0.978% and 1.259% of 5-minute returns exhibiting jumps respectively. Generally, the CSI 300 stock index futures jumps a bit more frequently and in larger size than the CSI 300 stock index does and both markets experience positive jumps more frequently than negative ones. Concretely, China’s spot market experienced 184 negative jumps and 294 positive jumps and the futures market experienced 280 negative and 353 positive ones, presenting an asymmetry between positive and negative jumps as well as demonstrating that China’s stock and futures markets tend to jump upward. In addition, there is significant evidence that cojumps exist in the two markets and that they are, on average, more frequent from 11:00 am to 11:30 am than other trading hours.

The realized covariance displays strong dependency but the jump covariation appears noticeably more erratic and less persistent than the total realized covariance. To explore the role of cojumps in realized covariance and variance forecasts, we built HAR-RcCov-JcCov and HAR-RV-JcCov models for daily, weekly, and monthly covariance and variance forecasts. The estimated results of these models confirm the existence of highly persistent covariance and variance dependence structure. Moreover, cojumps between spot and futures markets make a significant impact on covariance and variance forecasts over the daily, weekly, and monthly horizons.

Logistic regression on rare events (relogit model) is applied to study how the 12 macroeconomic news surprises explain cojumps. Among the 12 macroeconomic news announcements, surprises of GDP, industrial value-added, fixed investment, retail sales, electricity consumption and industrial profit can significantly explain the probability of cojumps. Specifically, GDP, industrial value-added, retail sales, and fixed investment are mirrors of real economic activity while industrial profit is directly related to investor benefit, and the electricity consumption, a reflection of production condition in different industries. Cojumps are more likely to occur when GDP, industrial value-added and fixed investment announcements have positive surprises and electricity consumption, industrial profit and retail sales have negative ones while GDP, retail sales and industrial value-added are more likely to produce cojumps than electricity consumption, fixed investment and industrial profit do if their surprises have the same absolute value.

---

*a* We do not consider potential lead and lag response to news due to the limited macroeconomic data in our sample and to the rare event dependent variable (cojump).
As for the future works, there are several interesting directions along different dimensions. One is to extend the current framework focusing on the contemporaneous jumps within the Chinese stock market to jump clustering in time as well as to jump contagion across international markets. Recent advances in mutually-exciting jump processes (Aït-Sahalia et al., 2015) opens doors to these interesting research questions. Another important direction is to understand the role of cojumps between spot and futures prices in the practice of risk management. A small but growing literature has provided evidence of the benefits of modeling cojumps in improving the portfolio allocation (Das and Uppal, 2004; Aït-Sahalia et al., forthcoming). Our findings of the statistical significance of forecasting realized variances by incorporating cojumps and of the impact of the scheduled news announcements on the occurrence of cojumps might shed light on the development of better risk models. A third direction is to investigate the impact of the market microstructure, such as liquidity and bid-ask spread, on cojumps between China’s spot and futures markets; see for example Balduzzi et al. (2001) and Jiang et al. (2011).

References


Table 8

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Marginal effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−5.532 (0.073)***</td>
</tr>
<tr>
<td>CPI</td>
<td>−0.572 (3.172)</td>
</tr>
<tr>
<td>EC</td>
<td>−6.301 (2.666)***</td>
</tr>
<tr>
<td>FL</td>
<td>10.110 (4.124)***</td>
</tr>
<tr>
<td>GDP</td>
<td>37.500 (10.61)***</td>
</tr>
<tr>
<td>IP</td>
<td>−14.170 (5.906)***</td>
</tr>
<tr>
<td>PPI</td>
<td>−1.365 (4.264)</td>
</tr>
<tr>
<td>PMI</td>
<td>−0.928 (2.446)</td>
</tr>
<tr>
<td>RS</td>
<td>−48.840 (4.785)***</td>
</tr>
<tr>
<td>TB</td>
<td>0.701 (3.604)</td>
</tr>
<tr>
<td>IV</td>
<td>26.890 (5.037)***</td>
</tr>
<tr>
<td>M2</td>
<td>−5.141 (4.263)</td>
</tr>
<tr>
<td>NL</td>
<td>1.955 (7.315)</td>
</tr>
</tbody>
</table>

Note: The table shows the impact of China macroeconomic announcements on cojumps defined by Eq. (22) and marginal effects calculated by Eq. (18) at the means of the regressors. The standard errors of regression coefficients are reported in parentheses.

⁎⁎⁎ Denote statistical significance at 5% level.

⁎⁎⁎⁎ Denote statistical significance at 1% level.

Please cite this article as: Wang, H., et al., Cojumps in China’s spot and stock index futures markets, Pacific-Basin Finance Journal (2015), http://dx.doi.org/10.1016/j.pacfin.2015.10.002