

LIHAR model for forecasting realized volatilities featuring long-memory and asymmetry

Jiwon Shin^a · Dong Wan Shin^{a,1}

^aDepartment of Statistics, Ewha Womans University

(Received July 15, 2016; Revised September 7, 2016; Accepted September 14, 2016)

Abstract

Cho and Shin (2016) recently demonstrated that an integrated HAR model has a forecast advantage over the HAR model of Corsi (2009). Recalling that realized volatilities of financial assets have asymmetries, we add a leverage term to the integrated HAR model, yielding the LIHAR model. Out-of-sample forecast comparisons show superiority of the LIHAR model over the HAR and IHAR models. The comparison was made for all the 20 realized volatilities in the Oxford-Man Realized Library focusing specially on the DJIA, the S&P 500, the Russell 2000, and the KOSPI. Analysis of the realized volatility data sets reveal apparent long-memory and asymmetry. The LIHAR model takes advantage of the long-memory and asymmetry and produces better forecasts than the HAR, IHAR, LHAR models.

Keywords: leverage, HAR model, asymmetry, long memory, nonstationarity, volatility forecasting

1. 서론

주가, 환율 등의 금융시계열의 변동성은 금융자산의 위험관리, 자산재분배, 파생상품가격결정 등 여러 측면에서 매우 중요하다. 그간 금융 변동성 추정치로 일별 수익률에 GARCH 류의 모형을 적용하여 구한 추정치를 주로 사용해왔다. 최근에는 일중 고빈도 거래 자료에 근거한 변동성 추정치인 실현변동성(realized volatility; RV)이 계산이 간단하면서도 정확하게 추정이 가능하여 주목받고 있다. 또한 실현변동성은 모형을 바탕으로 하지 않기 때문에 변동성 변수로 다룰 수 있다는 것도 큰 장점이다. 실현변동성이란 주가, 환율, 파생상품 등의 로그 가격 일일 변동성의 추정치이다. 종가를 P_t 로 나타내면 로그 가격은 $p_t = \ln P_t$, 일일 로그수익률은 $r_t = p_t - p_{t-1}$ 로 나타낼 수 있다. 실현변동성을 나타내기 위해서는 일중 M 개 등간격의 로그수익률 $r_{t-j\Delta} = p_{t-j\Delta} - p_{t-(j-1)\Delta}$, $j = 1, 2, \dots, M$, $\Delta = 1/M$ 계산이 필요하다. 이때 t 시일의 RV는 다음과 같이 나타낼 수 있다.

$$RV_t = \sqrt{\sum_{j=1}^M r_{t-j\Delta}^2}, \quad t = 1, 2, \dots,$$

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF)(2016R1A2B4008780).

¹Corresponding author: Department of Statistics, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Korea. E-mail: shindw@ewha.ac.kr

RV는 보통 일중 1분, 5분 혹은 10분 간격의 고빈도 데이터를 사용하여 계산되며 자산 리스크의 효과적인 평가 도구로 자산위험관리와 파생상품 가격책정, 자산 포트폴리오 작성 시 아주 중요하게 쓰인다. Andersen과 Bollerslev (1997)는 일일 변동성의 추정치로 5분 간격 RV가 우수한 추정치임을 지적한 바 있다. RV의 특성에 대한 논문은 Andersen 등 (2005), Banerjee와 Urga (2005), Huang와 Tauchen (2005), Deo 등 (2006), McAleer와 Medeiros (2008), Goncalves와 Meddahi (2009), Shephard와 Shephard (2010), Andersen 등 (2011), Chiriac과 Voev (2011), Todorov 등 (2011) 등이 있다.

미래 RV를 예측하는 것은 금융과 금융공학에서 매우 중요하고 실현변동성의 특성을 잘 반영하는 RV 예측모델링은 매우 의미 있는 일이며, 효율적인 예측을 위해서는 RV가 가지고 있는 중요한 성질을 적극적으로 반영하여야 한다. 실현변동성 예측에 관련된 논문으로 Corsi와 Renó (2009), Park과 Shin (2014), Souček과 Todorova (2014), Song과 Shin (2015), Yun과 Shin (2015) 등이 있다.

실현변동성의 가장 중요한 성질 중 하나는 장기기억성(long-memory)이다. 실현변동성의 표본자기상관함수는 시차가 증가함에 따라 느리게 감소하며 그 감소속도가 대수적으로 표현된다는 것이 여러 논문에서 보고된바 있다 (Ding과 Granger, 1996; Andersen 등, 2001). 예측에 있어서 장기기억성을 반영하는 방법 중 하나가 부분 누적(fractional integration)을 고려한 Granger와 Joyeux (1980)의 ARFIMA 모형이고 Dacorogna 등 (1993), Baillie (1996) 등의 논문에서 그 적용을 볼 수 있다. 다른 방법이 Corsi (2009), Hwang과 Shin (2014) 등에 의해 발표된 heterogeneous autoregressive(HAR) 모형이다. 이 모형은 실현변동성의 주간 이동평균, 월간 이동평균 등을 예측변수로 사용한 회귀모형으로 예측력이 좋을 뿐 아니라 예측함수의 의미가 명쾌해서 널리 사용되고 있다. Martens 등 (2009), Park과 Shin (2014) 등의 논문에서는 실현변동성이 장기기억성을 갖지만 정상계열임을 전제하여 예측방법이 구성되고 있다. 최근 Cho와 Shin (2016)은 실현변동성이 지니는 장기기억성이 매우 강하기 때문에 단위근을 부가한 비정상 HAR 모형, 즉 integrated HAR(IHAR) 모형이 더 예측력이 좋을 것을 보였다.

실현변동성에서 장기기억성 못지않게 중요한 것이 비대칭성(asymmetry)이다. Bekaert와 Wu (2000)는 금융시장에 비대칭성이 존재함을 발표하였고, McAleer와 Medeiros (2008), Corsi 등 (2012)은 장기기억성과 비대칭성을 모두 반영하는 레버리지(leverage)를 넣은 HAR 모형, leverage HAR(LHAR) 모형을 발표하였다. 이 논문들에서의 모형은 정상계열의 장기기억성을 전제한 것이다.

그러나 Cho와 Shin (2016)이 지적한 바와 같이 실현변동성의 장기기억성은 단위근을 부가한 비정상 모형에 의해 더 잘 반영되어 정상성을 가정한 모형 보다 뛰어난 예측력을 가진다는 점을 고려해 볼 때 비대칭성과 단위근을 동시에 고려했을 때 보다 우수한 예측을 달성할 수 있을 것이라 기대되어 이 부분을 집중 탐구하는 것이 본 논문의 중요 목적이다.

본 연구에서는 Cho와 Shin (2016)의 IHAR 모형에 비대칭성을 반영한 모형(LIHAR)을 제안하고 이 모형이 IHAR 모형보다 예측력이 뛰어난을 실증적으로 증명하고자 한다. 이를 위해서 주요 실현변동성 라이브러리 중 하나인 Oxford-Man Realized Library(<http://realized.oxford-man.ox.ac.uk/>)에 축적되고 있는 20개의 모든 주요 금융지수의 실현변동성 데이터를 사용하여 IHAR 모형과 LIHAR 모형의 예측력을 비교해 보일 것이다.

이후 본 논문은 다음과 같이 구성된다. 2절에서 논문에 쓰인 데이터에 대해 설명하고 3절에서 RV의 장기기억성 특성, 4절에서 RV의 비대칭 특성을 살펴본다. 5절에서는 새로 제안하는 LIHAR 모형을 소개하고 6절에서 기존의 모형들과 예측력 비교를 한 뒤, 7절에서 요약 및 결론을 낸다.

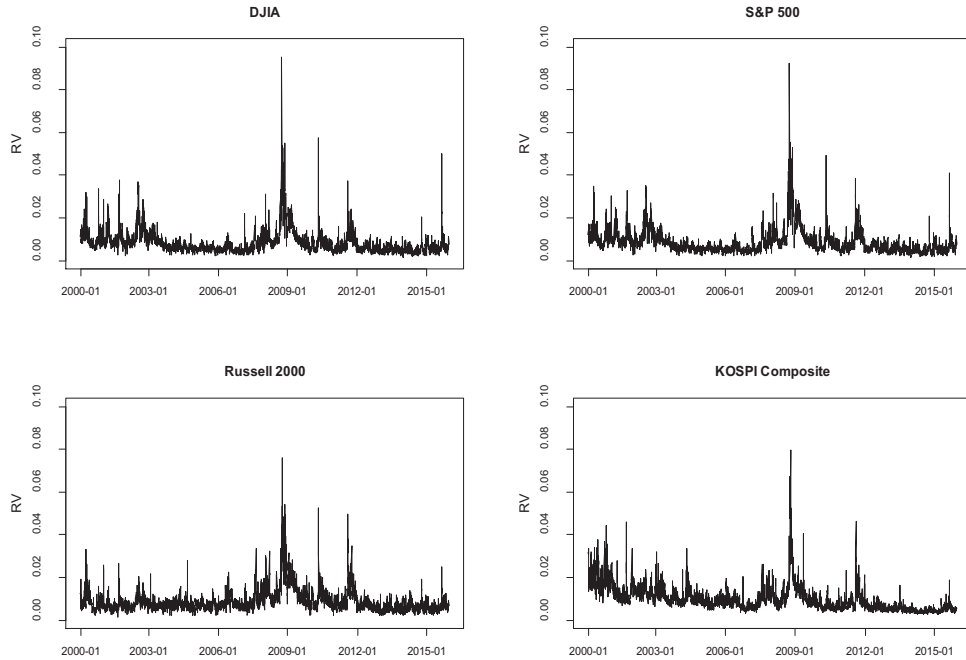
2. 데이터

실현변동성 예측 비교 데이터로는 Oxford-Man Realized Library 자료를 사용한다. 2000년부터

Table 2.1. Basic statistics of RVs

| RV | n | mean | se | sd | min | Q1 | med | Q3 | max | skew | kurt |
|--------------|------|-------|------|-------|-------|-------|-------|--------|--------|------|-------|
| DJIA | 3986 | 82.47 | 0.87 | 55.08 | 15.2 | 49.9 | 68.99 | 96.53 | 952.13 | 3.88 | 29.66 |
| S&P 500 | 3984 | 84.64 | 0.89 | 56.33 | 13.49 | 50.4 | 70.57 | 100.23 | 924.37 | 3.51 | 24.22 |
| Russell 2000 | 3985 | 86.35 | 0.84 | 53.15 | 15.91 | 55.95 | 72.66 | 98.34 | 758.21 | 3.31 | 19.18 |
| KOSPI | 3928 | 99.27 | 0.93 | 58.58 | 25.61 | 59.33 | 84.85 | 122.59 | 796.69 | 2.48 | 12.95 |

Unit = bp, 1bp = 0.01%, the scale BP applies only for mean, se, sd, min, Q1, med, Q3 and max.

**Figure 2.1.** Realized volatilities (RVs).

2015년까지의 기간으로 20개의 금융지수의 5분 간격 RV를 사용한다. 이중 먼저 미국 주가지수 3개와 한국 주가지수(DJIA, S&P 500, Russell 2000, KOSPI Composite)를 중점적으로 심도 있게 각각도 분석하고 후에 전체 20개 지수에 대해 예측력 비교 분석을 행한다.

Table 2.1은 네 지수의 실현변동성 기초통계량이다. 데이터 크기 n 은 1년간의 주식 개장일이 약 250일 정도이고 16년간의 데이터이기 때문에 약 4,000개로 나타남을 알 수 있다. 실현변동성의 mean은 80–90bp 정도로 하루의 0.8–0.9%를 나타낸다. 즉, 평균적으로 주가는 1% 미만의 변동을 나타냄을 알 수 있다. 또한 skewness가 3 정도로 비대칭성이 있는 것을 알 수 있고, kurtosis가 20 정도로 굉장히 큰 중심 집중 분포를 나타냄을 알 수 있다. 최솟값과 제1사분위수(Q1)의 차이에 비해 제3사분위수(Q3)와 최댓값의 차이가 훨씬 큰 것을 확인할 수 있다. 이처럼 실현변동성의 분포는 정규분포와는 거리가 있는 오른쪽으로 상당히 긴 꼬리를 가진 중심집중분포임을 알 수 있다.

Figure 2.1은 4가지 주요 지수의 실현변동성을 나타낸 그래프이다. 모든 지수에서 공통적으로 2007년 중반부터 2008년까지 큰 변동성이 나타나는 것을 확인할 수 있다. 이때가 2007–2008 세계금융위기로 서브프라임 모기지(subprime mortgage)사태로부터 시작되어 전 세계로 파급된 대규모의 금융위기 기

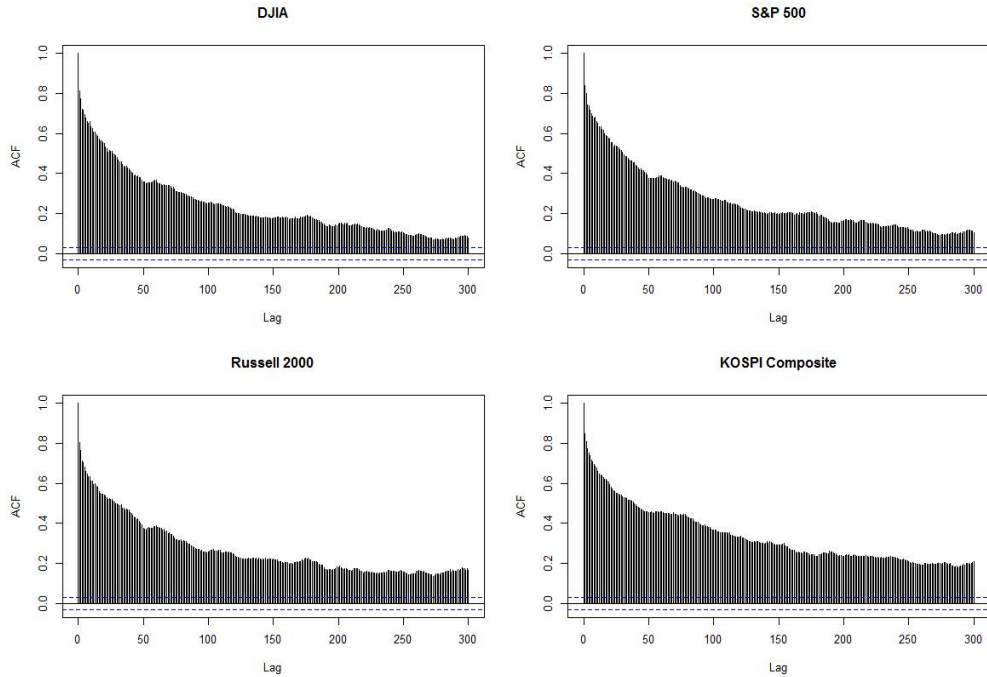


Figure 3.1. SACF of realized volatilities (RVs).

간이었다. 반면, 2010년 이후에는 세계금융위기 때에 비해서는 비교적 안정적인 변동성을 볼 수 있는 가운데 미국 지수의 경우 두 차례(2010년 5월경, 2011년 8월경) 중간 규모의 변동성 기간을 볼 수 있으며 KOSPI의 경우 한차례(2011년 8월경)의 중간 규모 변동성 기간을 볼 수 있다.

3. 장기기억과 비정상성

Figure 2.1의 RV 시도표에서 변동성이 클 때는 계속 변동성이 큰 경향이 있고, 변동성이 작을 때는 계속 작은 경향이 있음을 보고 실현변동성이 장기기억성을 가지고 있다고 추측할 수 있다. 장기기억성 특성은 Andersen 등 (2001)에서도 기술하고 있다. 또 Figure 3.1의 실현변동성의 표본자기상관함수(SACF)에서 장기기억성이 명확하다는 것을 알 수 있다. Figure 3.1에서 4개의 지수 모두 SACF가 시차가 300이 되도록 유의한 경향을 볼 수 있다. 이는 실현변동성이 장기기억성을 가진다는 것을 의미한다. 따라서 예측을 위한 모델링을 함에 있어 장기기억 특성을 반영해야 함을 알 수 있다.

이러한 특성을 반영하여 효율성과 정밀성을 인정받은 최근 가장 널리 사용되던 모델은 Corsi (2009)의 HAR 모형이고 다음과 같다.

$$RV_{t+1} = \beta_0 + \beta_1RV_t^d + \beta_2RV_t^w + \beta_3RV_t^m + w_t,$$

여기서 $RV_t^d = RV_t$ 이고 RV_t^w, RV_t^m 은 주간, 월간의 이동평균(moving average)로 다음과 같고

$$RV_t^w = \frac{1}{5} (RV_t + RV_{t-1} + \dots + RV_{t-4}),$$

$$RV_t^m = \frac{1}{22} (RV_t + RV_{t-1} + \dots + RV_{t-21})$$

Table 3.1. HAR model estimation results

| | DJIA | S&P 500 | Russell 2000 | KOSPI |
|-------------|---------------------|---------------------|---------------------|---------------------|
| | 4.500*** (0.969) | 0.416*** (0.925) | 4.784*** (1.025) | 4.450*** (1.010) |
| Day | 0.398*** (0.019) | 0.451*** (0.019) | 0.391*** (0.019) | 0.440*** (0.019) |
| Week | 0.380*** (0.029) | 0.339*** (0.028) | 0.373*** (0.029) | 0.363*** (0.272) |
| Month | 0.167*** (0.023) | 0.161*** (0.022) | 0.180*** (0.023) | 0.150*** (0.022) |
| Sum of Coef | 0.945 | 0.951 | 0.945 | 0.954 |

Note: Significance code *** for 0.1% levels. Numbers in parentheses are OLS standard errors.

Table 3.2. Unit root test results

| | ADF(AIC order) | KPSS | $\hat{d}(se)$ | GPH |
|--------------|----------------|----------|---------------|----------|
| DJIA | -4.996(25)** | 2.649** | 0.580(0.03) | 2.708** |
| S&P 500 | -4.861(25)** | 2.520** | 0.596(0.03) | 3.233*** |
| Russell 2000 | -5.183(20)** | 2.287** | 0.588(0.03) | 2.969** |
| KOSPI | -5.103(17)** | 12.623** | 0.598(0.03) | 3.296*** |

Significance code *** and ** for 0.1% and 1% levels.

w_t 는 오차항이다.

HAR 모형은 전기 값, 전기 주간 이동평균, 전기 월간 이동평균 항을 넣은 단순회귀모형이다. 네 가지 전제기간 데이터에 대한 HAR 적합 결과는 Table 3.1과 같다.

Table 3.1을 보면 모든 계수들이 0.1%에서 유의하여 잘 적합 되었으나 일별, 주별, 월별의 계수의 합이 약 0.95로 1에 매우 근접함을 알 수 있다. 이렇게 1에 가까운 계수 합이 얻어지는 것은 금융시계열의 RV가 단위근 계열이라는 가능성을 시사해준다.

단위근 여부 판단을 위해 Dickey와 Fuller (1979)의 Augmented Dickey-Fuller(ADF) 검정, Kwiatkowski 등 (1992)의 Kwiatkowski-Phillips-Schmidt-Shin(KPSS) 검정, Geweke와 Porter-Hudak (1983)의 GPH 검정을 적용하여 보았다.

Bisaglia과 Procidano (2002)의 연구에서 ADF 검정은 장기기억 비정상계열이나 부분누적모수가 $d > 0.5$ 이상의 부분누적 계열(fractional integration; $FI(d)$)의 검정에서 과기각 현상을 보인다는 결과 보고하였다. 이를 바탕으로 대안 검정인 KPSS 검정에 주목하였다. KPSS 검정은 Kwiatkowski 등 (1992)이 제안한 방법으로 흔히 알고 있는 ADF 검정과 같이 단위근 검정의 한 종류이다. KPSS 검정의 ADF 검정과는 반대로 귀무가설은 정상계열이며, 귀무가설이 기각될 때 비정상계열이라고 생각할 수 있다.

또 다른 검정인 GPH 검정은 부분누적 과정 $FI(d)$ 에서 $d < 0.5$ 일 때 정상계열이고 $d \geq 0.5$ 일 때 비정상계열임을 이용한 것으로 부분누적모수의 추정치 \hat{d} 를 0.5에서 표준화 시킨 통계량 $(\hat{d} - 0.5)/se(\hat{d})$ 이 귀무가설에서 표준정규분포라는 사실을 이용한 검정이다. GPH가 크면 정상성이 기각된다. GPH 추정치의 극한 분포는 Robinson (1995)이 발표하였으며, GPH 추정치의 편차나 mse 등의 특성은 Hurvich 등 (1998)의 논문에 잘 나타나 있다. 이 세 가지 단위근 검정 결과가 Table 3.2에 나타나 있다. Table 3.2에서 보면, AIC 기준으로 최적인 차수로 ADF 검정을 한 결과 모두 1% 유의수준에서 기각되므로 정상계열이라는 결과가 나왔다. 반면에 부분누적모수 d 의 값은 대체로 0.5-0.6 정도의 값을 가지고, DJIA,

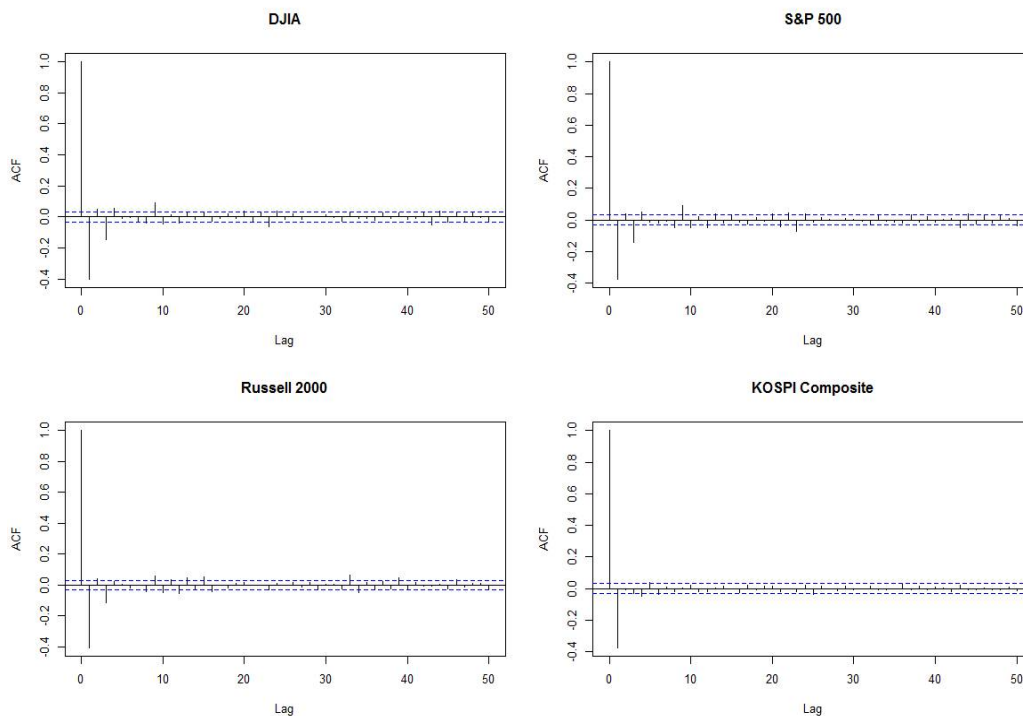


Figure 3.2. SACF of difference of realized volatilities (ΔRVs).

S&P 500, Russell 2000, KOSPI Composite 모두 GPH 검정 $(\hat{d} - 0.5)/se(\hat{d})$ 결과 5% 유의수준에서 정상계열이라는 귀무가설을 기각하고, 비정상계열이라는 결과를 얻었다. KPSS 검정 결과에서도 4가지 지수 모두에서 1% 임계값인 0.739를 한참 상회하므로 모두 1% 유의수준에서 정상계열이라는 귀무가설을 기각하고 비정상계열이라는 결과를 얻었다. ADF 검정 결과로 RV가 정상계열이라 판단했었지만 KPSS 검정과 GPH 검정 결과를 전체적으로 고려할 때 2000년부터 2015년까지의 각국의 금융지수의 RV는 비정상계열이라고 보고 이를 반영한 예측을 고려해볼 필요가 있다. 다음 Figure 3.2는 RV를 비정상계열이라고 보고 차분한 ΔRV 의 SACF이다. RV를 차분한 후의 SACF는 Figure 3.2의 RV의 SACF에 비하여 눈에 띄게 장기기억성이 감소한 것을 볼 수 있다. 이것은 모형에 단위근을 부과하는 자체만으로 장기기억성을 효과적으로 반영할 수 있다는 것을 의미한다.

기존의 연구들은 ADF 검정 결과만을 바탕으로 RV가 정상계열이라는 가정 하에 모형을 개발해 왔었다. 이 연구에서는 HAR 계수 합이 1에 근사하다는 것과, KPSS 검정 결과, 부분누적모수 d 추정값을 근거로 계수의 합을 1로 고정한 단위근 부과 모델(IHAR)을 사용한다. IHAR 모델에 대한 설명은 5절을 참고하기 바란다.

4. 비대칭성

금융지수의 RV에는 비대칭 변동성이 존재한다는 것은 오래전부터 알려진 사실이다. 금융시계열 비대칭변동성에 대한 연구는 Christie (1982), Campbell와 Hentschel (1992), Bekaert와 Wu (2000), McAleer와 Medeiros (2008), Asai 등 (2011), Corsi 등 (2012) 등의 많은 논문에서 다루졌다.

Table 4.1. Basic statistics of RV+, RV- which correspond to days of positive returns and negative returns, respectively

| r_{t-1} | nObs | | Mean | | Sd | | Min | | Median | | Max | |
|-------------|------|------|--------|-------|-------|-------|-------|-------|--------|-------|--------|--------|
| | - | + | - | + | - | + | - | + | - | + | - | + |
| DJIA | 1859 | 2127 | 90.56 | 75.41 | 62.69 | 46.31 | 19.05 | 15.20 | 74.83 | 63.17 | 952.13 | 525.79 |
| S&P 500 | 1875 | 2109 | 93.18 | 77.05 | 63.15 | 48.25 | 17.97 | 13.49 | 76.94 | 63.95 | 924.37 | 509.50 |
| Rusell 2000 | 1912 | 2073 | 94.54 | 78.81 | 59.33 | 45.46 | 21.77 | 15.91 | 78.74 | 67.23 | 758.21 | 435.87 |
| KOSPI | 1989 | 1939 | 101.20 | 97.30 | 60.70 | 56.26 | 26.68 | 25.61 | 85.73 | 83.67 | 674.42 | 796.69 |

Table 4.2. Correlations of RV with r_{t-1} , r_{t-1}^+ , r_{t-1}^-

| | Corr(r_{t-1} , RV_t) | Corr(r_{t-1}^+ , RV_t) | Corr(r_{t-1}^- , RV_t) |
|-------------|----------------------------|------------------------------|------------------------------|
| DJIA | -0.177 | -0.613 | 0.514 |
| S&P 500 | -0.196 | -0.643 | 0.517 |
| Rusell 2000 | -0.197 | -0.586 | 0.471 |
| KOSPI | -0.103 | -0.627 | 0.534 |

$$r_{t-1}^+ = \max(r_{t-1}, 0), \quad r_{t-1}^- = \min(r_{t-1}, 0).$$

금융시계열은 하루 전날의 금융지수에 하락이 나타나면 같은 크기로 상승한 날보다 그 다음날 금융지수의 변동성이 커지는 특징이 있다. 또한 변동성은 투자자에게 곧 리스크이기도 하다. 금융시계열의 변동성에서 비대칭성이 나타나는 이유는 금융자산 투자자에게 체감되는 리스크가 지수의 상승보다 하락에서 발생되기 때문이다. 다시 말하자면 시장 가격이 하락하면 레버리지가 증가하고, 이에 따른 금융자산 투자자들의 위험도가 높아지는 것이다. 여기서 레버리지 효과란 타인 자본을 이용한 자기자본이익률의 상승효과로 타인으로부터 빌린 차입금을 지렛대로 삼아 자기자본이익률을 높이는 것을 의미하며, 지렛대 효과라고도 한다.

금융지수의 RV에 비대칭 변동성이 존재한다는 것은 다음과 같은 결과들로도 확인할 수 있다. Table 4.1은 총 n 일 동안 각 날짜 t 일별로 전날의 수익률 r_{t-1} 이 음수(negative)인 경우와 양수(positive)인 경우로 나누어 각각의 RV의 관측수와 평균, 표준편차, 최솟값, 중앙값, 최댓값을 나타낸 표이다.

Table 4.1을 보면 4개의 지수 모두에서 전날의 수익률이 음수일 때가 양수일 때에 비하여 RV의 평균이 크고 표준편차도 큰 것을 알 수 있다. 수익률이 음수일 때 변동성이 평균적으로 더 크고 변동성의 변동성이 더 큰 경향이 있다는 것을 알 수 있다. Table 4.2는 전날의 수익률과 RV의 상관관계를 나타낸 표로 역시 음의 수익률을 가지는 경우와 양의 수익률을 가지는 경우로 나누어서 상관계수를 나타내었다.

Table 4.2에서 r_{t-1}^- 와 RV_t 는 음의 상관관계를 지니고 있는 것을 알 수 있으며, r_{t-1}^+ 와 RV_t 의 상관계수 절댓값이 r_{t-1}^- 와 RV_t 의 상관계수보다 더 큼을 알 수 있다. 이것은 음의 수익률이 RV의 값에 좀 더 영향을 많이 미치고 있음을 나타내고 절댓값이 큰 음의 수익률을 가지게 되었을 때 다음날의 RV값이 커지는 경향이 있음을 의미한다. 이것은 Figure 4.1에서 시각적으로도 확인할 수 있다.

Figure 4.1의 실선은 한 시차 전의 수익률이 음의 값(r_{t-1}^-)을 가지는 경우와 양의 값(r_{t-1}^+)을 가지는 경우로 나누어서 한 시차 전인 수익률 r_{t-1}^+ 또는 r_{t-1}^- 과 RV_t 와의 단순회귀 그래프이다. 각 자산별 그림에서 오른쪽 실선은 양의 수익률 r_{t-1}^+ 에 대한 RV_t 의 단순회귀이고, 왼쪽 실선은 음의 수익률 r_{t-1}^- 에 대한 RV_t 의 단순회귀이다. 각 색의 점선은 수익률 = 0을 기준으로 하여 대칭되도록 표시하여 실선의 기울기를 시각적으로 비교하기 좋게 나타낸 것이다. 4개의 지수에서 모두 왼쪽 실선의 기울기가 더 큰 것을 볼 수 있다. 음의 수익률일 때 RV가 전반적으로 더 크다는 것을 나타내며 이러한 점은 RV의 예측모형에 비대칭성을 반영하면 예측력의 개선을 기대해 볼 수 있게 한다.

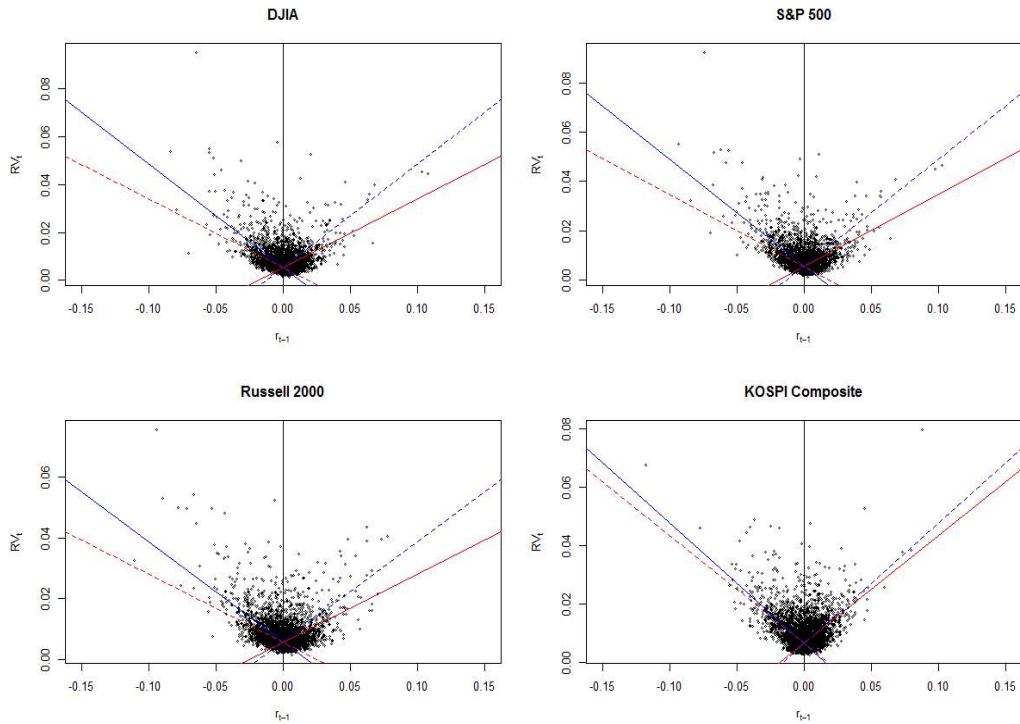


Figure 4.1. Regression plots of (RV_t, r_{t-1}^+) and (RV_t, r_{t-1}^-) .

이전에는 Engle (1982)의 Generalized Autoregressive Conditional Heteroskedasticity(GARCH) 류의 변동성 모형에 변동성의 비대칭 효과를 반영한 모형이 많이 사용되었다. Nelson (1991)의 Exponential GARCH(EGARCH) 모형과 Engle와 Ng (1993)의 Nonlinear GARCH(NGARCH) 모형, Glosten 등 (1993)의 Glosten-Jagannathan-Runkle GARCH(GJR-GARCH) 모형 등이 금융시계열 변동성에서의 비대칭성을 설명하기 위해 개발된 대표적인 모형이다. 그러나 Corsi 등 (2001)과 Corsi (2004)에서 제안한 HAR 모형이 GARCH 류의 모형보다 예측력이 더 뛰어나기 보고된 이후부터는 HAR 계통 모형도 널리 사용하게 되었다. 이후 Corsi (2009), Corsi 등 (2012), Souček과 Todorova (2014) 등에 의해 HAR 모형에 비대칭성을 고려하는 LHAR 모형에 대한 연구들이 진행되었다. HAR 모형에 대한 설명은 3절, LHAR 모형에 대한 설명은 5절을 참고하기 바란다.

5. LIHAR 모형

3절에서 파악된 RV의 비정상 장기기억성을 반영하여 HAR 모형을 변형시킨 것이 Cho와 Shin (2016)의 IHAR 모형으로 다음과 같다.

$$RV_{t+1} = \beta_0 + \beta_1 RV_t^d + \beta_2 RV_t^w + \beta_3 RV_t^m + w_t, \quad \beta_1 + \beta_2 + \beta_3 = 1. \quad (5.1)$$

이 모형의 $\beta_0, \beta_2, \beta_3$ 의 추정치와 이들의 표준오차 계산은 식 (5.1)을 오차수정모형 형태

$$RV_{t+1} - RV_t^d = \beta_0 + \beta_2 (RV_t^w - RV_t^d) + \beta_3 (RV_t^m - RV_t^d) + w_t \quad (5.2)$$

Table 5.1. LIHAR model estimation results

| | Coef | DJIA | S&P 500 | Russell 2000 | KOSPI |
|-------|------------|----------------------|----------------------|----------------------|----------------------|
| | β_0 | -5.522*** (0.554) | -5.601*** (0.530) | -6.071*** (0.552) | -3.707*** (0.585) |
| Day | β_1 | 0.275*** (0.019) | 0.323*** (0.019) | 0.284*** (0.019) | 0.361*** (0.019) |
| Week | β_2 | 0.404*** (0.029) | 0.382*** (0.028) | 0.390*** (0.029) | 0.385*** (0.028) |
| Month | β_3 | 0.321*** (0.023) | 0.295*** (0.022) | 0.326*** (0.022) | 0.254*** (0.021) |
| Day | γ_1 | -0.076*** (0.007) | -0.088*** (0.007) | -0.075*** (0.006) | -0.061*** (0.007) |
| Week | γ_2 | -0.194*** (0.021) | -0.167*** (0.019) | -0.132*** (0.015) | -0.121*** (0.018) |
| Month | γ_3 | 0.079 (0.044) | 0.109** (0.039) | 0.074* (0.030) | 0.150*** (0.034) |

Significance codes ***, **, and * for 0.1%, 1%, and 5% levels. Numbers in parentheses are OLS standard errors.

로 변환하여 추정하면 편리하다. 또 $\beta_0, \beta_1, \beta_2$ 의 추정치와 이들의 표준오차는

$$RV_{t+1} - RV_t^m = \beta_0 + \beta_1 (RV_t^d - RV_t^m) + \beta_2 (RV_t^w - RV_t^m) + w_t \quad (5.3)$$

를 추정하여 계산하면 편리하다. 이 모형은 단위근을 갖는 AR(22) 모형 중 하나이다. 이 모형은 HAR에 단위근을 부과한 모형으로 단위근 계열의 특성이자 금융시계열의 특성인 장기기억성을 기존의 HAR 모형보다 잘 나타낼 수 있기 때문에 모형의 예측력이 HAR 보다 우수하다는 것이 Cho와 Shin (2016)에 의해 입증된 바 있다.

4절에서 논한 바와 같이 RV는 매우 뚜렷한 비대칭성을 지니고 있다. 이를 적극적으로 반영하여 HAR 모형을 개선한 것이 Corsi와 Renó (2009), Asai 등 (2012) 등의 LHAR 모형으로 널리 사용되고 있으며 최근의 적용을 Souček와 Todorova (2014), Patton와 Sheppard (2015) 등의 논문에서 볼 수 있다. LHAR 모형은 다음과 같다.

$$RV_{t+1} = \beta_0 + \beta_1 RV_t^d + \beta_2 RV_t^w + \beta_3 RV_t^m + \gamma_1 r_t^{d-} + \gamma_2 r_t^{w-} + \gamma_3 r_t^{m-} + w_t,$$

여기서 $r_t^{d-} = r_t^d I(r_t^d < 0)$, $r_t^{w-} = (1/5)(r_t + r_{t-1} + \dots + r_{t-4})I(r_t + r_{t-1} + \dots + r_{t-4} < 0)$, $r_t^{m-} = (1/22)(r_t + r_{t-1} + \dots + r_{t-21})I(r_t + r_{t-1} + \dots + r_{t-21} < 0)$ 이고 $I(A)$ 는 사건 A의 지시함수이다. 즉, $r_t^{d-}, r_t^{w-}, r_t^{m-}$ 는 r_t^d, r_t^w, r_t^m 의 음수인 부분들만 나타낸 항이다.

앞서 3절에서 IHAR이 HAR에 비해 장기기억성을 잘 반영하는 모형이고 비정상계열로 보는 것의 타당성을 확인했으므로 IHAR 모형을 바탕으로 비대칭 변동성을 반영하도록 일별, 주별, 월별 레버리지 항 $r_t^{d-}, r_t^{w-}, r_t^{m-}$ 을 넣은 LIHAR 모형을 제안한다. 하루 전 수익률, 일주일 수익률과 한 달 수익률이 음인 경우를 나타내도록 모형에 포함시킴으로써 단기적 장기적인 레버리지를 모두 반영하도록 모형을 설계하였다. IHAR 모형에 레버리지를 넣은 모형(LIHAR)은 다음과 같다.

$$RV_{t+1} = \beta_0 + \beta_1 RV_t^d + \beta_2 RV_t^w + \beta_3 RV_t^m + \gamma_1 r_t^{d-} + \gamma_2 r_t^{w-} + \gamma_3 r_t^{m-} + w_t, \quad \beta_1 + \beta_2 + \beta_3 = 1. \quad (5.4)$$

이 모형의 추정은 모형 (5.4)를 모형 (5.2), (5.3)처럼 오차수정모형 형태로 바꾸어 추정하면 편리하다.

Table 6.1. Relative efficiencies of 1-step out-of-sample LIHAR forecasts relative to other HAR forecasts

| | HAR/LIHAR | | | IHAR/LIHAR | | | LHAR/LIHAR | | |
|--------------|-----------|-------|-------|------------|-------|-------|------------|---------|-------|
| | MAE | RMSE | MAPE | MAE | RMSE | MAPE | MAE | RMSE | MAPE |
| DJIA | 1.050** | 1.044 | 1.129 | 1.043** | 1.048 | 1.085 | 1.021 | 0.974 | 1.108 |
| S&P 500 | 1.060** | 1.050 | 1.143 | 1.051** | 1.054 | 1.095 | 1.010 | 0.965* | 1.104 |
| Russell 2000 | 1.057** | 1.014 | 1.132 | 1.047** | 1.017 | 1.091 | 1.006 | 0.973** | 1.064 |
| KOSPI | 1.048'' | 1.011 | 1.122 | 0.979 | 0.990 | 0.996 | 1.074** | 1.014 | 1.162 |

Significance codes for Diebold-Mariano test ***, **, *, and '' for 0.1%, 1%, 5%, and 10% levels.

Table 5.1에 LIHAR 모형으로 적합시킨 회귀계수의 추정결과를 나타내었다. Table 5.1에서 보면 4개 지수에서 거의 모든 변수들이 0.1% 유의수준에서 유의함을 알 수 있다. 특히 레버리지 항들의 계수 $\gamma_1, \gamma_2, \gamma_3$ 도 상당히 유의하여 비대칭성이 RV 모델링에 상당히 기여한다는 것을 알 수 있다. 다음 6절에서 LIHAR이 미래 RV 예측을 다른 모델에 비해 잘하는지 평가해본다.

6. 예측력비교

모델의 예측력을 평가하기 위해서 데이터는 우선적으로 4개의 금융지수를 2000년부터 2015년까지의 기간으로 사용하였다. 데이터 계열의 크기를 n 이라할 때 85%데이터($T = 0.85n$)부터 모델을 적합 시키고 최대 15%($m = 0.15n$)까지의 데이터로 예측력을 평가하였다. 평가방법으로는 표본 외 1, 5, 10, 15시차(out-of-sample h-step-ahead)의 예측 결과를 MAE, MAPE와 RMSE의 3가지 손실함수 기준으로 평가하였다. 예를들면 1-step의 경우 손실함수들은 다음과 같다.

$$\begin{aligned}
 MAE &= m^{-1} \sum_{T=n-m}^{n-1} \left| \widehat{RV}_{T+1|T} - RV_{T+1|T} \right|, \\
 MAPE &= m^{-1} \sum_{T=n-m}^{n-1} \left| \widehat{RV}_{T+1|T} - RV_{T+1|T} \right| / RV_{T+1}, \\
 RMSE &= \sqrt{m^{-1} \sum_{T=n-m}^{n-1} \left(\widehat{RV}_{T+1|T} - RV_{T+1|T} \right)^2},
 \end{aligned}$$

여기서 $\widehat{RV}_{T+1|T}$ 은 시점 T 까지의 자료에 근거한 RV_{T+1} 의 예측치이다. Table 6.1에서는 MAE, RMSE, MAPE의 상대효율(relative efficiency)를 나타내었다. HAR 모형에 대한 LIHAR 모형의 예측 MAE는 다음과 같이

$$MAE_{\text{상대 효율}} = \frac{\text{HAR 모형 MAE}}{\text{LIHAR 모형 MAE}}$$

계산하고 RMSE나 MAPE의 상대효율도 마찬가지로 계산된다. 상대효율이 1 이상의 값이 얻어지면 제안된 모형 LIHAR이 비교대상인 모형보다 예측력 측면에서 더 나은 모델이라고 볼 수 있다. 또한 예측치의 차이가 명확한지 비교하기 위해 MAE와 RMSE의 상대효율의 통계적인 유의성을 Diebold와 Mariano (1995)가 제안한 Diebold-Mariano(DM) 검정에 의해 계산하여 상대효율값에 표시하였다.

먼저 1-step 예측결과를 살펴본다. Table 6.1을 보면 HAR/LIHAR에서 MAE, RMSE, MAPE 기준에서 모두 1 이상으로 모두 LIHAR이 기본 HAR 모델보다 예측을 잘함을 알 수 있고, IHAR/LIHAR에서도 모두 1 이상이므로 IHAR 모델보다 IHAR 모델에 비대칭성까지 반영하였을 때 예측력이 개선되는

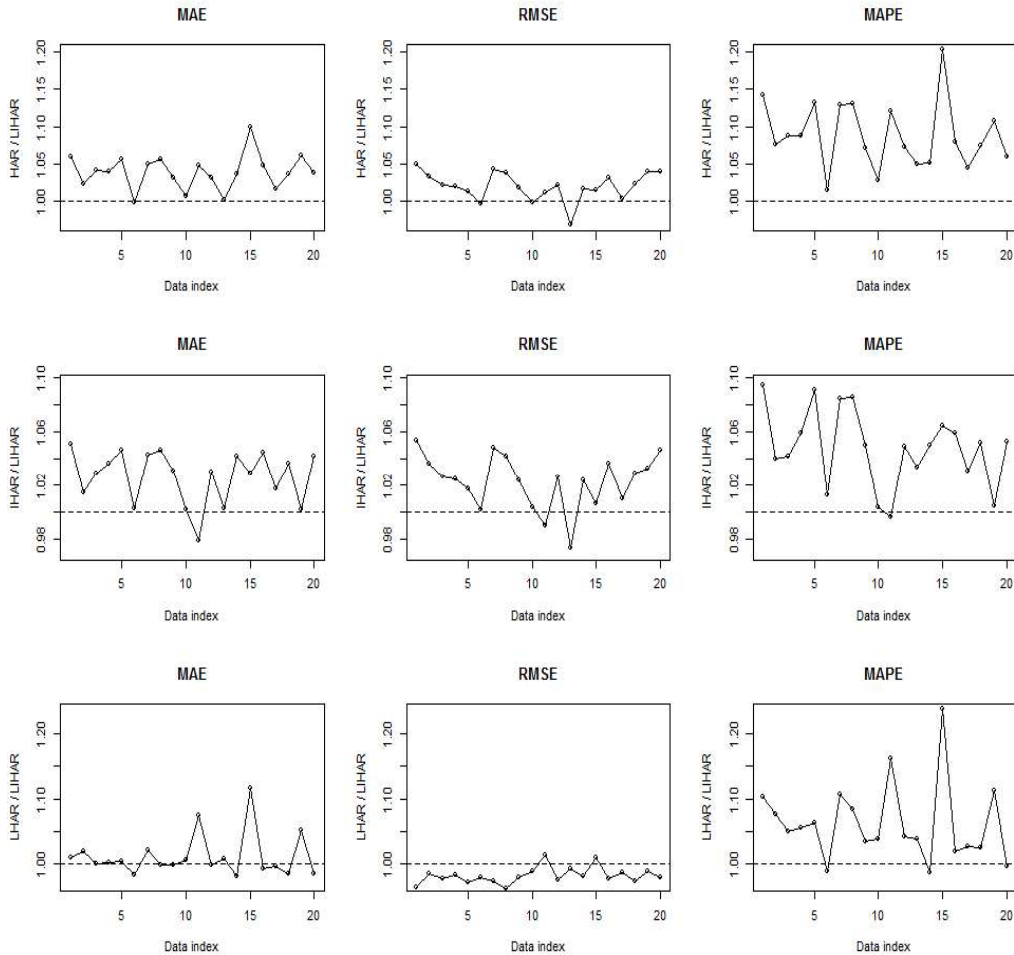


Figure 6.1. Relative efficiencies of out-of-sample LIHAR forecasts relative to HAR, IHAR and LHAR forecasts.

것을 확인할 수 있었다. 또한 기본 HAR에 비대칭성을 고려한 LHAR 모델보다 LIHAR이 전반적으로 우수함을 확인하였다. 아울러 MAE와 RMSE의 상대효율이 대체적으로 5%에서 통계적으로 유의함을 볼 수 있다.

Figure 6.1은 여러 가지 데이터로 LIHAR의 우수성을 검증해보기 위해 Oxford-Man Institute(<http://realized.oxfordman.ox.ac.uk/>)에 있는 20가지 주요 주가지수의 실현변동성 데이터 모두에 대해 MAE, RMSE, MAPE의 1단계 예측 상대효율을 구하여 시각적으로 나타낸 것이다. 각 주가지수 명과 이에 대응하는 번호는 Table 6.2에 명시하였다. Figure 6.1에서 상대효율이 전반적으로 모두 1위에 있는 것을 볼 수 있다. LAHR/LIHAR의 RMSE기준에서 일부 약간 열등한 경향이 있는 것을 제외하면, LIHAR이 기존의 모형들보다 전반적으로 우수하다는 것을 알 수 있다.

Figure 6.1에서 HAR/LIHAR과 IHAR/LIHAR의 경우 MAE, RMSE, MAPE 모두 거의 1 수평선 위쪽에 위치하는 것을 볼 수 있다. 그리고 LHAR/LIHAR의 경우에는 MAE와 MAPE가 상당 부분 1

Table 6.2. Data number & description table

| Number | Index | Number | Index |
|--------|---------------|--------|------------------|
| 1 | S&P 500 | 11 | KOSPI Composite |
| 2 | FTSE100 | 12 | AEX |
| 3 | Nikkei225 | 13 | SwissMarket |
| 4 | DAX | 14 | IBEX35 |
| 5 | Russell 2000 | 15 | CNXNifty |
| 6 | AllOrdinaries | 16 | IPC Mexico |
| 7 | DJIA | 17 | Bovespa |
| 8 | Nasdaq | 18 | Euro STOXX50 |
| 9 | CAC40 | 19 | FT Straits Times |
| 10 | Hangseng | 20 | FTSE MIB |

수평선 위고, RMSE에서만 1 수평선 살짝 아래에 전반적으로 존재함을 확인할 수 있었다. RMSE는 MAE나 MAPE에 비해 오차가 제곱 꼴로 계산되므로 몇 개의 시차에서 상대적으로 큰 오차가 발생하면 값이 커지게 된다. MAE와 MAPE는 우수하고 RMSE에서 부진한 것은 전체적으로 LIHAR의 성능이 우수하나 몇몇 소수의 경우 오차가 LHAR에 비해 매우 크다는 것을 의미한다. 그러나 전반적으로 HAR/LIHAR, IHAR/LIHAR과 LHAR/LIHAR 모두에서 3가지 손실함수의 상대효율이 전반적으로 1 수평선 이상을 상회하고 있으므로 LIHAR이 20개의 세계적 금융데이터들 모두에 대해 기존의 HAR, IHAR, LHAR모델에 비해 LIHAR의 예측 성능이 우수하다고 할 수 있다.

다음은 DJIA, S&P 500, Russell 2000, KOSPI에 대해 1, 5, 10, 15단계 예측 비교를 논한다. Figure 6.2는 HAR, IHAR, LHAR, LIHAR 모형의 장기예측 성능 그래프이고 $h = 1, 5, 10, 15$ 단계 예측을 결과를 나타내었다. 이때, 레버리지를 넣은 모형 LHAR과 LIHAR은 2단계 이상의 장기예측 시 미관측 r_{t+h} 를 이용한 $E_T(r_{T+h}^d)$, $E_T(r_{T+h}^w)$, $E_T(r_{T+h}^m)$ 의 추정치 필요하다. 이들은 시점 T 에서 r_{T+h}^d , r_{T+h}^w , r_{T+h}^m 의 조건부 기대치를 의미한다.

$E_T(r_{T+2}^w)$ 의 계산과정을 좀 더 자세히 보면 다음과 같다. $r_{T+2}^w = \min(0, (r_{T+2} + r_{T+1} + r_T + r_{T-1} + r_{T-2})/5)$ 이고, 이때 $r_T + r_{T-1} + r_{T-2} = A_t^w$, $r_{T+2} + r_{T+1} + r_T + r_{T-1} + r_{T-2} = B_t^w$ 라 하면 A_t^w 와 B_t^w 는 이변량 정규분포 $\begin{pmatrix} A_t^w \\ B_t^w \end{pmatrix} \sim N\left[\begin{pmatrix} 3\mu \\ 5\mu \end{pmatrix}, \begin{pmatrix} 3\sigma_t^2 & 3\sigma_t^2 \\ 3\sigma_t^2 & 5\sigma_t^2 \end{pmatrix}\right]$ 를 따름을 알 수 있다. 여기서 A_t^w 가 주어진 하에서 B_t^w 의 조건부 분포는 쉽게 유도할 수 있고 이로부터 $E_T(r_{T+h}^d)$, $E_T(r_{T+h}^w)$, $E_T(r_{T+h}^m)$ 을 계산하였다. 또 이들의 계산에 있어서 각각의 r_t 에 대해 $N(\mu, \sigma_t^2)$ 을 가정하였고, 정규분포 모수의 추정치로는 r_t 에 AR(0)-GARCH(1,1) 적합으로 추정된 모수를 사용하였다.

Figure 6.2의 장기예측 결과를 보면 레버리지를 추가한 모형 LHAR과 LIHAR의 성능이 레버리지를 추가하지 않은 모형 HAR과 IHAR 모두에 비해 성능이 월등히 우수하였고 이것은 DJIA, S&P 500, Russell 2000, KOSPI 4가지 데이터에 모두에서 확인할 수 있다. MAE기준이나 MAPE기준에서 LIHAR모형이 장기예측에서도 전반적으로 우수함을 확인할 수 있었다. 특히 KOSPI에서는 장기예측으로 갈수록 MAE, RMSE, MAPE 모든 기준에서 LIHAR의 우수성이 두드러졌다. 그런데 KOSPI를 제외한 3가지 데이터에서 IHAR도 HAR보다 RMSE기준의 성능이 떨어지는 경우를 확인할 수 있는데, 이것은 LIHAR이 LHAR에 RMSE기준에서 약간 부진한 이유와 같다. IHAR이 전반적으로 HAR보다 우수한 예측력을 가지고 있더라도, HAR에 비해 IHAR이 몇몇 소수의 큰 오차를 발생시킬 수 있음을 한 번 더 확인할 수 있었다. 이 비교 분석을 통해 금융변동성 예측에 있어 단위근과 레버리지가 포함된 LIHAR 모형의 활용이 특히 효과적일 수 있음을 확인하였다.

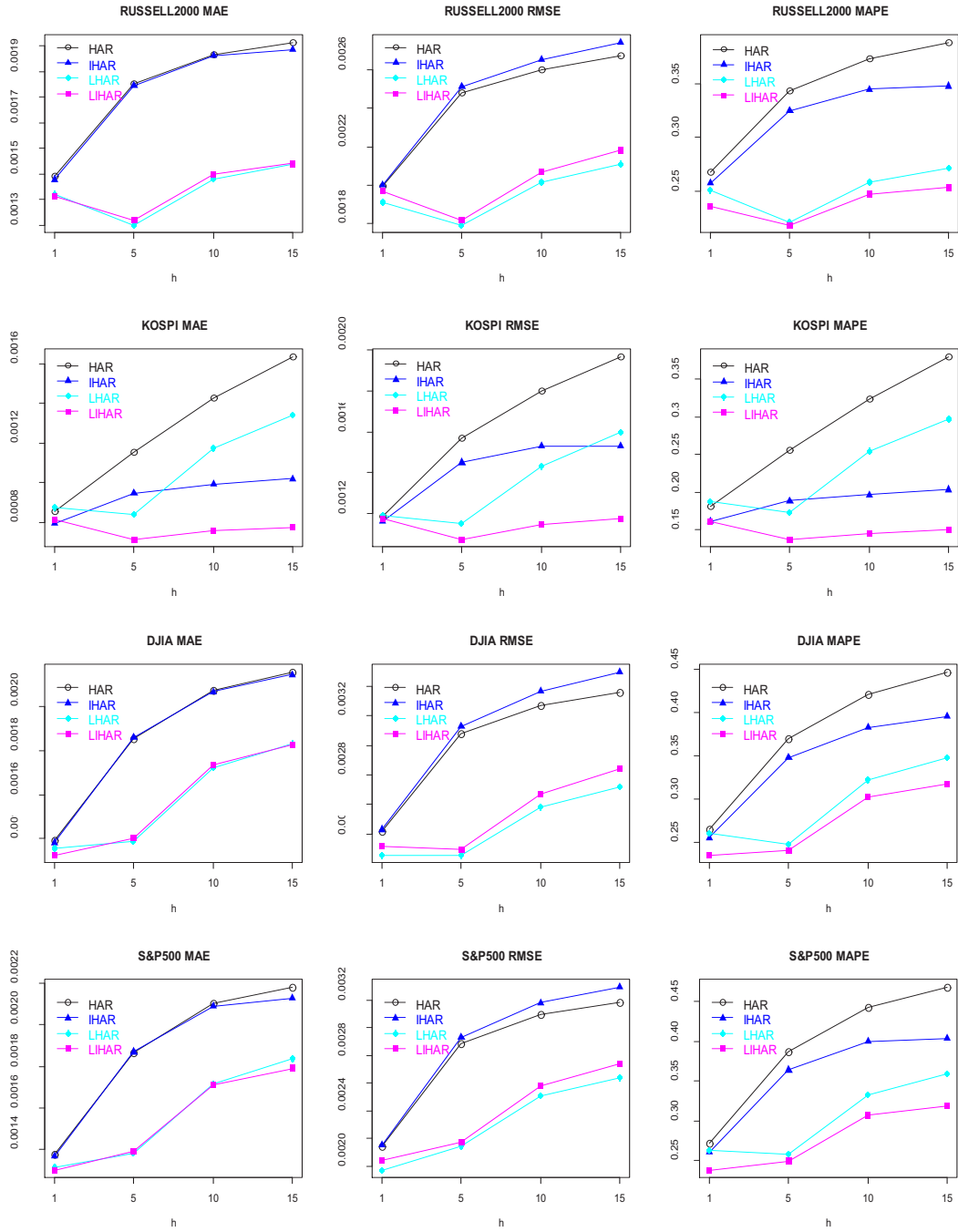


Figure 6.2. h -step forecasting results of HAR, IHAR, LHAR and LIHAR for DJIA, S&P 500, Russell 2000 and KOSPI.

7. 결론

자산리스크의 효과적인 평가 도구로 자산위험관리와 파생상품 가격책정, 자산 포트폴리오 작성 시 아주 중요하게 쓰이는 실현변동성을 정확하게 예측하기 위해 실현변동성이 지니는 특성을 살펴보았다. 실현변동성의 주요한 첫 번째 특성으로 장기기억성을 들 수 있었고 이것은 SACF의 시차가 300을 넘도록 유의한 점과 부분누적모수(fractional integration parameter) d 의 추정치가 0.5 이상인 것으로 확인 가능했다. 실현변동성의 두 번째 특성으로는 비대칭성을 들 수 있고 이것은 음의 수익률 r_{t-1}^- 와 RV_t 의 상관계수가 양의 수익률과 r_{t-1}^+ 와 RV_t 의 상관계수보다 큼을 살펴봄으로써 실현변동성의 비대칭성을 확인할 수 있었다. 장기기억성은 기존의 HAR (Corsi, 2009) 모형보다 단위근을 부과한 비정상 모형 IHAR (Cho와 Shin, 2016) 모형에 의해 더 잘 반영됨을 확인하고, 이 모형을 바탕으로 본 연구에서는 비대칭성까지 고려한 LIHAR 모형을 제안하였다. 제안한 LIHAR 모형은 금융 실현변동성의 특성인 장기기억성과 비대칭성을 잘 반영하는 모형으로 20개의 주요 금융시계열을 바탕으로 한 예측력평가에서 LIHAR 모형이 기존의 HAR, IHAR, LHAR 모형보다 전반적으로 우수함을 실증적으로 확인하였다. 또한 장기예측에서도 LIHAR 모형의 전반적인 우수성을 확인할 수 있었으며, 한국 주요지표인 KOSPI Composite에서 LIHAR 모형이 다른 모형에 비해 두드러지게 우수한 것은 한국 금융변동성 예측에 있어 LIHAR 모형이 상당한 가치가 있음을 시사한다.

References

- Andersen, T. G. and Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns, *The Journal of Finance*, **52**, 975–1005.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of realized stock return volatility, *Journal of Financial Economics*, **61**, 43–76.
- Andersen, T. G., Bollerslev, T., and Meddahi, N. (2005). Correcting the errors: volatility forecast evaluation using high-frequency data and realized volatilities, *Econometrica*, **73**, 279–296.
- Andersen, T. G., Bollerslev, T., and Meddahi, N. (2011). Realized volatility forecasting and market microstructure noise, *Journal of Econometrics*, **160**, 220–234.
- Asai, M., McAleer, M., and Medeiros, M. C. (2012). Asymmetry and long memory in volatility modeling, *Journal of Financial Econometrics*, **10**, 495–512.
- Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics, *Journal of Econometrics*, **73**, 5–59.
- Banerjee, A. and Urga, G. (2005). Modelling structural breaks, long memory and stock market volatility: an overview, *Journal of Econometrics*, **129**, 1–34.
- Bekaert, G. and Wu, G. (2000). Asymmetric volatility and risk in equity markets, *The Review of Financial Studies*, **13**, 1–42.
- Bisaglia, L. and Procidano, I. (2002). On the power of the augmented Dickey-Fuller test against fractional alternatives using bootstrap, *Economics Letters*, **77**, 343–347.
- Campbell, J. Y. and Hentschel, L. (1992). No news is good news: an asymmetric model of changing volatility in stock returns, *Journal of Financial Economics*, **31**, 281–318.
- Chiriac, R. and Voev, V. (2011). Modelling and forecasting multivariate realized volatility, *Journal of Applied Econometrics*, **26**, 922–947.
- Cho, S. and Shin, D. W. (2016). An integrated heteroscedastic autoregressive model for forecasting realized volatilities, *Journal of the Korean Statistical Society*, **45**, 371–380.
- Christie, A. A. (1982). The stochastic behavior of common stock variances: value, leverage and interest rate effects, *Journal of Financial Economics*, **10**, 407–432.
- Corsi, F. (2004). A simple long memory model of realized volatility, Available from: <http://dx.doi.org/10.2139/ssrn.626064>
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility, *Journal of Financial Econometrics*, **7**, 174–196.

- Corsi, F., Audrino, F., and Renó, R. (2012). HAR modeling for realized volatility forecasting, *Handbook of Volatility Models and Their Applications* (pp. 363–382), John Wiley & Sons, New Jersey.
- Corsi, F. and Renó, R. (2009). HAR volatility modelling with heterogeneous leverage and jumps, Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.375.5804&rep=rep1&type=pdf>
- Corsi, F., Zumbach, G., Müller, U. A., and Dacorogna, M. M. (2001). Consistent high-precision volatility from high frequency data, *Economic Notes*, **30**, 183–204.
- Dacorogna, M. M., Müller, U. A., Nagler, R. J., Olsen, R. B., and Pictet, O. V. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market, *Journal of International Money and Finance*, **12**, 413–438.
- Deo, R., Hurvich, C., and Lu, Y. (2006). Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment, *Journal of Econometrics*, **131**, 29–58.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, **74**, 427–431.
- Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy, *Journal of Business & Economic Statistics*, **13**, 134–144.
- Ding, Z. and Granger, C. W. (1996). Modeling volatility persistence of speculative returns: a new approach, *Journal of Econometrics*, **73**, 185–215.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1007.
- Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility, *The Journal of Finance*, **48**, 1749–1778.
- Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models, *Journal of Time Series Analysis*, **4**, 221–238.
- Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks, *The Journal of Finance*, **48**, 1779–1801.
- Goncalves, S. and Meddahi, N. (2009). Bootstrapping realized volatility, *Econometrica*, **77**, 283–306.
- Granger, C. W. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis*, **1**, 15–29.
- Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price variance, *Journal of Financial Econometrics*, **3**, 456–499.
- Hurvich, C. M., Deo, R., and Brodsky, J. (1998). The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series, *Journal of Time Series Analysis*, **19**, 19–46.
- Hwang, E. and Shin, D. W. (2014). Infinite-order, long-memory heterogeneous autoregressive models, *Computational Statistics & Data Analysis*, **76**, 339–358.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?, *Journal of Econometrics*, **54**, 159–178.
- Martens, M., Van Dijk, D., and De Pooter, M. (2009). Forecasting S&P 500 volatility: long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements, *International Journal of Forecasting*, **25**, 282–303.
- McAleer, M. and Medeiros, M. C. (2008). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries, *Journal of Econometrics*, **147**, 104–119.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach, *Econometrica*, **59**, 347–370.
- Park, S. and Shin, D. W. (2014). Modeling and forecasting realized volatilities of Korean financial assets featuring long memory and asymmetry, *Asia-Pacific Journal of Financial Studies*, **43**, 31–58.
- Patton, A. J. and Sheppard, K. (2015). Good volatility, bad volatility: signed jumps and the persistence of volatility, *Review of Economics and Statistics*, **97**, 683–697.
- Robinson, P. M. (1995). Log-periodogram regression of time series with long range dependence, *The Annals of Statistics*, **23**, 1048–1072.
- Shephard, N. and Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility

- (HEAVY) models, *Journal of Applied Econometrics*, **25**, 197–231.
- Song, H. and Shin, D. W. (2015). Long-memories and mean breaks in realized volatilities, *Applied Economics Letters*, **22**, 1273–1280.
- Souček, M. and Todorova, N. (2014). Realized volatility transmission: the role of jumps and leverage effects, *Economics Letters*, **122**, 111–115.
- Todorov, V., Tauchen, G., and Gryniv, I. (2011). Realized Laplace transforms for estimation of jump diffusive volatility models, *Journal of Econometrics*, **164**, 367–381.
- Yun, S. and Shin, D. W. (2015). Forecasting the realized variance of the log–return of Korean won US dollar exchange rate addressing jumps both in stock-trading time and in overnight, *Journal of the Korean Statistical Society*, **44**, 390–402.

장기기억성과 비대칭성을 띠는 실현변동성의 예측을 위한 LIHAR모형

신지원^a · 신동완^{a,1}

^a이화여자대학교 통계학과

(2016년 7월 15일 접수, 2016년 9월 7일 수정, 2016년 9월 14일 채택)

요약

최근에 Cho와 Shin (2016)가 변동성 예측 모형으로 유명한 HAR (Corsi, 2009) 모형보다 단위근을 부과한 IHAR 모형이 더 우수하다는 것이 보고하였다. 금융시계열에 비대칭 변동성이 존재한다는 것은 널리 알려져 있다. 이 논문에서는 IHAR 모형에 레버리지를 고려한 LIHAR 모형을 제안한다. LIHAR 모형과 IHAR 모형 기존의 HAR 모형, LHAR 모형과의 예측력 비교를 통해 LIHAR 모형의 우수성을 보인다. 모형을 평가하기 위해 Oxford-Man 라이브러리 20개의 실현변동성 데이터를 이용하였다. 특히 DJIA, S&P 500, Russell 2000, KOSPI Composite 데이터는 다양한 분석을 하였다. 주가와 같은 금융지수의 변동성에는 장기기억성과 비대칭 변동성이 존재하고, 이런 특징을 LIHAR 모형이 HAR, IHAR, LHAR 모형보다 적절하게 반영하고 있는 것을 확인 하였다. 또한 예측력도 LIHAR 모형이 가장 우수하였다. 금융시계열의 실현변동성에 장기기억성, 비대칭변동성, 비정상성을 모두 반영하여 예측하는 것이 상당한 가치가 있음을 확인하였다.

주요용어: 레버리지, HAR 모형, 비대칭 변동성, 장기기억성, 비정상성, 변동성 예측

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) (2016R1A2B4008780).

¹교신저자: (03760) 서울특별시 서대문구 이화여대길 52, 이화여자대학교 통계학과.

E-mail: shindw@ewha.ac.kr