

# A Bootstrap Lagrangian Multiplier Test for Market Microstructure Noise in Financial Assets

Hyo Jin Kim<sup>a</sup> · Dong Wan Shin<sup>a,1</sup> · Jonghun Park<sup>b</sup> · Sang-goo Lee<sup>c</sup>

<sup>a</sup>Department of Statistics, Ewha Womans University

<sup>b</sup>Department of Industrial Engineering, Seoul National University

<sup>c</sup>Department of Computer Science and Engineering, Seoul National University

(Received March 9, 2015; Revised March 21, 2015; Accepted March 25, 2015)

---

## Abstract

Stationary bootstrapping is applied to a Lagrangian multiplier (LM) test to test market microstructure noise (MMN) in financial asset prices. A Monte-Carlo experiment shows that the bootstrapping method improves the size of the original LM test which has some size distortion for conditional heteroscedastic models. The proposed test is illustrated for real data sets like KOSPI index and Won-Dollar exchange rate.

Keywords: Lagrangian multiplier test, market microstructure noise, mixed effect model, realized volatility.

---

## 1. 서론

관측된 금융 자산 가격 과정은 실제로는 관측할 수 없는 잠재 가격 과정과 잡음 과정의 두 부분으로 구성된다. 여기서 후자를 시장 미시구조 잡음(MMN; Market Microstructure Noise)이라 하는데 이는 비동시적 거래, 매수호가와 매도호가의 차이 등의 결과로 나타난다.

시장 미시구조 잡음은 로그 수익률에 계열 상관을 발생시키고 이로 인하여 금융 자산에 대한 통계적 관점에서 몇 가지 문제를 야기시킨다. 특히 실현변동성의 편의(bias)를 발생시키는데 추출 간격이 짧을수록 편의는 더 커진다. 또한 시간이 더 긴 추출 간격에 대한 실현변동성은 큰 분산을 가지게 된다. 따라서 시장 미시구조 잡음의 영향을 최소화하기 위하여 편의와 분산 간 균형을 고려한 최적 추출 간격을 결정하는 것이 중요하다. 이에 대해 Zhang 등 (2005), Bandi와 Russell (2008), Hansen과 Lunde (2006) 등은 최적 샘플링 간격에 대한 연구를 발표한 바 있다. 또 Oh와 Shin (2012)은 한국 주요 상장사 주가의 실현 변동성 추정 시 시장 미시구조 잡음과 최적 추출 빈도수의 관계를 연구한 바 있다. 또한 Bandi 등 (2008)과 McAleer과 Medeiros (2008) 등의 실현변동성에 대한 조사 연구도 살펴볼 필요가 있다.

추출 간격에 대한 최적 이론이 있으나 “5분~30분 사이의 추출 간격을 임의 선택하는 방법(임의 선택법)”이 단순하기 때문에 빈번하게 쓰인다. 임의 선택법의 근거는 추출 간격이 그 정도 되면 시장 미시구조 잡음으로 발생하는 로그 수익률의 계열 상관은 무시해도 될 정도이기 때문이다. 이러한 임의 선택

---

This research is supported by the Science Research Center program funded by Ministry of Education, Science and Technology (2011-0030811).

<sup>1</sup>Corresponding author: Department of Statistics, Ewha Womans University, Seoul 120-750, Korea.

E-mail: [shindw@ewha.ac.kr](mailto:shindw@ewha.ac.kr)

법은 문헌상에서도 드물지 않게 발견되는데, 예를 들면 Bandi 등 (2008), Giot 등 (2010), Busch 등 (2011)에서 임의 선택법의 적용 예를 볼 수 있다.

Shin과 Hwang (2015)은 시장 미시구조 잡음에 대한 라그랑지 승수 검정(이하 LM 검정)을 개발하였다. LM 검정은 임의 선택법에 통계적 타당성을 부여한다. LM 검정은 매우 단순하며 귀무가설 하에서의 극한 분포는 단순 카이 제곱 분포를 가진다. 그러나 몬테 카를로 실험은 제안된 검정이 몇 가지 현실성 있는 조건부 이분산 변동성 모형에 대하여 타당한 유의수준과 검정력을 보임에도 불구하고, 몇몇의 또 다른 현실성 있는 조건부 이분산 변동성 모형에 대해서는 유의수준 왜곡 문제를 보여 이를 개선할 여지가 있다.

일반적인 가설검정, 신뢰구간 등의 통계적 추론이 대표본 근사이론에 근거하여 구성되는 경우 유한 표본에서는 유의수준 왜곡, 신뢰도 왜곡 등의 문제를 내포하고 있는 경우가 빈번하고 이 문제에 대한 해결 방법으로 종종 부트스트래핑이 적용된다. 부트스트래핑을 이용하여 대표본 이론에 근거한 검정 및 신뢰구간을 개선하는 방법은 통계학의 모든 분야에서 널리 나타난다. 이 중 실현 변동성과 관련된 최근의 연구 결과를 몇 개 소개한다. Goncalves와 Meddahi (2009)과 Hwang과 Shin (2013)에서는 Barndorff-Nielsen과 Shephard (2002), Barndorff-Nielsen 등 (2006)의 실현변동성의 중심극한 정리에 근거한 적분변동성 신뢰구간이 유한 표본에서 왜곡을 가지며 부트스트래핑을 적용할 경우 왜곡 문제가 상당히 개선됨을 보였다. Hwang과 Shin (2014)은 Barndorff-Nielsen과 Shephard (2006)의 중심극한정리에 근거하여 실현변동성에 내포된 점프 성분 탐지 검정이 유한 표본에서 유의수준 왜곡을 가지며 부트스트래핑을 적용하였을 때 이것이 개선됨을 보인 바 있다.

본 논문에서는 Shin과 Hwang (2015)의 LM 검정의 유의수준 왜곡 문제의 개선책으로 Politis와 Romano (1994)가 고안한 정상적 부트스트래핑(stationary bootstrapping)을 이용하려 한다. 몬테 카를로 실험을 통해 LM 검정이 유의수준 문제를 완화시킴을 보이겠다. 또한 2008년부터 2012년까지의 KOSPI 지수에 정상적 부트스트래핑 LM 검정을 적용하여 실현 변동성의 샘플링 간격으로 11분이 적절함을 보이겠다. 추가로 같은 기간의 원-달러 환율에 위 방법을 적용하여 실현 변동성의 샘플링 간격으로 8분이 적절함을 보이겠다.

본 논문의 구성은 다음과 같다. 2절에서는 LM 검정에 대해 설명하고, 3절에서는 정상적 부트스트래핑에 LM 검정을 적용하며, 4절에서는 유한 샘플 모의 실험을 하고, 5절에서는 실증 분석을 통하여 실제 데이터에 적용한다.

## 2. LM 검정

어떤 자산의  $t$ 시점 로그 가격을  $\tilde{p}_t$ 라 하고 다음

$$d\tilde{p}_t = \mu_t dt + \sigma_t d\omega_t \quad (2.1)$$

과 같은 일반화 위너프로세스를 따른다고 가정하자. 이때  $\omega_t$ 는 표준 브라운 운동,  $\mu_t$ 는 평균과정이고,  $\sigma_t$ 는 변동성(volatility) 과정이다. 이 실제 로그과정은 시장 미시구조 잡음 때문에 직접 관측 못하는 잠재과정이며 관측된 로그 가격은  $p_t = \tilde{p}_t + u_t$ 와 같이 잡음이 부가된 형태로 나타낼 수 있다. 여기서  $u_t$ 는 *i.i.d.*  $N(0, \sigma_u^2)$ 를 따르며  $\{\omega_t\}$ 와 독립인 시장 미시구조 잡음이다. 이처럼 관측 로그 가격  $p_t$ 는 잠재적인 실제 로그 가격인  $\tilde{p}_t$ 와 측정 오차 과정인  $u_t$ 의 두 부분으로 구성된 단순 혼합 효과 모형이다.

우리의 관심은 시장 미시구조 잡음의 존재성을 따져보는 가설 검정에 있다. 이에 대한 귀무가설은

$$H_0 : \sigma_u^2 = 0$$

이다. 이 가설에 대해 Shin과 Hwang (2015)이  $\mu_t = \mu$ 와  $\sigma_t = \sigma$ 로 동질성을 가정한 후 LM 검정을 개발한 바 있다. 또 모의 실험을 통해 현실성 있는 여러 변동성 모형에서 적절한 유의수준과 검정력을 가짐을 보였다. 그러나 그 실험에서 보여준 바에 의하면 실제 조건부 이분산성을 갖는 중요한 변동성 모형에 대해 어느 정도의 유의수준 왜곡이 있음을 알 수 있다. 즉 변동성 과정이 GARCH(1, 1) 확산 모형과 로그 정규 확산 모형의 경우, 5% 정도의 명목 유의수준 하에서 실제 유의수준은 10% 정도가 되는 상황이 보고되었다.

본 논문에서는  $T$ 일 동안의 로그 가격 데이터의 표본이 있는 경우를 다룬다. 이때  $\Delta t = 1$ 은 단위 시간의 간격(길이)을 나타내며 보통 하루이다. 또한  $t = 0, 1, 2, \dots, (T - 1)$ 는 관측 시작 시점을 나타내며 보통 관측일을 의미한다. 일자별 표본은  $n$ 개의 같은 시간 간격으로 구성되어 있다고 가정한다. 그러면 관측된 로그 가격의 집합은

$$\left\{ p_{t+\frac{0}{n}}, p_{t+\frac{1}{n}}, \dots, p_{t+\frac{n}{n}} \right\}, \quad t = 0, 1, 2, \dots, (T - 1)$$

으로 나타낼 수 있다.  $t$ 일에  $i$ 번째로 관측된 실제 로그가격을  $p_{t+i/n}$ 라 하면, 실제 로그 수익률  $\tilde{r}_{ti} = \tilde{p}_{t+i/n} - \tilde{p}_{t+(i-1)/n}$ 은 직접 관측되지 못하고 1차 이동 평균 오차  $v_{ti} = u_{t+i/n} - u_{t+(i-1)/n}$ 와 혼합되어  $r_{ti} = \tilde{r}_{ti} + v_{ti}$ 처럼 나타나며 관측된 로그 수익률 자료 집합은 다음과 같이 표현된다.

$$\left\{ r_{ti} = p_{t+\frac{i}{n}} - p_{t+\frac{i-1}{n}}, \quad i = 2, 3, \dots, n, \quad t = 0, 1, \dots, (T - 1) \right\},$$

여기서  $i = 1$ 에 대응하는 로그수익률  $r_{t1}$ 은 포함하지 않았음을 유의하자. 그 이유는 실현변동성 계산 시 보통 하룻밤 사이의(overnight) 로그수익률, 즉 전날 종가와 당일 시가 차이에 대한 로그수익률은 반영하지 않기 때문이다.

Shin과 Hwang (2015)은  $n$ 과  $T$ 가 클 경우, LM 검정의 검정 통계량은 다음의 단순 통계량과 같음을 보였다.

$$LM = \frac{nT^{-1} \left( \sum_{t=0}^{T-1} \sum_{i=3}^n r_{ti} r_{t,i-1} \right)^2}{\hat{\sigma}^4}, \quad \hat{\sigma}^2 = \frac{\sum_{t=0}^{T-1} \sum_{i=2}^n r_{ti}^2}{T}. \quad (2.2)$$

LM 검정은  $\sigma_u^2$ 이 0에 가까운 곳에서 국지적으로 최강 검정이다. Shin과 Hwang (2015)은 LM 검정이 1차 상관계수의 합으로 구성된 단순 함수로 표현됨을 주목하였다. 즉

$$\bar{\rho}_P = \frac{\sum_{t=0}^{T-1} \sum_{i=3}^n r_{ti} r_{t,i-1}}{\sum_{t=0}^{T-1} \sum_{i=2}^n r_{ti}^2}$$

이라 하면,  $LM = nT\bar{\rho}_P^2$ 가 된다. 또 큰  $n$ 과  $T$ 에 대한 LM의 귀무가설 극한 분포는 자유도가 1인 카이 제곱 분포임을 증명하였다.

### 3. 부트스트래핑 검정

우리는 LM 검정이  $r_{ti}$ 와  $r_{t,i-1}$ 의 계열 상관을 확인하는 것임에 주목하였다. 그러므로 부트스트랩 표본은 부트스트랩 LM 검정이 충분한 검정력을 확보하기 위하여  $r_{ti}$ 와  $r_{t,i-1}$ 사이의 상관 구조를 유지해야 한다. 그래서 *i.i.d.* 부트스트래핑은 적절하지 못하며 블록 부트스트래핑이 방법이 보다 적절하다. 우리는 Politis와 Romano (1994)에서 제안된 가장 널리 쓰이는 블록 부트스트래핑의 하나인 정상적 부트스트래핑 방법을 사용하고자 한다.

### 3.1. 정상적 부트스트래핑 LM 검정

먼저 정상적 부트스트래핑 방법을 살펴본다. 전체 로그수익률 표본은  $\{r_{t2}, \dots, r_{tn}, t = 1, \dots, T\}$ 이다. 각  $t$ 일의 부표본

$$\{r_{t2}, \dots, r_{tn}\} \quad (3.1)$$

에 대하여 Politis와 Romano (1994)의 정상적 부트스트래핑을 적용하면  $t$ 일에 대한 정상적 부트스트래핑 부표본

$$\{r_{t2}^*, \dots, r_{tn}^*\} \quad (3.2)$$

을 얻을 수 있다. 대략적으로 말하면 정상적 부트스트래핑 부표본 (3.2)는  $t$ 일의 부표본 (3.1)로부터 기하분포를 따르고 크기  $L_k$ 를 갖는 블록  $\tau$ 개를 랜덤하게 뽑아 결합한 것이다. 여기서  $k = 1, \dots, \tau$ 이다. 자세한 설명은 아래 3.2절에 기술하였다. 이 과정을 통해 우리는 전체  $T$ 개의 정상적 부트스트래핑 부표본  $\{r_{t2}^*, \dots, r_{tn}^*, t = 1, \dots, T\}$ 을 결합하여 다음의 정상적 부트스트래핑 표본을 얻었다.

$$\{r_{t2}^*, \dots, r_{tn}^*, t = 1, \dots, T\}$$

그러면 LM의 부트스트랩 버전은

$$\begin{aligned} LM^* &= \frac{nT^{-1} \left( \sum_{t=0}^{T-1} \sum_{i=3}^n r_{ti}^* r_{t,i-1}^* - \sum_{t=0}^{T-1} \sum_{i=3}^n r_{ti} r_{t,i-1} \right)^2}{\hat{\sigma}^{*4}} \\ \hat{\sigma}^{*2} &= \frac{\sum_{t=0}^{T-1} \sum_{i=2}^n r_{ti}^{*2}}{T} \end{aligned} \quad (3.3)$$

이다. 위 식 (3.3)에서  $X_i = \sum_{t=0}^{T-1} r_{t,i-1} r_{ti}$ 를 분자에서 빼준 이유는 다음과 같다.

먼저  $X_i^* = \sum_{t=0}^{T-1} r_{t,i-1}^* r_{ti}^*$ 라 하자. 이때  $X_i^*, \dots, X_n^*$ 는

$$E^* X_i^* = X_i, \quad \text{즉} \quad E^* \sum_{t=0}^{T-1} r_{ti}^* r_{t,i-1}^* = \sum_{t=0}^{T-1} r_{ti} r_{t,i-1}; \quad \text{Var}^*(X_i^*) = \hat{\sigma}^2$$

를 만족하는 *i.i.d.* 확률변수이다. 여기서  $E^*$ 와  $\text{Var}^*$ 는 각각  $\{r_{t2}, \dots, r_{tn}, t = 1, \dots, T\}$ 가 주어진 상태에서의 조건부 기댓값, 조건부 분산을 나타낸다. 또  $\hat{\sigma}^{*2}$ 와  $\hat{\sigma}^2$ 은 모두  $\sigma^2$ 에 대해서 일치성을 지니므로  $T$ 와  $n$ 이 모두 무한대로 갈 때,  $LM^*$ 의 극한분포는 자유도가 1인 카이 제곱 분포이다. 이로써 귀무가설 하에서  $LM^*$ 의 분포는 LM의 분포를 일치성 있게 추정함을 알 수 있다.

위에서 논의한 부트스트랩 분포의 일치성으로부터 부트스트랩 검정을 구성할 수 있다. 주어진 표본  $\{r_{ti}, i = 2, \dots, n, t = 1, \dots, T\}$ 에 대하여, 명목 수준  $\alpha \in (0, 1)$ , 정수  $B > 0$ 가 주어지면  $B$ 개의 부트스트랩 LM 검정  $\{LM^{*(1)}, \dots, LM^{*(B)}\}$ 을 구성한다. 여기서  $LM^*(\alpha)$ 을  $B$ 개의 부트스트랩 LM값에 대한  $(1 - \alpha)$  실증 분위수로 정의하면, 명목 수준  $\alpha$ 에서의 LM 검정은

$$LM > LM^*(\alpha) \quad (3.4)$$

인 경우 귀무가설  $H_0$ 를 기각하는 것이다. 대립가설 하에서는 시장 미시구조 잡음 때문에  $\rho_P \neq 0$ 이고 LM은 무한대로 발산하므로 위 식 (3.4)의 확률이 1로 가며 따라서 부트스트랩 검정은 검정 일치성을 가짐을 보일 수 있다.

### 3.2. 정상적 부트스트래핑 방법

부표본 (3.1)로부터 정상적 부트스트래핑 부표본 (3.2)를 구성하는 방법을 설명한다. 먼저 다음과 같은 관측 데이터 집합의 주기 확장으로 얻은 새로운 시계열  $\{r_{ti} : i \geq 1\}$ 를 정의하자. 이때 각  $i \geq 1$ 에서 어떤 정수  $q$ 에 대하여  $i = qn + j$  조건을 만족하는  $r_{ti} := r_{tj}$ 를 정의한다. 그러면 (3.1)로부터 확장된 데이터 계열

$$\{r_{t2}, r_{t3}, \dots, r_{tn}, r_{t2}, r_{t3}, \dots\} \quad (3.5)$$

을 얻는다. 확장 계열 (3.5)로부터 임의의 위치  $i \in \{2, \dots, n\}$ 에서부터 크기  $\ell$ 인 블록

$$B(i, \ell) = \{r_{ti}, \dots, r_{t(i+\ell-1)}\}$$

을 랜덤하게 추출한다. 이때 블록 시작 위치  $i$ 는  $\{2, \dots, n\}$ 에서의 일양분포를 따르고 블록 크기  $\ell$ 은 확률  $p$ 를 갖는 기하분포를 따르는 임의의 난수라고 하자. 이를 위해 다음의 세 가지 성질을 만족하는 확률 변수  $I_1, I_2, \dots$ 와  $L_1, L_2, \dots$ 를 생성한다.

- (i)  $I_1, I_2, \dots$ 는  $\{2, \dots, n\}$ 사이의 값을 가지며  $P(I_1 = i) = 1/(n-1)$ ,  $i = 2, \dots, n$ 를 만족하는 *i.i.d.* 이산 균등분포를 따른다.
- (ii)  $L_1, L_2, \dots$ 는 모수  $p \in (0, 1)$ 와  $P(L_1 = \ell) = p(1-p)^{\ell-1}$ ,  $\ell = 1, 2, \dots$ 를 만족하는 기하 분포를 따르는 *i.i.d.* 확률 변수이다.
- (iii)  $\{I_1, I_2, \dots\}$ 와  $\{L_1, L_2, \dots\}$ 는 서로 독립이다.

이제  $\tau = \inf\{k \geq 1 : L_1 + \dots + L_k \geq n\}$ 이라 하자. 그 뒤  $\tau$ 개의 블록  $B(I_1, L_1), \dots, B(I_\tau, L_\tau)$ 을 선택한다.  $L_1 + \dots + L_\tau$ 개의 로그수익률이 블록  $B(I_1, L_1), \dots, B(I_\tau, L_\tau)$ 에서 다시 추출된 것에 주목하자. 여기서 마지막  $L_1 + \dots + L_\tau - (n-1)$  로그수익률을 제거하여  $(n-1)$ 개로 이루어진 부트스트랩 관측치  $\{r_{t2}^*, \dots, r_{tn}^*\}$ 을 얻을 수 있다.  $\{r_{t2}, \dots, r_{tn}\}$  조건 하에서  $\{r_{ti}^*, i = 2, \dots\}$  과정은 정상성을 가진다.

## 4. 몬테 카를로 연구

Shin과 Hwang (2015)에서 제안한 LM 검정은 변동성이 일정하다는 가정 하에서 구성된 것이기에 어떤 조건부 이분산성 하에서는 유의수준 왜곡을 지님이 보고되었다. 본 절에서는 부트스트랩 검정이 현실성 있는 이질적 변동성 과정 하에서 LM 검정이 갖는 유의수준 왜곡문제를 개선한다는 점을 보이고자 한다. 이를 위해  $\omega_{1t}, \omega_{2t}, \omega_{3t}$ 가 독립적인 표본 브라운 운동을 따르는 모형

$$d\tilde{p}_t = \mu dt + \sigma_t \left( \rho_1 d\omega_{1t} + \rho_2 d\omega_{2t} + \sqrt{1 - \rho_1^2 - \rho_2^2} d\omega_{3t} \right)$$

를 이용하여 살펴본다. 동질적 변동성 과정

$$V_0 : d\sigma_t^2 = dt,$$

뿐만 아니라, 다음과 같은 세 종류의 비동질적 변동성 과정도 고려한다.

$$V_1 : d\sigma_t = 0.04(0.64 - \sigma_t^2) dt + 0.14\sigma_t^2 d\omega_{1t},$$

$$V_2 : d \log \sigma_t^2 = -0.014(0.84 + \log \sigma_t^2) dt + 0.11 d\omega_{1t},$$

$$V_3 : \sigma_t^2 = \sigma_{1t}^2 + \sigma_{2t}^2,$$

**Table 4.1.** Empirical size(%) of the LM test at  $\sigma_u = 0$ . Nominal size = 5%.

$\sigma_u$	$T$	$n$	$V_0$		$V_1$		$V_2$		$V_3$	
			LM	LM*	LM	LM*	LM	LM*	LM	LM*
Model with no trend and no leverage										
.00	200	20	4.2	4.9	9.4	5.8	9.2	5.0	4.7	4.9
.00	200	100	5.0	6.9	9.3	6.6	12.1	6.8	5.8	6.8
.00	200	500	6.0	7.1	11.0	7.5	12.3	6.6	6.7	7.2
.00	1000	20	4.8	1.3	8.4	2.6	10.3	2.6	5.2	2.0
.00	1000	100	5.3	5.3	9.0	6.2	11.1	5.3	5.6	6.2
.00	1000	500	3.8	4.8	9.1	5.5	12.1	6.1	4.0	4.9
Model with trend and leverage										
.00	200	20	5.7	5.8	8.4	5.9	9.5	6.8	5.2	5.2
.00	200	100	4.5	5.3	7.8	5.9	10.7	6.8	5.3	6.1
.00	200	500	4.1	5.3	8.8	5.5	11.7	6.2	6.2	7.2
.00	1000	20	5.4	1.9	9.6	2.5	11.8	3.0	5.3	1.7
.00	1000	100	3.6	4.1	7.3	4.6	10.2	6.0	4.6	4.2
.00	1000	500	5.2	6.2	7.7	5.2	10.6	5.6	5.2	5.8

※ 반복 수 = 1,000, 부트스트랩 반복 수  $B = 100$ ;

$V_0 =$  동질적 변동성,  $V_1 =$  GARCH(1, 1) 확산 변동성,  $V_2 =$  로그 정규 확산 변동성,  $V_3 =$  두 요인 확산 변동성

여기서

$$\begin{aligned} d\sigma_{1t}^2 &= 0.57(0.33 - \sigma_{1t}^2)dt + 0.23\sigma_{1t}^2 d\omega_{1t}, \\ d\sigma_{2t}^2 &= 0.08(0.18 - \sigma_{2t}^2)dt + 0.11\sigma_{2t}^2 d\omega_{2t}. \end{aligned}$$

위 모형에서  $V_1, V_2, V_3$ 를 각각 GARCH(1, 1) 확산 모형, 로그 정규 확산 모형, 두 요인 확산 모형이라 부른다. 이러한 변동성 모형은 Bollerslev와 Zhou (2002), Andersen 등 (2002), Goncalves와 Meddahi (2009), Hwang과 Shin (2013, 2014), Shin과 Hwang (2015) 등의 실증 분석에 쓰인 바 있다.

본 논문의 몬테 카를로 연구에서는 추세가 있는 모형과 없는 모형 모두를 고려하였다. 추세가 없는 모형은  $\mu = 0$ 인 경우이며, 이때  $(\rho_1, \rho_2) = (0, 0)$ 로 하여  $\tilde{p}_t$ 의 레버리지(leverage)가 배제된 형태를 부여하였다. 추세가 있는 모형은  $\mu = 0.33$ 이며, 동질적 변동성, GARCH(1, 1) 확산 변동성, 로그 정규 확산 변동성, 두 요인 확산 변동성에 대해 각각

$$(\rho_1, \rho_2) = (-0.6, 0), (-0.6, 0), (-0.3, -0.3), (-0.3, -0.3).$$

조건을 지정하여  $\tilde{p}_t$ 에 레버리지를 부여하였다. 또 GARCH(1, 1) 변동성, 로그 정규 변동성, 두 요인 확산 변동성에 대해서는 변동성의 초기값을 지정해 주어야 하는데 이에 대해서는 위 모형 순서대로  $\sigma_0^2 = 0.64$ ,  $\log \sigma_0^2 = -0.84$ ,  $(\sigma_{10}^2, \sigma_{20}^2) = (0.33, 0.18)$ 를 고려하였다. 모든 정규 오차는 IMSL의 FORTRAN 서브루틴, RNNOA로 생성하였고 이산 균등난수, 기하난수는 RNUND, RNGEO를 사용하여 발생시켰다.

5% 수준에서 LM 검정의 기각 빈도는 Table 4.1과 Table 4.2에 있다. 이때 각각 조합의 실험 조건은  $n = 20, 100, 500$ ,  $T = 200, 1000$ ,  $\sigma_u = 0, 0.01, 0.02$ 이며 실험 반복수는 1,000, 부트스트랩 반복 수는  $B = 100$ 로 하였다. 블록 크기 모수는  $p = 0.05(n/100)^{-1/3}$ 로 하여  $n = 20, 100, 500$ 일 때, 평균 블록 크기는  $p^{-1} = 11.7, 20.0, 34.2$ 가 된다.

**Table 4.2.** Empirical size(%) of the LM test at  $\sigma_u = 0.01, 0.02$ . Nominal size = 5%.

$\sigma_u$	$T$	$n$	$V_0$		$V_1$		$V_2$		$V_3$	
			LM	LM*	LM	LM*	LM	LM*	LM	LM*
Model with no trend and no leverage										
.01	200	20	4.8	4.9	8.5	4.8	9.0	5.4	5.3	5.0
.01	200	100	5.8	6.4	9.0	6.8	12.0	7.6	6.6	6.5
.01	200	500	13.8	15.0	23.0	17.9	31.3	21.4	29.0	30.5
.01	1000	20	4.6	1.6	9.7	3.4	12.2	2.6	5.2	1.5
.01	1000	100	6.0	5.0	10.5	5.7	15.2	6.7	9.1	6.2
.01	1000	500	34.0	35.1	66.5	58.9	77.3	67.6	85.7	85.4
.02	200	20	6.2	5.2	8.2	5.9	10.1	5.4	6.4	5.1
.02	200	100	10.0	10.4	16.9	12.5	22.9	15.2	19.4	18.7
.02	200	500	79.1	80.8	98.2	97.0	99.3	98.4	100.0	100.0
.02	1000	20	6.2	0.9	10.4	2.0	14.6	2.7	8.3	0.9
.02	1000	100	21.3	17.4	51.1	34.6	61.5	43.9	67.9	59.4
.02	1000	500	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Model with trend and leverage										
.01	200	20	4.3	4.9	7.6	5.2	8.9	5.9	4.8	4.4
.01	200	100	4.8	5.9	8.4	6.2	12.0	6.8	6.2	7.4
.01	200	500	8.8	11.3	21.4	16.5	31.2	21.8	27.1	29.0
.01	1000	20	5.0	2.8	9.3	3.2	12.3	2.8	5.2	2.2
.01	1000	100	6.2	6.2	11.4	6.8	14.0	5.9	7.3	5.3
.01	1000	500	31.9	34.6	65.3	57.5	77.3	65.8	83.9	83.9
.02	200	20	6.2	5.4	8.0	4.9	9.8	5.8	7.0	6.3
.02	200	100	8.1	8.4	15.8	11.1	22.2	13.7	19.1	18.4
.02	200	500	79.4	79.4	97.7	96.4	99.1	98.0	100.0	100.0
.02	1000	20	6.1	1.8	12.1	1.8	13.8	2.0	7.5	1.5
.02	1000	100	21.4	17.5	47.5	32.4	58.0	41.2	65.9	56.1
.02	1000	500	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

※ 반복 수 = 1,000, 부트스트랩 반복 수  $B = 100$ ;

$V_0 =$  동질적 변동성,  $V_1 =$  GARCH(1, 1) 확산 변동성,  $V_2 =$  로그 정규 확산 변동성,  $V_3 =$  두 요인 확산 변동성

Table 4.1의  $\sigma_u = 0$ 에서 유의수준 결과를 보자. 먼저 추세가 없는 모형과 레버리지가 없는 모형을 통해, GARCH(1, 1) 확산 모형( $V_1$ )과 로그 정규 확산 모형( $V_2$ )에서 LM 검정의 유의수준이 왜곡됨을 주목하자. 이 경우 LM 검정의 유의수준은 9.0%에서 12.3% 사이로 나타났다. 반면 부트스트랩 검정을 이용하면 유의수준 왜곡이 완화되고 있음을 볼 수 있다. 이때 LM\*의 유의수준은 2.6%에서 7.5% 사이로 나타났다. 서로 다른 두 변동성 모형, 동질적 모형( $V_0$ )과 두 요인 확산 모형( $V_3$ )에서는 LM\*가 여전히 수용할 수 있는 유의수준을 가짐을 볼 수 있다. 추세와 레버리지가 있는 모형에서도 추세가 없는 모형에서와 비슷한 양상을 보인다. 즉, GARCH(1, 1) 확산 모형( $V_1$ )과 로그 정규 확산 모형( $V_2$ )에서 LM은 유의수준 왜곡을 보이는데 LM\*는 개선된 유의수준을 갖고 또 다른 두 변동성 모형, 동질적 모형( $V_0$ )과 두 요인 확산 모형( $V_3$ )에서 LM\*는 받아 들일만한 유의 수준을 갖는다.

Table 4.2의  $\sigma_u = 0.01, 0.02$ 에서의 검정력 결과를 보자. 전체 8개의 모형 중에서  $n, T, \sigma_u$ 가 각각 증가함에 따라 검정력이 유사하게 증가함을 알 수 있다. 이는 LM 검정과 LM\* 검정 모두 비슷한 검정력을 가짐을 의미한다. 어떤 경우 겉보기에는 LM의 경우가 더 높은 검정력 갖는 것처럼 보이지만 이는 LM의 높은 유의수준 왜곡에 기인하는 것이다. 이러한 유의수준 차이가 보정된다면 LM의 검정력과

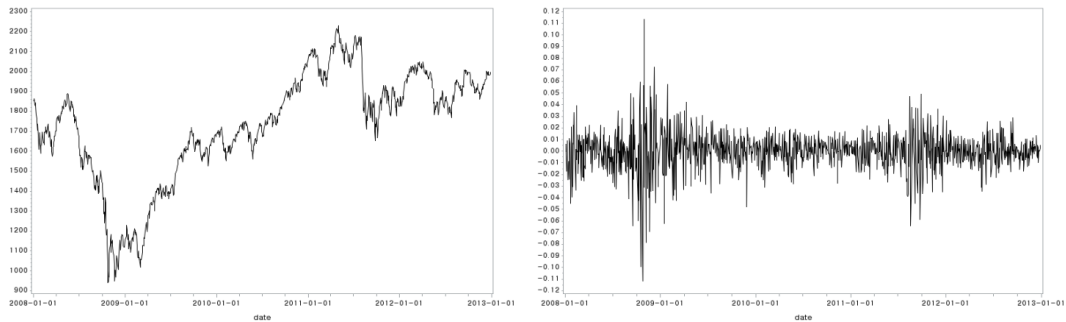


Figure 5.1. Daily KOSPI and its log return

LM\*의 검정력은 비슷해질 것이다.

이러한 유한 표본 연구를 통하여 부트스트랩 LM\*이 조건부 이분산 모형에서 LM이 지니고 있는 유의수준 문제를 의미 있게 개선했다고 볼 수 있다.

## 5. 실증 분석

KOSPI 지수와 원-달러 환율의 2008년부터 2012년까지의 고빈도 데이터에 근거하여 일별 실현변동성을 계산하기 위하여 샘플링 간격을 정할 때, 위에서 제안한 부트스트랩 LM 검정이 유용하게 사용됨을 보이고자 한다.

먼저 실현변동성을 2절의 기호를 사용하여 기술한다. 하루 관측치 중 샘플링을  $n$  번 한다고 할 때, 2008년부터 2012년 중 어떤 날  $t$  일에서의 실현변동성(RV)은 다음 식처럼 표현할 수 있다.

$$RV_t = \sqrt{\sum_{i=2}^n r_{ti}^2}.$$

위의 실현변동성을 계산하기 위해서는  $n$ 이 지정되어야 한다. 이를 위해서 부트스트랩 LM 검정을 사용하였다.

### 5.1. KOSPI 지수

KOSPI 자료의 전체 관측 일수는 1,248일이며 하루 동안의 최소 관측 간격은 1분이다. 본 논문에서는 거래소가 열리는 오전 9시부터 오후 3시까지, 총 6시간 동안의 데이터만을 이용하여 실현변동성을 계산한다.

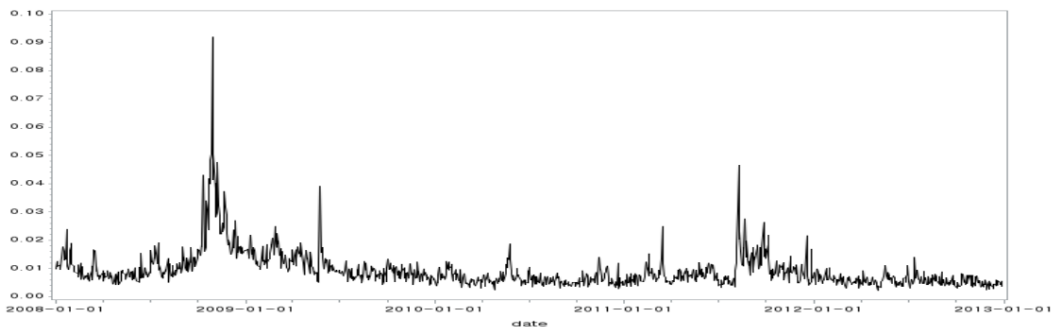
실현변동성 계산에 앞서 KOSPI 지수와 수익률의 개요를 살펴본다. Figure 5.1은 일별 증가(좌)와 일별 로그수익률(우)의 시도표이다. 왼쪽에 있는 일별 KOSPI 지수의 시도표를 살펴보면 2008년 5월부터 급격히 하락하다 2008년 10~11월에는 지수가 1000 이하로 내려간 날도 있다. 그 이후 1년에 걸쳐 2008년 1월 수준을 회복하는 것을 볼 수 있다. 2011년 8월부터 잠시 하락세를 보이다 2012년 말까지 증감을 반복하고 있다. 오른쪽에 있는 일별 로그수익률의 특징은 평균은 거의 0으로 일정하지만 변동폭은 크게 움직이는 구간과 비교적 작게 움직이는 구간으로 뚜렷하게 구분되는 것을 확인할 수 있다. 특히 2008년 9~11월과 2011년 8~11월 사이에 변동폭이 크게 나타났다.



**Table 5.1.** LM test results for KOSPI (2008.01.01.~2012.12.31.)

$h$ , sampling interval (min)	LM	$p$ -value%	bootstrap $p$ -value%	$h$ , sampling interval (min)	LM	$p$ -value%	bootstrap $p$ -value%
1	1095.8	.0	.0	9	18.2	.0	9.9
2	1172.2	.0	.0	10	11.8	.1	9.2
3	389.9	.0	.0	11	11.3	.1	13.4
4	53.5	.0	.3	12	4.3	3.9	32.0
5	.7	39.1	72.8	13	4.6	3.2	32.4
6	9.8	.2	27.5	14	1.1	28.6	57.3
7	22.9	.0	5.4	15	5.8	1.6	27.7
8	16.5	.0	8.1	16	.1	81.4	89.8

※ 부트스트랩 반복 수  $B = 1000$



**Figure 5.2.** Realized volatility for KOSPI based on 11 minute sampling

3.1절에서 제안한 부트스트랩 LM 검정을 KOSPI 지수에 적용해 보았다. Table 5.1은 KOSPI 지수 데이터에 적용한 부트스트랩 LM 검정 결과를 나타낸다. 각 열은 표본 추출 간격  $h = 1, 2, \dots, 16$ (분)에서 LM값, 유의확률, 부트스트랩 유의확률이다. 보통 “임의 선택법”을 사용할 때 9.5분과 같이 소수점 이하를 포함하는 샘플링 간격은 사용하지 않기 때문에, 우리는 정수 시간 간격에 대해서만 고려하였다. Table 5.1에서 10% 수준의 LM 검정을 적용한다면 5, 14, 16분의 추출 간격 중 하나를 선택하면 된다. 이들 중 어느 것을 택할지에 대해서는 추가적인 조사가 필요할 것이다. 또한 부트스트랩 LM 검정에 의하면, 추출 간격 11분이 적절해 보인다. Figure 5.2에 KOSPI 지수의 11분 간격 샘플링에 근거한 일별 실현변동성 시도표가 나타나 있다.

### 5.2. 원-달러 환율

원-달러 환율 데이터의 관측 일수는 1,249일이며 하루 중 최소 관측 간격은 1분이다. KOSPI 지수와 마찬가지로 개장 시간(6시간) 동안의 데이터만을 이용하여 실현 변동성을 계산하고자 한다.

먼저 원-달러 환율 데이터의 개요를 살펴본다. Figure 5.3은 일별 증가(좌)와 일별 로그수익률(우)의 시도표이다. 왼쪽에 있는 일별 원-달러 환율의 시도표를 살펴보면 2008년 1월부터 차츰 증가하다 2009년 3월 1600원에 가까운 최고점을 기록하며 2010년까지 천천히 하락하는 움직임을 보인다. 이는 KOSPI 지수의 급락과 마찬가지로 미국발 금융위기의 영향을 받은 것으로 해석할 수 있다. 오른쪽에 있는 일별 로그수익률의 특징은 평균은 거의 0으로 일정하지만 변동폭은 크게 움직이는 구간과 비교적 작게 움직이는 구간으로 뚜렷하게 구분되는 것을 확인할 수 있다. 특히 2008년 9~11월과 2010년 3~5월, 2011년

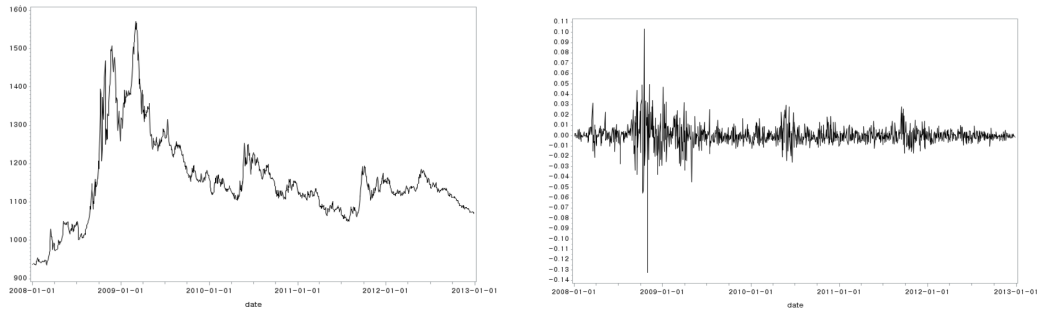


Figure 5.3. Won-Dollar exchange rate and its log return (Daily closing price (left), Daily log return (right))

Table 5.2. LM test results for won-dollar exchange rate (2008.01.01.~2012.12.31.)

$h$ , sampling interval (min)	LM	$p$ -value%	bootstrap $p$ -value%	$h$ , sampling interval (min)	LM	$p$ -value%	bootstrap $p$ -value%
1	11197.3	.0	.0	9	389.2	.0	18.7
2	914.4	.0	.2	10	269.5	.0	19.0
3	405.6	.0	.0	11	39.3	.0	10.0
4	174.5	.0	9.8	12	13.8	.0	25.3
5	147.4	.0	2.8	13	2.3	12.6	70.4
6	161.5	.0	.7	14	.3	57.1	87.5
7	254.6	.0	3.0	15	.0	85.9	94.1
8	263.2	.0	10.4	16	.9	34.0	73.5

※ 부트스트랩 반복 수  $B = 1000$

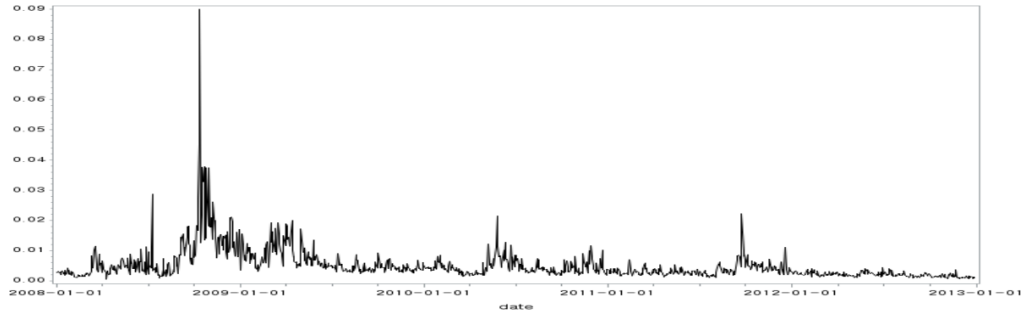


Figure 5.4. Realized volatility for Won-Dollar exchange rate on 8 minute sampling

8~11월 사이에 변동폭이 크게 나타났다.

Table 5.2의 LM 검정 결과에서 10% 수준을 적용한다면, 13분의 추출 간격을 선택하면 된다. 또 부트스트랩 LM 검정 결과에서 10% 수준을 적용한다면 8분의 추출 간격이 적절하다고 할 수 있을 것이다. Figure 5.4에 8분 간격에 근거한 일별 실현 변동성 시도표가 나타나 있다.

### 6. 결론 및 논의

본 연구에서는 기존 LM 검정의 한계점인 조건부 이분산 모형에서의 유의수준 왜곡 문제를 개선하기 위해 부트스트래핑 방법을 적용하였다. 로그 수익률간의 상관 구조를 유지하는 정상적 부트스트래핑 표

본을 이용하여 개선된 부트스트래핑 LM 검정을 제안하였다. 다음으로, KOSPI와 대미 원-달러 환율의 실현변동성에 대해 부트스트래핑 LM 검정을 이용하여 최적 샘플링 간격을 선택하는 것을 예시하였다. KOSPI 지수의 경우 10% LM 검정을 적용하면 5, 14, 16분이 선택되는데 비해 부트스트래핑 LM 검정에 의하면 11분 간격이 선택되었다. 또한 원-달러 환율의 경우 10% LM 검정을 적용하면 13분이 선택되거나 부트스트래핑 LM 검정에 의하면 8분이 선택되었다.

## References

- Andersen, T. G., Benzoni, L. and Lund, J. (2002). An empirical investigation of continuous-time equity return models, *The Journal of Finance*, **57**, 1239–1284.
- Bandi, F. M. and Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling, *Review of Economic Studies*, **75**, 339–369.
- Bandi, F. M., Russell, J. R. and Yang, C. (2008). Realized volatility forecasting and option pricing, *Journal of Econometrics*, **147**, 34–46.
- Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J. and Shephard, N. (2006). Limit theorems for bipower variation in financial econometrics, *Econometric Theory*, **22**, 677–719.
- Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society: Series B*, **64**, 253–280.
- Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation, *Journal of Financial Econometrics*, **4**, 1–30.
- Bollerslev, T. and Zhou, H. (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility, *Journal of Econometrics*, **109**, 33–65.
- Busch, T., Christensen, B. J. and Nielsen, M. O. (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets, *Journal of Econometrics*, **160**, 48–57.
- Giot, P., Laurent, S. and Petitjean, M. (2010). Trading activity, realized volatility and jumps, *Journal of Empirical Finance*, **17**, 168–175.
- Goncalves, S. and Meddahi, N. (2009). Bootstrapping realized volatility, *Econometrica*, **77**, 283–306.
- Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise, *Journal of Business and Economic Statistics*, **24**, 127–161.
- Hwang, E. and Shin, D. W. (2013). Stationary bootstrapping realized volatility, *Statistics and Probability Letters*, **83**, 2045–2051.
- Hwang, E. and Shin, D. W. (2014). A bootstrap test for jumps in financial economics, *Economics Letters*, **125**, 74–78.
- McAleer, M. and Medeiros, M. C. (2008). Realized volatility: A review, *Econometric Reviews*, **27**, 10–45.
- Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap, *Journal of the American Statistical Association*, **89**, 1303–313.
- Shin, D. W. and Hwang, E. (2015). Lagrangian multiplier test for market microstructure noise with applications to sampling interval determination for realized volatilities, *Economics Letters*, Accepted.
- Zhang, L., Mykland, P. A. and Ait-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association*, **100**, 1394–1411.

# 금융자산의 시장 미시구조 잡음에 대한 부트스트래핑 라그랑지 승수 검정

김효진<sup>a</sup> · 신동완<sup>a,1</sup> · 박종현<sup>b</sup> · 이상구<sup>c</sup>

<sup>a</sup>이화여자대학교 통계학과, <sup>b</sup>서울대학교 산업공학과, <sup>c</sup>서울대학교 컴퓨터공학과

(2015년 3월 9일 접수, 2015년 3월 21일 수정, 2015년 3월 25일 채택)

## 요약

본 논문에서는 정상적 부트스트래핑을 금융 자산 가격에서 시장 미시구조 잡음에 대한 라그랑지 승수 검정에 적용한다. 몬테 카를로 실험을 통해 부트스트래핑 방법이 조건부 이분산 모형을 적용한 기존 라그랑지 승수 검정의 유의수준 왜곡 문제를 개선함을 보인다. 이 검정을 KOSPI 지수와 원-달러 환율과 같은 실제 데이터에 적용한다.

주요용어: 라그랑지 승수 검정(Lagrangian multiplier test), 시장 미시구조 잡음, 혼합 효과 모형, 실현 변동성.

이 연구는 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (과제번호: 2011-0030811).

<sup>1</sup>교신저자: (120-750) 서울특별시 서대문구 대현동 11-1, 이화여자대학교 통계학과.

E-mail: shindw@ewha.ac.kr