先物硏究 제22권 제1호(2014년 02월) 117~139

고빈도 자료를 활용한 독성 주문 흐름 측정 : KOSPI200 지수 선물 시장

옥 기 율*
오 명**

<초 록>

본 연구의 목적은 새로운 정보 거래 확률 측정치인 VPIN(Volume-Synchronized Probability of Informed Trading)을 사용하여 고빈도 KOSPI200 지수 선물 시장의 독성 주문 흐름을 확인하는 것이다. VPIN은 PIN(Probability of Information-based Trading)과 같이 최우추정법으로 관찰 불가능한 모수들을 추정할 필요가 없고 거래량 시간(volume time)을 기준으로 구성하기 때문에 고빈도 시장에 적용하기 용이하다. 고빈도 KOSPI 200 지수 선물 시장의 VPIN과 선물 시장 변동성 간의 관계를 상관계수와 조건부 확률 분포로 분석하였는데, KOSPI200 지수 선물 VPIN이 지속적으로 높을 때, 선물 시장 변동성이 커질 가능성이 높은 것으로 나타났다. 이는 VPIN이 독성 주문 흐름으로 인한 선물 시장 변동성 예측에 유용함을 나타낸다.

주제어 : KOSPI200 지수 선물, 고빈도 자료, VPIN, 독성 주문 흐름, 변동성

논문접수일 : 2013년 10월 21일, 수정일 : 2014년 01월 13일, 게재확정일 : 2014년 02월 04일

* **교신저자**, 부산대학교 경영대학 경영학과, **주소** : 609-735, 부산광역시 금정구 장전동 산 30번지 부산대학교 경영대학, **E-mail** : kyohk@pusan.ac.kr, **전화** : 051-510-2552.

** 부산대학교 경영대학 경영학과, **E-mail** : wuming5@pusan.ac.kr.

Ⅰ. 서 론

현대 재무이론에 있어서 대부분의 시장 미시구조 모형(market microstructure models)들은 투자자들 간에 정보 비대칭(information asymmetry) 현상이 존재한다고 가정한다.[1] 정보 비대칭 하에서 투자자들을 정보 거래자와 비정보 거래자로 나눌 수 있고, 이러한 비대칭적 정보로 인해 주문 흐름 불균형(order flow imbalance) 현상과 비정보 거래자들의 역선택(adverse selection) 문제 등이 존재하게 된다. 정보 거래자들은 사적 정보를 이용하여 비정보 거래자로부터 이익을 얻게 되고, 이와 반대로, 비정보 거래자들은 손실을 입게 된다. 이러한 관점에서 Easley et al.(2010, 2011a)은 정보거래확률이 높을수록 주문 흐름의 독성(toxicity of order flow)이 크다고 하였다.

많은 연구들은 정보거래확률을 측정하기 위하여 Easley et al.(1996)의 PIN(Probability of Information-based Trading) 측정치를 사용하였다. PIN을 추정하기 위해서는 최우추정법(maximum likelihood estimation)으로 관찰 불가능한 모수들을 추정해야 한다. 또한, 하루 시작 시 정보가 한 번만 발생한다는 가정 하에서 1일 PIN을 추정하는 것이기에 고빈도 시장의 독성 주문 흐름을 측정하는 데에 적용시킬 수 없다는 한계점이 존재한다.[2] 이러한 한계점을 보안하기 위하여 Easley et al.(2010, 2011a)은 고빈도 시장에 적용시킬 수 있는 수정 PIN인 VPIN을 사용하여 선물 시장에서의 정보거래확률을 추정하였다.[3] VPIN 측정치의 장점은 PIN과 같이 관찰 불가능한 모수들을 측정할 필요가 없이 거래량 시간(volume time) 기준으로 고빈도 시장에서의 독성 주문 흐름을 쉽게 추정할 수 있다는 데에 있다.[4]

LÓpez de Prado(2012)는 VPIN이 PIN의 고빈도 추정치(high frequency estimate)로써 정보거래확률을 측정할 수 있다고 하였다. Easley et al.(2012a)은 고빈도 선물 시장의 S&P 500 E-mini와 WTI crude oil을 대상으로 분석한 결과, VPIN이 주문 흐름의 독성을

1) Glosten and Milgrom(1985), Kyle(1985), Easley and O'Hara(1987, 1992), 옥기율, 이상구(2012), Ohk(2011), Ohk(2013), 장국현, 윤병조(2011), 홍정효(2012), 장하성(1995).

2) PIN에 대한 자세한 내용은 Easley et al.(1996)와 최혁, 양철원(2006)의 논문을 참조.

3) Easley et al.(2010, 2011a)은 독성 주문 흐름의 측정치인 VPIN을 최초로 소개하였다.

4) 고빈도 시장에서 거래는 단위 시간마다 연속적으로 발생하는 것이 아니라 불규칙한 빈도(irregular frequency)로 발생한다. Easley and O'Hara(1992)은 고빈도 시장에서 거래 타이밍은 새로운 정보가 시장에 도달하는 타이밍과 상관관계가 있음을 고려하여 거래량을 정보 도달(information arrival)의 대용치(proxy)로 사용하였다. 이러한 관점에서 Easley et al.(2011a, 2012a)은 상수 거래량 시간(constant volume time) 기준으로 분석하면 새로운 정보(news of comparable relevance)가 시장에 도달하는 타이밍을 모방(mimic)할 수 있다고 하였다. 여기서 상수 거래량 시간이란 누적 거래량이 사전에 정한 상수 거래량에 도달할 때마다 기록한 시간을 의미한다. 이러한 거래량 시간에 대한 예시는 <표 2>의 1열, 2열, 3열에 나타나 있다.

측정하는데 적절하다고 하였다. 또한, VPIN과 선물 시장의 변동성은 양(+)의 상관관계를 가진다고 하였다. Liu and Zhang(2013)은 중국의 CSI300 Index 선물 시장을 대상으로 분석하였는데, Easley et al.(2012a)와 유사한 결론을 얻었다. Abad and Yagüe(2012)와 Yildiz et al.(2013)은 VPIN 측정치가 고빈도 주식 시장에서도 적용이 가능하다고 하였다.

이처럼 VPIN 측정치의 긍정적 측면을 주장한 연구가 존재하지만, 반대로 부정적 견해도 존재한다. Shen(2012)은 대만의 TAIEX 고빈도 선물 시장을 대상으로 분석한 결과, 시장 변동성의 예측 능력이 미약하다고 하였다. Andersen(2013)은 Easley et al.(2010, 2011a)의 VPIN이 주문 흐름의 독성을 정확히 나타내지 못한다고 하였다.[5] 최근, Easley et al.(2014)는 Easley et al.(2010, 2011a)의 VPIN을 측정하는 방법을 제대로 적용하지 못한 결과임을 지적하여, Andersen(2013)의 주장을 반박하였다. 이처럼 국외 선행연구들은 전반적으로, VPIN이 시장 변동성을 예측하는 대용치로 적합하다는 결론을 제시하고 있다.

하지만 국내 선물시장을 대상으로 한 VPIN의 적절성을 판단한 연구는 제대로 이루어지지 않고 있다. 이에 본 연구에서는 VPIN으로 코스피200 선물 시장의 독성 주문 흐름을 측정할 수 있는지 확인하고자 한다. 또한, VPIN이 독성 주문으로 인한 선물 시장 변동성에 대한 예측 능력을 가지는지 분석하고자 한다. 마지막으로 VPIN 측정치가 고빈도 거래(high frequency trading)에서 가지는 경제적 의의에 대해서 설명하고자 한다.

Easley et al.(2012a)은 고빈도 시장에서 거래량 시간 기준으로 측정한 VPIN과 변동성 간의 관계를 분석할 때, 시계열 검증 모형을 이용하는 것이 적합하지 않다고 하였다. 이러한 한계점에 의해 Easley et al.(2012a)는 피어슨 상관계수와 조건부 확률 분포를 이용하여 분석하였는데, 본 연구도 동일한 검증을 실행하였다. 주요 분석결과는 KOSPI 200 지수 선물의 VPIN 추정치가 지속적으로 높을 때, 선물 시장 변동성이 클 가능성이 높다는 것이다. 이는 VPIN이 독성 주문 흐름으로 인한 선물 시장 변동성을 예측하는 유용한 측정치라는 것을 의미한다.

본 연구의 공헌은 다음과 같다. 첫째, PIN과 같이 복잡한 모수추정 과정을 거치지 않고도 고빈도 시장의 독성 주문 흐름을 측정할 수 있는 VPIN 측정치를 최초로 국내 선물 시장에 적용시켰다. 둘째, 거래 방향(trade direction)으로 거래량을 분류하는 전통적

5) VPIN을 추정하기 위해서는 거래량을 매수, 매도로 분류시키는 과정이 필요하다. Easley et al.(2010, 2011a)은 거래 방향(trade direction)으로 거래량을 분류하는 전통적인 방법보다 더 용이한 새로운 바켓 분류(bucket classification) 방법으로 거래량을 분류시킨 후에 VPIN을 구성하였다. Andersen(2013)은 Easley et al.(2010, 2011a) VPIN의 유용성을 검증하기 위하여, 최우선 매수, 매도(best-bid-offer) 데이터를 이용하여 거래 방향을 확인 후, 거래량을 분류시켜 VPIN을 추정하였다. 이러한 방법으로 구한 VPIN을 Easley et al.(2010, 2011a) VPIN과 비교한 결과, 상이한 추세가 나타났다.

인 방법보다 더 용이한 바켓 분류(bucket classification)의 실무적 활용성을 확인하였다.[6] 셋째, Kirilenko et al.(2011)와 Abad and Yagüe(2012)은 고빈도 거래자가 사실상 시장조성자 역할을 한다고 하였다. 국내 KOSPI200 선물 시장에서 고빈도 거래비중은 38%로 높은 편인데, VPIN이 고빈도 거래자들의 사전적 투자 전략의 근거가 될 수 있다는 점에서도 의의가 있다.[7]

논문의 구성은 다음과 같다. 제 Ⅱ장에서는 자료와 VPIN 측정치의 구성방법에 대해 설명한다. 제 Ⅲ장은 실증분석을 통해 고빈도 시장에서 KOSPI200 지수 선물 VPIN과 선물 시장 변동성 간의 관계를 설명하였고, 마지막으로 제 Ⅳ장에서는 연구의 결과를 정리하고자 한다.

Ⅱ. 자료와 VPIN 구성

1. 자료

본 연구는 KOSPI200 지수 선물 시장에서 유동성이 가장 풍부한 최근월물(nearest-to-maturity futures)을 분석 대상으로 선정하였고, 표본 기간은 2010년 1월 4일부터 2013년 6월 28일까지이다.[8] 자료는 ㈜코스콤의 CHECK에서 제공하는 KOSPI200 지수 최근월물의 1분 단위 거래량 및 종가 데이터를 사용하였다.[9]

2. VPIN 구성

VPIN은 고빈도 시장에서 정보거래확률을 나타내는 측정치이다. 이는 Easley et al.

6) VPIN을 측정하기 위해서는 거래량을 매수, 매도로 분류시켜 주문 불균형을 확인하여야 한다. 본 연구에서는 Easley et al.(2012a, 2012b)의 바켓 분류 방법을 사용하여 거래량을 분류시켰다. 이는 누적 분포 함수(cumulative distribution function)를 이용하여 매입 거래량과 매도 거래량이 차지하는 비중을 각각 구한 후에 거래량을 분류시키는 방법이다.

7) 국내 주식시장에서는 거래비용 문제 때문에 고빈도 거래가 힘들지만 파생상품 시장에서는 활발히 이루어지고 있다. 2012년 기준으로 KOSPI200 선물 시장에서의 고빈도 거래비중은 38%, 지수 옵션 시장에서의 고빈도 거래비중은 44%이다.

8) Lee(2013)도 본 연구와 동일하게 KOSPI200 지수 선물 시장에서 거래가 가장 활발한 최근월물(nearest-to-maturity futures)을 분석대상으로 하여 한국 선물 시장에서의 고빈도 거래의 거래행태(trading behavior)에 대해 분석하였다.

9) KOSPI200 지수 선물의 만기일은 3, 6, 9, 12월의 두 번째 목요일이고 만기일의 거래시간은 9시부터 14시 50분까지이다. 대입수능시험일의 KOSPI200 지수 선물의 거래시간은 1시간 늦춰 10시부터 시작하여 16시 15에 종료되고 1년의 첫 거래일의 거래시간은 1시간 늦춰 10시부터 시작하여 15시 15분에 종료된다.

(1996)의 PIN을 수정하여 고빈도 시장에 적용시킨 것이다. VPIN의 장점은 PIN과 같이 최우추정법으로 관찰 불가능한 사적 정보 모수들을 추정하는 과정을 거치지 않고도 주문 흐름의 독성을 측정할 수 있다는 데에 있다. 본 연구에서는 Easley and O'Hara(1992)와 같이 거래량 시간(volume time) 기준으로 VPIN을 측정하였다. VPIN을 측정하기 위해서는 거래량을 매수, 매도로 분류시켜 주문 불균형을 구해야 하는데, 제공되는 KOSPI200 지수 선물의 거래량 데이터는 분류되어 있지 않다. 따라서 본 연구에서는 Easley et al.(2012b)의 바켓 분류 방법을 사용하여 거래량을 분류시켰다. 이러한 방법은 고빈도 시장에 적용시키는 데에 적합하고 거래 방향(trade direction)으로 거래량을 분류하는 전통적인 방법보다 더 용이하다.

본 연구에서는 VPIN 측정치를 보다 잘 이해하기 위하여 Easley et al.(2012a)의 연구에서 제시된 VPIN에 대해 구체적인 예시를 들어 설명한다.

2.1 타임 바

VPIN을 구성하기 위한 첫 번째의 중요한 요소는 바 사이즈(bar size)이다. Easley et al.(2012a)에 의하여 본 연구에서는 VPIN을 추정하기 위하여 1분 타임 바(time bar)를 사용하였다. <표 1>은 VPIN 구성의 첫 단계(타임 바 구성)를 구체적인 예시로 보여준 것이다. 여기에서는 2010년 1월 4일 10시 1분부터 10시 9분까지의 총 8개의 타임 바와 각 타임 바에 대응하는 KOSPI200 지수 선물의 가격 변동 및 거래량을 보여주고 있다.

<표 1> VPIN 구성 과정-타임 바

아래 표는 2010년 1월 4일 10시 1분부터 10시 9분까지의 타임 바와 각 타임 바에 대응하는 가격변동과 거래량을 보여준다.

타임 바	가격변동(포인트)	거래량(계약)
10:01:01~10:02:00	223.50-222.95 = 0.55	4258
10:02:01~10:03:00	223.50-223.50 = 0.00	2335
10:03:01~10:04:00	223.50-223.50 = 0.00	1577
10:04:01~10:05:00	223.50-223.50 = 0.00	1391
10:05:01~10:06:00	222.95-223.50 = -0.55	3447
10:06:01~10:07:00	223.30-222.95 = 0.35	1185
10:07:01~10:08:00	223.10-223.30 = -0.20	859
10:08:01~10:09:00	223.05-223.10 = -0.05	2263

2.2 거래량 바켓 구성과 바켓 분류

VPIN 측정치를 구성하기 위한 두 번째 단계는 거래량 바켓(volume bucket)을 구성하는 것이다. Easley et al.(2012a)에 의하여, 한 개 거래량 바켓의 사이즈는 샘플기간 동안의 일별 평균 거래량을 50으로 나눈 상수이다.[10] 본 연구에서 샘플기간 동안의 일별 평균 거래량은 280,250계약이고 한 개 거래량 바켓의 사이즈는 5,605계약이다. 한 개의 거래량 바켓을 구성하기 위해서는 거래량의 합이 5,605계약에 도달할 때까지 연속적인 타임 바에 있는 KOSPI200 지수 선물의 거래량을 누적시킨다. 여기에서 한 개의 바켓을 구성하기 위한 마지막 타임 바의 거래량을 누적시켜 5,605계약이 넘을 경우, 남은 거래량은 다음 바켓을 구성할 때 사용한다. 그러나 어떤 시장충격으로 인해 한 개의 타임 바(즉, 1분 동안)에 대응된 거래량이 많을 경우, 한 개의 타임 바에서도 여러 개의 바켓을 구성할 수 있다.

<표 2>는 구체적인 데이터로 거래량 바켓 구성과 바켓 분류과정을 보여준다. 여기서 1열, 2열, 3열, 4열은 거래량 바켓을 구성하는 과정을 보여주고 있다. 예를 들어, 첫 번째 타임 바(10:01:01~10:02:00)에 대응된 거래량은 4,258계약이고 두 번째 타임 바(10:02:01~10:03:00)에 대응된 거래량은 2,335계약이다. 바켓 1을 구성하기 위해서는 위 두 개의 타임 바가 필요 하는데, 2개 타임 바에 대응된 거래량의 합(6,593계약)이 5,605계약을 넘기에 남은 988계약은 바켓 2를 구성 시 사용한다. 이와 같은 방법으로 거래량 바켓을 구성한 후, VPIN 측정치를 구성하기 위한 세 번째 단계는 거래량 바켓을 매입 거래량과 매도 거래량으로 분류시키는 것이다. Easley et al.(2012a, 2012b)에 의하면, 한 개 바켓 안의 매입과 매도 거래량은 다음과 같은 방법으로 분류된다.

$$V_{\kappa}^{B} = \sum_{i=t(\kappa-1)+1}^{t(\kappa)} V_i \cdot Z\left(\frac{P_i - P_{i-1}}{\sigma_{\triangle P}}\right) \tag{1}$$

$$V_{\kappa}^{S} = \sum_{i=t(\kappa-1)+1}^{t(\kappa)} V_i \cdot \left[1 - Z\left(\frac{P_i - P_{i-1}}{\sigma_{\triangle P}}\right)\right] = V - V_{\kappa}^{B} \tag{2}$$

10) Easley et al.(2012a)의 Appendix(A.2 MONTE CARLO VALIDATION)에서는 몬테카를로 시뮬레이션으로 VPIN 추정치의 정확도를 측정하였다. 분석 결과, 샘플길이(n)를 50으로 설정하여 추정한 VPIN 값의 오차가 매우 작았다. 이에 Easley et al.(2012a)은 VPIN을 추정하기 위하여, 샘플길이(n)를 50으로 설정하였고 한 개 거래량 바켓의 사이즈는 일별 평균 거래량을 50으로 나눈 상수로 설정하였다. Abad and Yagüe(2012)와 Andersen and Bondarenko(2013) 역시 동일한 방법으로 VPIN을 추정하였다.

<표 2> VPIN 구성 과정-거래량 바켓 구성과 바켓 분류

아래 표는 구체적인 예시로 거래량 바켓 구성과 바켓 분류과정을 보여주고 있다. 여기서는 2010년 1월 4일 10시 1분부터 10시 10분까지의 타임 바와 각 타임 바에 대응하는 거래량, 누적 거래량, 바켓, 가격변동, 매입 비중($Z(\triangle P/\sigma_{\triangle P})$)과 매도 비중($1-Z(\triangle P/\sigma_{\triangle P})$), 매입 거래량과 매도 거래량을 보여주고 있다. 1열, 2열, 3열, 4열은 거래량 바켓을 구성하는 과정이고 5열, 6열, 7열, 8열, 9열은 거래량을 매입과 매도로 분류하는 과정이다.

타임 바	거래량	누적 거래량	바켓	가격변동	매입 비중	매도 비중	매입	매도
10:01:01~10:02:00	4258	4258	1	0.55	0.9996	0.0004	4256.30	1.70
10:02:01~10:03:00	**1347**	**5605**	**1**	**0.00**	**0.5000**	**0.5000**	**673.50**	**673.50**
10:02:01~10:03:00	988	988	2	0.00	0.5000	0.5000	494.00	494.00
10:03:01~10:04:00	1577	2565	2	0.00	0.5000	0.5000	788.50	788.50
10:04:01~10:05:00	1391	3956	2	0.00	0.5000	0.5000	695.50	695.50
10:05:01~10:06:00	**1649**	**5605**	**2**	**-0.55**	**0.0004**	**0.9996**	**0.66**	**1648.34**
10:05:01~10:06:00	1798	1798	3	-0.55	0.0004	0.9996	0.72	1797.28
10:06:01~10:07:00	1185	2983	3	0.35	0.9839	0.0161	1165.92	19.08
10:07:01~10:08:00	859	3842	3	-0.20	0.1105	0.8895	94.92	764.08
10:08:01~10:09:00	**1763**	**5605**	**3**	**-0.05**	**0.3798**	**0.6202**	**669.59**	**1093.41**
10:08:01~10:09:00	500	500	4	-0.05	0.3798	0.6202	189.90	310.10
10:09:01~10:10:00	935	1435	4	0.25	0.9370	0.0630	876.10	58.90

여기에서, V_{κ}^{B}은 바켓 κ안의 매입 거래량이고 $t(\kappa)$은 바켓 κ을 구성하기 위한 마지막 타임 바이며, V_i은 타임 바 i안의 거래량이다. Z는 표준 정규 분포의 누적 분포 함수를 의미한다. P_i은 타임 바 i의 KOSPI200 지수 선물 종가이고 $\triangle P$는 타임 바 i과 $i-1$ 사이의 가격변동을 나타낸다. $\sigma_{\triangle P}$는 전체 샘플기간 동안의 KOSPI200 지수 선물가격 변동의 표준편차이고 V_{κ}^{S}는 바켓 κ안의 매도 거래량이다. 만약에 $\triangle P=0$이라면 한 개 타임 바 안의 매도와 매입 거래량은 동일하다. 그러나 $\triangle P>0$이라면 한 개 타임 바 안의 매입 거래량 비중이 더 크다. 반대로, $\triangle P<0$이면 한 개 타임 바 안의 매도 거래량 비중이 더 크다.

<표 2>에서 5열, 6열, 7열, 8열, 9열은 각 타임 바에 대응하는 거래량을 매입과 매도로 분류시키는 과정을 보여주고 있다. 전체 샘플기간 동안의 $\sigma_{\triangle p}$은 0.1634이다. 예를 들어, 첫 번째 타임 바(10:01:01~10:02:00)의 매입 비중은 0.9996(즉, $Z(\frac{0.55}{0.1634})$)이고 매입 거래량은 4256.30(즉, 0.9996×4258) 계약이며 매도 거래량은 총 거래량에서 매입 거래량을 뺀 값이다.

2.3 주문 불균형

VPIN을 구성하기 위한 네 번째 단계는 각 바켓의 주문 불균형의 절대치를 구하는 것이다. 주문 불균형은 매입 거래량과 매도 거래량의 차이로 구할 수 있다. <표 3>은 각 바켓의 주문 불균형을 구하는 과정을 보여주고 있다. 예를 들어, 바켓 1의 매입 거래량은 4,256.30+673.50 = 4929.80계약이고 매도 거래량은 1.70+673.50 = 675.20계약이다. 따라서 바켓 1의 주문 불균형은 4929.80-675.20 = 4254.60계약이 된다.

<표 3> VPIN 구성 과정-주문 불균형

아래 표는 각 바켓에 대응되는 매입 거래량, 매도 거래량과 주문 불균형 절대치를 나타내고 있다. 주문 불균형은 매입 거래량과 매도 거래량의 차이이다.

바켓	매입 거래량	매도 거래량	주문 불균형
1	4929.80	675.20	4254.60
2	1978.66	3626.34	1647.68
3	1931.15	3673.85	1742.70
4	2613.56	2991.44	377.88
5	2919.59	2685.41	234.18
6	3474.09	2130.91	1343.18
7	1883.75	3721.25	1837.50
8	1731.65	3873.35	2141.70

2.4 VPIN 측정치 구성

Easley et al.(2011a, 2012a)에 의하면 VPIN 측정치는 다음과 같다.

$$VPIN \approx \frac{\sum_{\kappa=1}^{n} OI_{\kappa}}{nV} = \frac{\sum_{\kappa=1}^{n} |V_{\kappa}^{S} - V_{\kappa}^{B}|}{nV} \quad (3)$$

여기에서, OI_{κ}는 바켓 κ의 주문 불균형이고 n은 샘플길이이며 V는 한 개 거래량 바켓의 사이즈이다(상수). V_{κ}^{S}은 바켓 κ안의 매도 거래량이며 V_{κ}^{B}은 바켓 κ안의 매입 거래량이다. Easley et al.(2012a)에 의하여, 한 개 거래량 바켓의 사이즈($V = V_{\kappa}^{S} + V_{\kappa}^{B}$)는 샘플기간 동안의 일별 평균 거래량을 50으로 나눈 값이고 샘플길이(n)는 50이다. 따라서 50개의 바켓으로 한 개의 VPIN 측정치를 구성할 수 있다.

<표 4>는 VPIN의 추정과정을 보여주고 있다. 예를 들어, 첫 번째 VPIN은 1~50 바켓으로 구성되고 51번째의 바켓이 형성되면 2~51 바켓을 이용하여 두 번째 VPIN을 구성한다. 따라서 샘플기간 동안 일별 평균 50개의 VPIN을 구성할 수 있다.

<표 4> VPIN 측정치의 구성

아래 표는 VPIN 측정치의 구성을 위해 사용한 바켓들을 보여주고 있다. 여기서 샘플길이는 50이다. 즉 50개의 바켓으로 한 개의 VPIN을 측정한다.

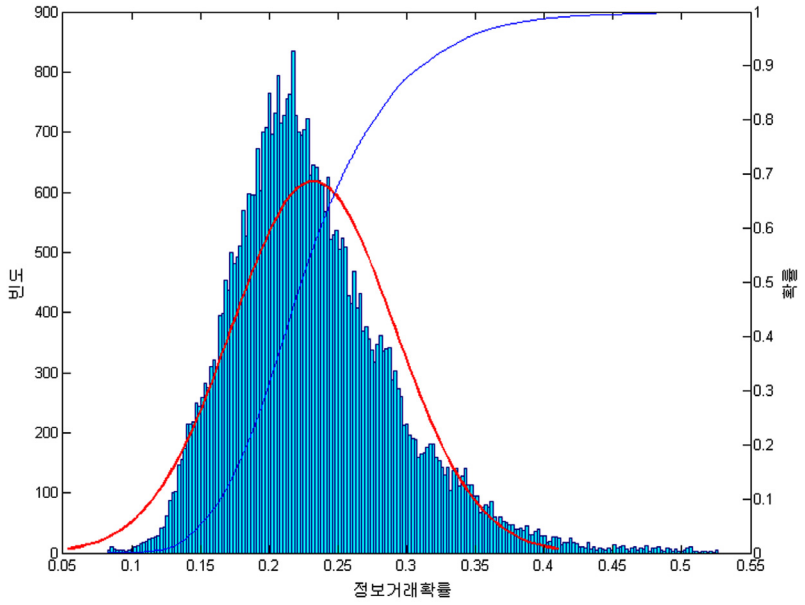
관측치	VPIN	바켓
1	0.2249	1-50
2	0.2102	2-51
3	0.2064	3-52
4	0.2017	4-53
5	0.2046	5-54
6	0.2060	6-55
7	0.2085	7-56
8	0.2029	8-57
9	0.1980	9-58
10	0.1962	10-59

<그림 1>은 샘플기간 동안 KOSPI200 지수 선물 VPIN의 분포와 VPIN의 누적 분포함수(CDF)를 보여주고 있다.[11] KOSPI200 지수 선물의 VPIN 분포 패턴은 로그 정규(log-normal) 형에 가깝다.

11) Easley et al.(2011a, 2012a)은 어느 시점의 주문 흐름의 독성을 분석하려면 그 시점의 절대적인 VPIN 수준만 보는 것이 아니라 전체 샘플기간 동안의 VPIN 분포에서의 상대적 수준을 분석하는 것도 중요하다고 보고하고 있다.

<그림 1> KOSPI200 지수 선물의 VPIN 분포와 CDF(VPIN)

아래의 그림은 샘플기간 동안의 KOSPI200 지수 선물 VPIN의 분포와 VPIN의 누적 분포 함수를 보여주고 있다. 여기서 X축은 정보거래확률(VPIN)이고 Y(좌)축은 빈도이며 Y(우)축은 확률이다.



<표 5>는 샘플기간 동안 KOSPI200 지수 선물 VPIN의 기초 통계량이다. VPIN은 정보거래확률을 나타내기에 0과 1사이의 값을 가진다. VPIN 추정치가 클수록 주문 불균형 현상이 심하고 주문 흐름의 독성이 크다는 것을 의미한다.

<표 5> VPIN의 기초 통계량

아래 표는 2010년 1월 4일부터 2013년 6월 28일까지의 KOSPI200 지수 선물의 VPIN 측정치에 대한 기초 통계량을 정리한 것이다.

통계량	KOSPI200 지수 선물의 VPIN
평균	0.2321
표준편차	0.0593
첨도	4.4254
왜도	0.9101
중앙값	0.2233
최소값	0.0822
최대값	0.5266
관측수	43450

<표 6>은 전체 샘플기간 및 2011년 8월 8일 전, 후 2개의 하위기간으로 나누어 분석할 때, KOSPI200 지수 선물의 CDF(VPIN)이 대응되는 VPIN 측정치들을 비교한 것이다.[12] 여기서, VPIN(샘플기간)이 0.1679일 때의 CDF(VPIN)가 0.1이라는 것은 전체 샘플기간 동안 KOSPI200 지수 선물의 VPIN 측정치가 0.1679 이하인 확률이 0.1이라는 것을 의미한다.

<표 6> KOSPI200 지수 선물의 CDF(VPIN)

아래 표는 KOSPI200 지수 선물의 CDF(VPIN)가 대응하는 VPIN 측정치들을 비교한 것이다.

CDF(VPIN)	VPIN(샘플기간)	VPIN(폭락 전)	VPIN(폭락 후)
0.1	0.1679	0.1646	0.1620
0.2	0.1858	0.1838	0.1818
0.3	0.1999	0.1984	0.1970
0.4	0.2115	0.2109	0.2103
0.5	0.2242	0.2233	0.2226
0.6	0.2383	0.2371	0.2362
0.7	0.2552	0.2537	0.2526
0.8	0.2804	0.2764	0.2738
0.9	0.3229	0.3112	0.3015

<표 6>에서는 전체 샘플기간 및 2개의 하위기간에서, 각 CDF(VPIN)가 대응하는 VPIN 값들이 비슷하다는 결과를 얻을 수 있다. 이는 VPIN 변수가 안정적 시계열이라는 것을 의미한다.[13]

2.5 시장 폭락일 때의 VPIN

지난 2011년 8월 5일(금요일) 미국 신용등급의 강등으로, 8월 8일(월요일) 한국, 중국, 대만, 홍콩 등 아시아 주요 증시는 2~% 폭락하며 아시아 '블랙 먼데이'를 연출했다. 국내 증시를 볼 때, 8일 오후 1시 10분 코스닥 종합지수가 10.41% 폭락하여 서킷 브레이커를 발동하였다. 코스닥 시장에 서킷 브레이커가 발동된 것은 지난 2006년 1월 미국 증시 악화 및 테마주 급락을 시작으로 처음 발생했으며 2007년 8월 서브프라임

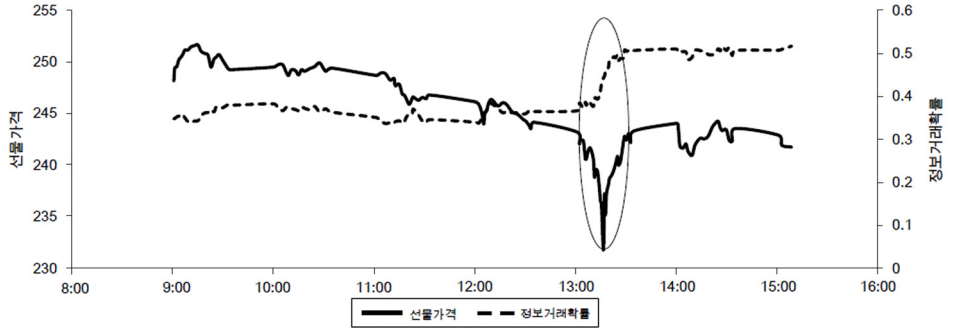
12) 2011년 8월 5일(금요일) 미국 신용등급의 강등으로 2011년 8월 8일 증시는 아시아 '블랙 먼데이'를 연출했다. <표 6>에서는 전체 샘플기간을 아시아 '블랙 먼데이' 전, 후로 나누어서도 분석하였다.

13) VPIN을 구성하는 주문 불균형의 시계열 데이터가 5% 유의수준에서 단위근을 가지고 있다는 귀무가설을 기각함으로서 안정적 시계열임을 보였다.

위기확산, 2008년 10월 두 번 연속 미국 등 글로벌 증시 동시하락으로 발생한 이후 3년여 만이다. 그리고 8일 오후 1시 23분에는 KOSPI200 선물가격 하락(전날 보다 5.52% 하락)으로 사이드카(프로그램 매도호가 일시 효력정지)를 발동했다. <그림 2>는 2011년 8월 8일 시장 폭락 하에서 KOSPI200 지수 선물 VPIN과 선물가격의 변화 추이이다.[14]

<그림 2> KOSPI200 지수 선물의 가격과 정보거래확률

아래 그림은 2011년 8월 8일(아시아 '블랙 먼데이')의 KOSPI200 지수 선물 종가와 정보거래확률(VPIN)의 변화 추이를 보여주고 있다. X축은 시간이고 Y(좌)축은 KOSPI200 지수 선물의 종가이며 Y(우)축은 정보거래확률이다.



여기에서, 시장폭락 전부터 VPIN 추정치가 점차적으로 커지는 것을 확인할 수 있다. 즉, 주문 흐름의 독성이 커지고 있다는 것을 의미한다. 이는 고빈도 시장에서 VPIN이 독성 주문 흐름으로 인한 시장 변동성을 예측하는 유용한 측정치라는 증거가 될 수 있다.[15] 이어서 제 Ⅲ장에서는 전체 샘플기간을 대상으로 VPIN과 선물 시장 변동성 간의 관계를 분석하고자 한다.

Ⅲ. VPIN 측정치와 선물 시장 변동성

고빈도 시장에서 KOSPI200 지수 선물의 주문 흐름 독성이 크다는 것은 주문 불균형 현상이 심하고 정보거래확률이 높다는 것이다. 이에 VPIN 측정치가 높을수록 선물 시장

14) 샘플기간 동안 일별 평균 50개의 VPIN을 구성하였지만, 2011년 8월 8일에는 총 127개의 VPIN을 구성하였다. 이는 그 날의 주문 흐름의 독성이 매우 크다는 것을 의미한다.

15) VPIN은 독성 주문 흐름으로 인한 변동성을 예측할 수 있지만 가격 변화의 방향(sign of the price changes)을 예측할 수 없다. 이는 VPIN이 방향성 지표(directional indicator)가 아니라는 것을 의미한다.

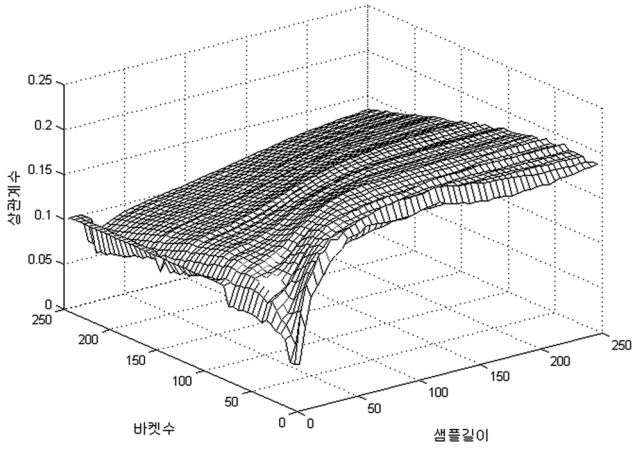
변동성이 클 것이라고 기대할 수 있다.[16] 본 연구에서는 거래량 시간 기준으로 VPIN을 추정하였기에 시계열 검증 모형을 적용시켜 고빈도 KOSPI200 선물 시장 VPIN과 선물 시장 변동성 간의 관계를 분석하기에는 적합하지 않다.[17] 따라서 Easley et al.(2012a)와 같이 피어슨 상관계수와 조건부 확률 분포를 이용하여 두 변수 간의 관계를 확인하였다.

1. VPIN 측정치와 선물 시장 변동성 간의 상관관계

VPIN과 선물 시장 변동성 간의 관계를 상관계수로 분석하였다. <그림 3>은 샘플길이와 일별 평균 바켓 수의 다양한 조합 하에서 구한 VPIN과 절대 수익률(즉, 시장 변동성)의 상관계수를 보여주고 있다.[18] 상관계수는 $\rho(VPIN_{\kappa}, \left|\frac{P_{\kappa+1}}{P_{\kappa}}-1\right|)$을 이용하여 구하였다. 여기서 P는 KOSPI200 지수 선물의 종가를 나타내고 κ는 거래량 바켓을 나타낸다.

<그림 3> KOSPI200 지수 선물 VPIN과 절대 수익률의 상관계수

아래 그림은 다양한 조합 하에서 구성한 KOSPI200 지수 선물 VPIN과 절대 수익률의 상관계수를 나타낸다.



16) Wang(2002)은 고빈도 선물 시장을 대상으로 분석 시, 새로운 정보는 불규칙한 빈도(irregular frequency)로 시장에 도달하고, 정보 거래(informed trading)로 인해 정보가 가격에 반영되므로 시장 변동성이 증가한다고 하였다. 따라서 고빈도 시장 하에서, 정보거래확률(VPIN)이 높을 때 미래 단기 변동성(short-run volatility)은 증가한다. 그러나 저빈도(low frequency) 선물 시장을 대상으로 분석 시, 정보 거래자들은 과대(과소) 평가된 가격 상황에서 매도(매수) 거래를 통해 증가된 변동성을 감소시키는 역할을 수행한다. 따라서 저빈도 시장 하에서, 정보거래확률이 높을 때 미래 변동성은 낮아진다.

17) 거래량 시간을 기준으로 VPIN을 구성하면 새로운 정보(news of comparable relevance)가 시장에 도달할 때마다 VPIN을 추정하는 것과 같다. 이와 달리 시계 시간(clock time)을 기준으로 VPIN을 구성하면 이질적인 정보량(heterogeneous amounts of information) 하에서 VPIN을 추정하는 것과 같다.

18) <그림 3>은 총 2,500개의 조합(combination)으로 구성된 것이다. VPIN이 고빈도 KOSPI200 지수 선물 시장의 독성 주문 흐름을 추정할 수 있는 유용한 측정치인지를 재확인하기 위하여 부록에 강건성 검증을 하였다.

<그림 3>은 샘플길이와 일별 평균 바켓 수의 다양한 조합 하에서 VPIN과 시장 변동성의 상관관계가 모두 1% 유의수준에서 양(+)이라는 것을 보여주고 있다. 이는 KOSPI200 지수 선물의 VPIN 측정치가 높을수록 선물 시장 변동성이 크다는 것을 의미한다. 이어서 본 연구에서는 조건부 확률 분포를 사용하여 두 변수 간의 관계를 추가적으로 분석하였다. 즉, VPIN이 높을 때와 VPIN이 지속적으로 높을 때 각각의 경우, 시장 변동성에 미치는 영향을 분석하고자 한다. 본 연구에서는 주로 지속적으로 높은 KOSPI200 지수 선물의 VPIN이 선물 시장 변동성에 미치는 영향을 통해 VPIN이 독성 주문 흐름으로 인한 변동성을 예측하는 유용한 측정치라는 것을 보여주고자 한다.

2. 조건부 확률 분포

2.1 VPIN이 높을 때의 시장 변동성

KOSPI200 지수 선물의 VPIN이 높을 때 선물 시장 변동성에 미치는 영향을 분석하였다. 여기서는 VPIN 대신 CDF(VPIN)을 사용하여 절대 수익률과의 관계를 보여주고 있다. <표 7>은 CDF(VPIN)의 조건부 확률 분포이다.

여기에서 절대 수익률은 0.25% 단위로, CDF(VPIN)은 0.1단위로 그룹핑(grouping)하여 총 10개 CDF(VPIN)의 조건부 확률 분포를 보여주고 있다. 즉 샘플기간 동안에 바켓 κ의 CDF(VPIN)가 각 그룹에 속할 때, 바켓 $\kappa+1$의 절대 수익률의 분포를 나타내고 있다(즉, $P(\left|\frac{P_{\kappa+1}}{P_{\kappa}}-1\right| | CDF(VPIN_{\kappa}))$). <표 7>로부터 다음과 같은 결과를 얻을 수 있다. 첫째, κ의 CDF(VPIN)이 낮을 때 $\kappa+1$의 절대 익률도 낮을 가능이 높다는 것을 알 수 있다. 예를 들어, CDF(VPIN)이 0~0.2일 때 절대 수익률이 0~0.25%일 가능성은 70% 넘는다. 둘째, κ의 CDF(VPIN)이 낮을 때 $\kappa+1$의 절대 수익률이 높을 확률은 0에 근사하지만 0이 아닌 값도 존재한다. 이는 시장 변동성이 큰 원인은 모두 독성 주문 흐름에 인한 것은 아니라고 해석할 수 있다. 셋째, κ의 CDF(VPIN)이 높을 때 $\kappa+1$의 절대 수익률이 낮은 것이 대부분이다. 이는 Easley et al.(2012a)의 결과와 일치하다. 이러한 실증 결과만으로 VPIN이 선물 시장 변동성을 예측하는데 유용한 측정치라고 주장하기에는 무리가 있다. 이에 본 연구에서는 지속적으로 높은 VPIN이 선물 시장의 변동성에 미치는 영향을 추가적으로 분석하여 VPIN의 유용성을 관찰하고자 한다.

<표 7> CDF(VPIN)의 조건부 확률 분포(a)

아래 표에서 절대 수익률은 0.25% 단위로, CDF(VPIN)은 0.1단위로 그룹핑(grouping)하여 총 10개 CDF(VPIN)의 조건부 확률 분포를 보여주고 있다. 여기서는 샘플기간 동안에 바켓 κ의 CDF(VPIN)가 각 그룹에 속할 때의 바켓 $\kappa+1$의 절대 수익률 분포를 나타내고 있다.

CDF(VPIN)	0.25(%)	0.50(%)	0.75(%)	1.00(%)	1.25(%)	1.50(%)	1.75(%)	2.00(%)	>2.00(%)
0.10	75.58	18.57	3.39	1.29	0.47	0.35	0.23	0.00	0.12
0.20	70.09	22.08	5.72	0.93	0.47	0.47	0.00	0.00	0.24
0.30	67.02	24.85	4.95	1.06	1.06	0.59	0.35	0.00	0.12
0.40	62.92	28.16	5.45	1.27	1.16	0.58	0.00	0.12	0.34
0.50	62.90	28.03	5.42	1.18	0.94	0.35	0.71	0.24	0.23
0.60	63.81	26.22	6.61	1.16	0.93	0.70	0.46	0.00	0.11
0.70	61.29	27.72	7.49	1.29	0.82	0.35	0.47	0.35	0.22
0.80	64.10	27.16	4.95	1.38	0.46	0.23	0.58	0.46	0.68
0.90	56.94	29.05	10.04	1.98	0.58	0.47	0.23	0.12	0.59
1.00	42.92	29.00	14.73	5.34	3.02	1.51	0.93	0.70	1.85

<표 8> CDF(VPIN)의 조건부 확률 분포(b)

아래 표에서 절대 수익률은 0.25% 단위로, CDF(VPIN)은 0.1단위로 그룹핑(grouping)하여 총 10개 CDF(VPIN)의 조건부 확률 분포를 보여주고 있다. 여기서는 CDF(VPIN)이 0.1 범위 내에서 유지될 때의 절대 수익률 최대치의 분포이다.

CDF(VPIN)	0.25(%)	0.50(%)	0.75(%)	1.00(%)	1.25(%)	1.50(%)	1.75(%)	2.00(%)	>2.00(%)
0.10	24.24	24.24	9.09	3.03	6.06	9.09	6.06	3.03	15.16
0.20	23.08	19.78	20.88	6.59	10.99	8.79	2.20	2.20	5.49
0.30	23.81	26.19	21.43	5.56	7.94	4.76	3.97	0.79	5.55
0.40	27.41	27.41	17.04	11.85	5.19	3.70	2.22	1.48	3.70
0.50	20.14	23.02	25.90	12.95	7.91	5.04	2.16	0.72	2.16
0.60	20.61	25.19	17.56	15.27	4.58	7.63	4.58	1.53	3.05
0.70	19.33	31.93	14.29	6.72	6.72	5.04	3.36	5.88	6.73
0.80	10.84	15.66	20.48	12.05	7.23	6.02	6.02	2.41	19.29
0.90	16.00	12.00	8.00	8.00	2.00	4.00	6.00	6.00	38.00
1.00	25.00	0.00	0.00	16.67	0.00	0.00	0.00	8.33	50.00

2.2 VPIN이 지속적으로 높을 때의 시장 변동성

본 연구에서는 Easley et al.(2012a)에 의하여, CDF(VPIN)이 일정 범위에 도달하여 지속적으로 유지될 때, 그 기간 내에서의 최대 변동성을 측정하여 두 변수 간의 관계를 보여주었다. 예를 들어, 만약에 CDF(VPIN)의 지속 범위를 0.1단위로 설정하고 바켓 1의 CDF(VPIN)가 0.81이고 바켓 2의 CDF(VPIN)가 0.85, 바켓 3의 CDF(VPIN)가 0.86, 바켓 4의 CDF(VPIN)가 0.92일 때, 바켓 1과 바켓 2, 바켓 2와 바켓 3, 바켓 1과 바켓 3에 대응하는 선물종가로 각각의 절대 수익률을 측정한다. 그 중에서 제일 큰 절대 수익률 값이 그 기간 내의 최대 변동성이다. Easley et al.(2012a)은 위와 같은 과정을 다음과 같은 수식으로 나타냈다.

$$\max\left|\frac{P_\theta}{P_m}-1\right| \quad (4)$$
$$\kappa(i) \le m < \theta,$$
$$\kappa(i) < \theta \le \kappa(i+1)$$

여기에서, P는 KOSPI200 지수 선물의 종가이고 $\kappa(i+1)-\kappa(i)$ 은 동일 범위 내에 있는 바켓 수이다.[19] <표 8>은 총 10개의 CDF(VPIN) 조건부 확률 분포를 보여주고 있다. 즉, CDF(VPIN)가 일정 범위 내에서 유지될 때 절대 수익률 최대치의 분포이다. 여기서 CDF(VPIN)가 높은 범위(즉, 0.6~0.7, 0.7~0.8, 0.8~0.9, 0.9~1.0) 내에서 유지될 때, 절대 수익률이 0.75%를 초과할 확률은 평균 67.31%라는 것을 보여주고 있다.[20] 즉, VPIN이 지속적으로 높을 때 선물 시장 변동성이 클 가능성이 크다는 것을 의미한다.[21] 이는 VPIN이 독성 주문 흐름으로 인한 시장 변동성을 예측하는 유용한 측정치임을 주장한 Easley et al.(2012a)의 연구결과에 부합한다.

Ⅳ. 결 론

본 연구에서는 Easley et al.(2010, 2011a)이 제안한 정보거래확률 측정치인 VPIN을

19) 위의 예에서 k는 2 또는 3의 값을 가지고 m은 1 또는 2의 값을 가진다.

20) <표 8>에서 CDF(VPIN)가 0.6~0.7, 0.7~0.8, 0.8~0.9, 0.9~1.0 내에서 지속될 때, 절대 수익률이 0.75%를 초과할 확률은 각각 48.74%, 73.5%, 72%, 75%이다.

21) 예를 들어, <그림 2>에서도 2011년 8월 8일 시장폭락 시점 전부터 VPIN 측정치는 지속적으로 높았다.

이용하여 고빈도 시장에서 KOSPI200 지수 선물의 독성 주문 흐름을 확인하였다. VPIN 측정치의 장점은 PIN과 같이 관찰 불가능한 모수들을 측정할 필요가 없다. 또한 거래량 시간 기준으로 VPIN을 구성하기에 고빈도 시장에 적용하기 용이하다.

본 연구는 고빈도 시장에서의 KOSPI200 선물 VPIN과 선물 시장 변동성 간의 관계를 상관계수와 조건부 확률 분포로 분석하였다. 주요 분석결과는 VPIN 측정치가 지속적으로 높을 때 시장 변동성이 커질 가능성이 높다는 것이다. 이는 KOSPI200 선물 VPIN이 독성 주문 흐름으로 인한 변동성을 예측하는 유용한 측정치임을 주장한 Easley et al.(2012a)의 연구와 동일한 결과를 보여주었다.

국내 KOSPI200 지수 선물 시장에서의 고빈도 거래비중은 38%로 높은 편이다. Kirilenko et al.(2011)와 Abad and Yagüe(2012)은 고빈도 거래자가 사실상 시장조성자 역할을 한다고 보고하였다. 주문 흐름이 균형 시, 고빈도 거래자들은 주문과 체결을 무한정 반복해 거래량을 늘리고 큰 이익을 얻을 수 있다. 반대로, 주문 흐름이 불균형 시 고빈도 거래자들은 역선택 문제로 손실을 입게 된다. 이에 시변(time-varying) 독성 측정치인 VPIN은 고빈도 거래자(high frequency trader)들의 사전적 투자 전략의 근거로 될 수 있다. 그러나 VPIN을 시장 안정화 지표로 사용하면 시장 효율성을 저해할 수 있다. 즉, 새로운 정보가 가격에 반영되는 과정에서 VPIN과 시장 변동성은 증가하게 되며, 이러한 변동성의 증가는 시장 효율성을 증가시키는 긍정적 현상이다. 따라서 VPIN을 통하여 변동성 증가를 예측하고 이를 차단하는 정책을 수행한다면, 가격의 정보 반영 속도를 저해하여 시장 효율성을 저해할 수 있다.

다만, 본 연구에서는 다양한 방법론을 고려하지 못한 한계가 존재한다. 만약 차후에 VPIN을 이용한 시계열 분석의 계량적 발전이 이루어지면 VPIN의 유용성이 더욱 제고될 것으로 기대한다.

참 고 문 헌

장국현, 윤병조, “KOSPI200 현·선물 거래량 변동성과 주가변동성에 관한 연구,” 「선물연구」, 19, 2, 2011, 149-173.

장하성, “KOSPI200에 대한 지수추적 포트폴리오의 추적효과와 유동성에 관한 연구,” 「선물연구」, 10, 3, 1995, 1-44.

최 혁, 양철원, “한국주식시장에서의 정보비대칭 측정치 비교,” 「증권학회지」, 35, 5, 2006, 1-44.

홍정효, “국제통화선물시장에서의 거래량정보의 비대칭적 특성에 관한 연구: 선진국 vs 개도국시장 비교,” 「선물연구」, 20, 2, 2012, 237-264.

Abad, D. and J. Yagüe, “From PIN to VPIN : An Introduction to Order Flow Toxicity,” *The Spanish Review of Financial Economics*, 10, 2, 2012, 74-83.

Andersen, T. G. and O. Bondarenko, “VPIN and the Flash Crash,” available at SSRN : http://ssrn.com/abstract=1881731, 2013.

Easley, D. and M. O’Hara, “Price, Trade Size, and Information in Securities Markets,” *Journal of Financial Economics*, 19, 1, 1987, 69-90.

Easley, D. and M. O’Hara, “Time and the Process of Security Price Adjustment,” *Journal of Finance*, 47, 2, 1992, 576-605.

Easley, D., M. LÓpez de Prado, and M. O’Hara, “Bulk Classification of Trading Activity,” 2012b, Working paper.

Easley, D., M. LÓpez de Prado, and M. O’Hara, “Flow Toxicity and Liquidity in a High Frequency World,” *Review of Financial Studies*, 25, 5, 2012a, 1457-1493.

Easley, D., M. LÓpez de Prado, and M. O’Hara, “Measuring Flow Toxicity in a High Frequency World,” available at SSRN : http://ssrn.com/abstract=1695596, 2010.

Easley, D., M. LÓpez de Prado, and M. O’Hara, “The Microstructure of the ‘Flash Crash’ : Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading,” *The Journal of Portfolio Management*, 37, 2, 2011a, 118-128.

Easley, D., M. LÓpez de Prado, and M. O’Hara, “VPIN and the Flash Crash : A rejoinder,” *Journal of Financial Markets*, 17, 2014, 47-52.

Easley, D., N. Kiefer, M. O’Hara, and J. Paperman, “Liquidity, Information, and Infrequently Traded Stocks,” *Journal of Finance*, 51, 4, 1996, 1405-1436.

Glosten, L. R. and P. Milgrom, “Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders,” *Journal of Financial Economics*, 14, 1, 1985, 71-100.

Kirilenko, A., A. P. Kyle, M. Samedi, and T. Tuzun, “The Flash Crash : The Impact of High-Frequency Trading on an Electronic Market,” 2011, Working paper.

Kyle, A. S., “Continuous Auctions and Insider Trading,” *Journal of the Econometric Society*, 53, 6, 1985, 1315-1335.

Lee, E. J., “High Frequency Trading in the Korean Index Futures Market,” *The Journal of Futures Market*, forthcoming, 2013.

Liu, W. W. and H. J. Zhang, “Measuring Flow Toxicity in a High Frequency World : Empirical Research Based on the CSI300 Index,” *Economic Issues in China*, 276, 1, 2013, 81-91.

LÓpez de Prado, M., “Low-Frequency Traders in a High-Frequency World : A Survival Guide,” available at SSRN : http://ssrn.com/abstract=2150876, 2012.

Ohk, K. Y., “Asymmetric Relationships between Unexpected Trading Activity and Stock Market Volatility,” *Journal of the Korean Data Analysis Society*, 15, 4(A), 2013, 1801-1809.

Ohk, K. Y., “Investor’s Trading Activity and Stock Market Volatility International Evidence,” *Journal of the Korean Data Analysis Society*, 13, 5(A), 2011, 2253-2261.

Shen, H. M., “The Study of Volume-Synchronized Probability of Informed Trading : Case of TAIEX Futures,” 2012, Working paper.

Wang, C. Y., “Information, Trading Demand, and Futures Price Volatility,” *The Financial Review*, 37, 2002, 295-316.

Yildiz, S., R. A. Van Ness, and B. F. Van Ness, “Analysis Determinants of VPIN, HFTs’ Order Flow Toxicity and Impact on Stock Price Variance,” 2013, Working paper.

<부 록> 다양한 조합 하에서 구성한 VPIN

VPIN이 고빈도 KOSPI200 지수 선물 시장의 독성 주문 흐름을 추정할 수 있는 유용한 측정치인지를 재확인하기 위하여 강건성 검증을 하였다. 즉, 다양한 조합(서로 다른 샘플길이와 일별 평균 바켓 수) 하에서 구성한 VPIN 측정치의 기초 통계량, 시계열 추이를 추가하여 분석하였다.

1. 다양한 조합 하에서 구성된 VPIN의 기초 통계량

일별 평균 바켓 수가 40, 샘플길이가 50일 때의 VPIN와 일별 평균 바켓 수가 50, 샘플길이가 250일 때의 VPIN을 추가로 구성하였다. <표 1>는 다양한 조합 하에서 구성한 VPIN의 기초 통계량을 보여주고 있다. VPIN-50-50은 일별 평균 바켓 수가 50이고 샘플길이가 50일 때의 VPIN을 나타내고 VPIN-50-250은 일별 평균 바켓 수가 50이고 샘플길이가 250일 때의 VPIN을 나타낸다. VPIN-40-50은 일별 평균 바켓 수가 40이고 샘플길이가 50일 때의 VPIN을 나타낸다.

<표 1> 다양한 조합 하에서 구성한 VPIN의 기초 통계량

아래 표는 서로 다른 샘플길이와 일별 평균 바켓 수로 구성된 KOSPI200 지수 선물의 VPIN 측정치에 대한 기초 통계량을 정리한 것이다. 여기에서 VPIN-50-50은 일별 평균 바켓 수가 50이고 샘플길이가 50일 때의 VPIN을 나타내고 VPIN-50-250은 일별 평균 바켓 수가 50이고 샘플길이가 250일 때의 VPIN을 나타낸다. VPIN-40-50은 일별 평균 바켓 수가 40이고 샘플길이가 50일 때의 VPIN을 나타낸다.

통계량	VPIN-50-50	VPIN-40-50	VPIN-50-250
평균	0.2321	0.2233	0.2320
표준편차	0.0593	0.0551	0.0464
첨도	4.4254	4.1323	5.4613
왜도	0.9101	0.8085	1.2158
중앙값	0.2233	0.2158	0.2247
최소값	0.0822	0.0771	0.1398
최대값	0.5266	0.4835	0.4715
관측수	43450	39105	43250

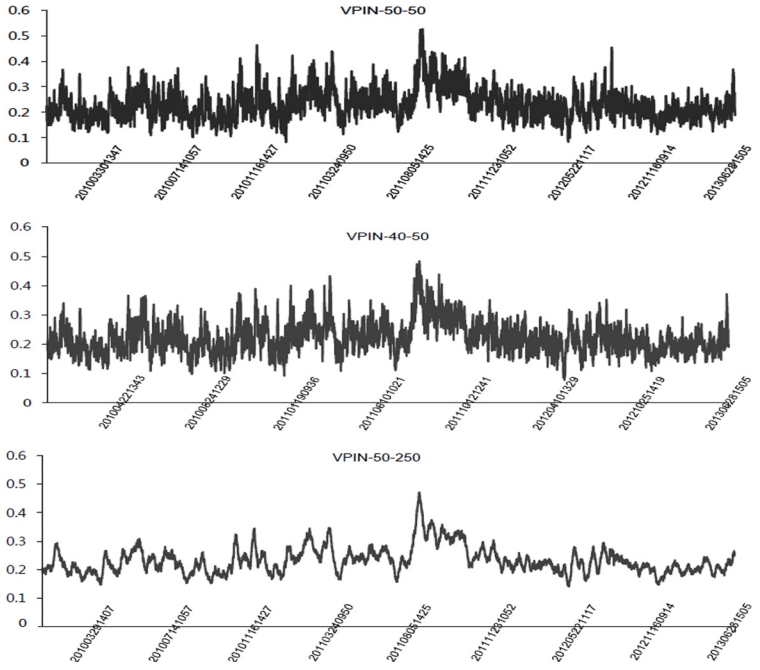
<표 1>는 다양한 조합으로 VPIN을 구성하여도 기초 통계량에 큰 영향을 미치지 않는다는 것을 보여주고 있다. 여기에서 관측수가 서로 상이한 원인은 VPIN을 구성하는 샘플길이와 일별 평균 바켓 수가 서로 다르기 때문이다.

2. 다양한 조합 하에서 구성한 VPIN의 시계열 추이

<그림 1>는 다양한 조합 하에서 구성한 VPIN의 시계열 추이를 보여주고 있다. 표본기간은 2010년 1월 4일부터 2013년 6월 28일까지이다. 상단 그림은 일별 평균 바켓 수가 50이고 샘플길이가 50일 때의 VPIN 시계열 추이이고 중단 그림은 일별 평균 바켓 수가 40이고 샘플길이가 50일 때의 VPIN 시계열 추이이다. 하단 그림은 일별 평균 바켓 수가 50이고 샘플길이가 250일 때의 VPIN 시계열 추이이다.

<그림 1> 다양한 조합 하에서 구성한 VPIN의 시계열 추이

아래 그림은 다양한 조합(즉, 서로 다른 샘플길이와 일별 평균 바켓 수) 하에서의 VPIN 시계열 추이를 나타낸 것이다. 상단 그림은 일별 평균 바켓 수가 50이고 샘플길이가 50일 때의 VPIN 시계열 추이이고 중단 그림은 일별 평균 바켓 수가 40이고 샘플길이가 50일 때의 VPIN 시계열 추이이다. 하단 그림은 일별 평균 바켓 수가 50이고 샘플길이가 250일 때의 VPIN 시계열 추이이다. X축은 시간을 나타내고 Y축은 정보거래확률(VPIN)을 나타낸다.



<그림 1>에서 각 그림들은 추정된 VPIN의 개수가 달라서 서로 다른 그림처럼 보일 뿐 그 형태는 비슷하다.[22]

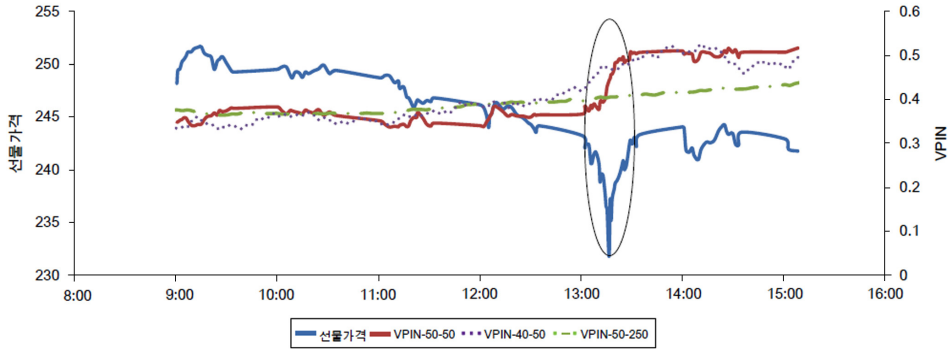
22) 한 개의 VPIN-50-50 또는 한 개의 VPIN-40-50을 구성하기 위해서는 50개의 바켓이 필요하다. 그러나 한 개의 VPIN-50-250을 구성하기 위해서는 250개의 바켓이 필요하다. 따라서 다양한 조합 하에서 추정된 VPIN의 개수는 다르다.

3. 시장 폭락 하에서의 VPIN

<그림 2>는 선물 시장 폭락(2011년 8월 8일 아시아 '블랙 먼데이') 하에서 추정된 VPIN을 분석한 것이다.

<그림 2> 시장 폭락 하에서의 KOSPI200 지수 선물가격과 VPIN

아래 그림은 2011년 8월 8일(아시아'블랙 먼데이')의 KOSPI200 지수 선물종가와 다양한 조합 하에서 구성한 정보거래확률(VPIN) 변화 추이를 보여주고 있다. X축은 시간이고 Y(좌)축은 KOSPI200 지수 선물의 종가이며 Y(우)축은 정보거래확률(VPIN)이다.



<그림 2>에서는 시장 폭락 전부터 VPIN 측정치들이 모두 상승하고 있다는 것을 확인할 수 있다. 즉, 주문 흐름의 독성이 커지고 있다는 것을 의미한다. 이는 서로 다른 샘플길이와 바켓 수로 VPIN을 구성하여도, VPIN은 고빈도 시장의 독성 주문 흐름을 추정할 수 있는 유용한 측정치라는 것을 보여주고 있다.

Korean Journal of Futures and Options
Volume 22 Number 1
February 2014

Measuring the Toxicity of Order Flow using High Frequency Data : The Case of KOSPI200 Index Futures

Ki Yool Ohk*
Ming Wu**

< Abstract >

This study presents a new informed trading probability measure VPIN (Volume-Synchronized Probability of Informed Trading) to estimate toxic order flow of KOSPI200 index futures in a high frequency world. This measure does not require to estimate non-observable parameters as PIN. Also, it is estimated based on volume time, so it can estimate toxicity of order flow in a high frequency world. We show a relation between KOSPI200 index futures VPIN and futures market volatility using correlation and conditional probability distribution. A main empirical result is that persistently high VPIN signifies a high risk of subsequent large futures market volatility. It means that VPIN is a useful measure to estimate a toxicity induced volatility.

JEL classification : C12, G13
Keywords : KOSPI200 Index Futures, High Frequency Data, VPIN, Toxic Order Flow, Volatility

Received : 21 Oct. 2013; Revised : 14 Jan. 2014; Accepted : 4 Feb. 2014

* **Corresponding Author. Address** : Department of Business Administration, College of Business Administration, Pusan National University, San30, Jangjeon-dong, Geumjeong-gu, Busan 609-735, Republic of Korea; **E-mail** : kyohk@pusan.ac.kr; **Tel** : +82-51-510-2552.

** Department of Business Administration, College of Business Administration, Pusan National University; **E-mail** : wuming5@pusan.ac.kr.