

# 경제/금융 변수를 이용한 한국 주식시장의 변동성 분석 및 예측 \*

이승희 †

한희준 ‡

## Abstract

본 논문에서는 반모수 단일지표(semiparametric single index) 모형을 이용하여 한국 주식시장 수익률의 변동성을 분석하였다. 반모수 단일지표 모형은 변동성의 단기적 변동을 설명하는 GARCH 형태의 모형과 변동성의 장기적 변동을 설명하는 단일지표 모형이 곁해진 형태이다. 이 모형은 관측 가능한 경제/금융 변수들 중 비정상(nonstationary) 변수들을 단일지표 모형에 포함시킴으로써 금융시계열의 비조건부 분산이 시간에 따라 변하는 비정상성을 설명할 수 있으며, 또한 모형을 반모수적으로 추정하므로 모형 설정 오류의 우려를 낮출 수 있다. KOSPI 지수 수익률의 변동성을 분석한 결과는 경기동행지수, VKOSPI지수, 주택매매가격지수를 이용할 때 한국 주식시장 변동성의 장기적 변동을 가장 잘 설명할 수 있음을 나타냈다. 표본 기간 내 및 기간 외 변동성 예측력을 평가한 결과 기존의 모형들보다 반모수 단일지표 모형이 우월하다는 것을 확인하였다.

This version: Dec 4, 2015

*JEL Classification:* C22, C50, G12

*Key words and phrases:* volatility, semiparametric single index model, GARCH-X, nonstationarity

---

\*본 논문에 대해 유익한 논평과 조언을 해주신 김창식 교수님(성균관대학교), 박준용 교수님(Indiana University 및 성균관대학교)과 2015 한국계량경제학회 참석자분들, 90th Annual conference of the West Economic Association International(Honolulu) 참석자분들 그리고 2015 Workshop of Sungkyun Institute of Economic Research 참석자분들께 감사드립니다.

†성균관대학교 경제학과 석사과정, e-mail: prious@skku.edu

‡교신저자, 성균관대학교 경제학과 부교수, e-mail: heejoonhan@skku.edu

# 1 서론

최근 금융시계열 변동성 모형에 관련된 연구에서는 다음과 같은 특징들을 살펴볼 수 있다. 첫째, 금융시장의 변동성을 설명할 때 실물경제상황 등의 경제적 요인들을 반영하는 모형이 큰 관심을 받고 있다. 금융시장의 변동성을 설명하는 경제적 요인을 밝히고자하는 노력은 Schwert(1989) 등에 의해 오래전부터 시도되어왔으나, 특히 최근에는 경제/금융 변수들을 이용해 변동성의 장기적 변동을 설명하는 시도가 주목을 받고 있다. 둘째, 추가수익률 등의 금융시계열이 정상적(stationary)라는 가정이 비현실적이므로 금융시계열의 비조건부 분산(unconditional variance)이 시간에 따라 변하는 비정상성(nonstationarity)을 허용하는 변동성 모형들이 새롭게 제안되고 있다. 이는 최근 2008년 리먼 브라더스사의 파산을 전후한 세계경제위기 그리고 2013년의 유럽 국채위기 등 경제/금융위기 시에 금융시장의 변동성이 급격히 높아지고 상당기간 높은 수준이 유지되는 현상이 관찰됨에 따라 이를 변동성 모형에 반영하기 위해서이다. 셋째, 이러한 점들과 관련하여 변동성의 장기적 변동을 설명하는 부분(long run component)과 단기적 변동을 설명하는 부분(short run component)을 곱하는 형태(multiplicative form)의 변동성 모형들에 대한 연구가 활발히 진행되고 있다.

이러한 특징들을 반영하면서 최근 많은 주목을 받고 있는 변동성 모형들은 Engle and Rangel(2008)의 spline-GARCH 모형, Hafner and Linton(2010)의 multivariate spline-GARCH 모형, Amado and Teräsvirta(2013, 2014)의 time-varying GARCH 모형 그리고 Engle et al.(2013)의 GARCH-MIDAS 모형 등이다. 이 모형들은 변동성의 단기변동부분을 GARCH 형태의 모형으로 모형화를 하는 공통점을 가지고 있는 반면 장기변동부분의 모형화에는 제각기 상이한 방법들을 채택하고 있다. spline-GARCH 모형과 time-varying GARCH 모형은 변동성의 장기변동부분을 시간의 함수(deterministic function of time)로 정의하고 있는데, Engle and Rangel(2008)과 Hafner and Linton(2010)은 비모수(nonparametric) 모형을 선택한 반면 Amado and Teräsvirta(2013)는 logistic 함수를 이용하고 있다. GARCH-MIDAS 모형은 저빈도(low frequency)의 경제변수들을 특정한 가중합 방식으로 이루어진 함수로 장기변동부분을 설명하고 있다.<sup>1</sup>

본 논문은 경제적 요인들을 반영하고 비정상성을 설명하는 등 최근의 주요 관심사항들을 반영하는 변동성 모형을 이용하여 한국의 KOSPI 지수 수익률의 변동성을 분석 및 예측하였다.<sup>2</sup> 구체적으로 본

<sup>1</sup> 최근에는 이러한 구조적 모형 대신 실현변동성(realized volatility)의 시계열을 직접 이용한 Corsi (2009)의 HAR 모형 등 축약형 모형(reduced form model) 또한 널리 연구되고 있다. 변동성의 예측만이 목적이라면 이러한 축약형 모형을 이용하는 것이 가능하나, 본 논문에서는 다음과 같은 이유로 축약형 모형 대신 금융시계열의 구조적 모형을 다루고 있다. 첫째, 축약형 모형을 사용하기 위해서는 고빈도 데이터(high frequency data)를 이용한 실현변동성을 먼저 계산해야 하지만 관심이 있는 금융시계열의 고빈도 데이터를 구하기 어려울 수가 있다. 둘째, 구조형 모형의 경우 GARCH 옵션가격모형을 통해 옵션가격의 결정에 직접적으로 이용될 수 있는 등 파생상품 가격결정에 더 적합할 수 있다.

<sup>2</sup> 기존의 한국 주식시장의 변동성을 다른 경제/금융 변수를 이용하여 분석하려 한 선행 연구로는 오현탁·이현상·이치송(2000)과 김세완(2009), 김영일(2013)의 연구 등을 들 수 있다. 오현탁·이현상·이치송(2000)의 연구에서는 AR-GJR(1,1)-M 모형과 AR-EGARCH(1,1)-M 모형을 이용하여 한국 주식시장 수익률의 비대칭적 변동성이 유의하게 나타나며, 시장 상승기보다 시장 하락기에서 변동성이 음의 충격에 더 민감하게 반응함을 연구하였다. 김세완(2009)은 경기동행지수를 기반으로 하여 경기 상황을 나타내는 데미 변수로 변환하여 AR-EGARCH(1,1)-M 모형을 이용하여 주식시장의 변동성을 분석하였다. 김영일(2013)의 연구는 GARCH-MIDAS 모형을 이용하였는데 변화하는 비조건부 분산을 잡기 위한 저빈도 변동성 부분에 대해서는 주식시장

논문에서는 반모수 단일지표(semiparametric single index) 변동성 모형을 채택하고 있는데, 이는 앞서 설명한 모형들처럼 변동성의 단기변동부분은 GARCH 형태의 모형으로 모형화하는 한편 이에 곱해지는 장기변동부분은 관측 가능한 경제/금융 변수들을 공변량(covariate)으로 이용하는 반모수 단일지표 모형으로 모형화하는 특징을 지니고 있다. 이 모형의 장점은 1) 어떠한 경제/금융 지표들이 해당 금융시계열 변동성의 장기변동에 영향을 미치는지에 대한 정보를 제공할 수 있으며, 2) 경제/금융 변수들 중 비정상 변수들을 포함시킴으로써 실제 금융시계열이 나타내는 비조건부 분산이 시간에 따라 변하는 비정상성을 설명할 수 있으며, 3) 모형을 반모수적으로 추정하므로 모형 설정 오류(model misspecification)의 우려를 낮출 수 있는 점 등이 있다.

본 논문에서는 반모수 단일지표 모형을 이용하여 한국 주식시장의 변동성에 영향을 미치는 외생적인 공변량(exogenous covariate)으로 실물경제상황과 금융시장여건을 반영하는 다양한 경제/금융 지표들을 살펴보았다. 그 중 경기동행지수, VKOSPI지수, 주택매매가격지수 등 세 가지 변수를 동시에 사용할 경우 한국 주식시장 변동성을 가장 잘 설명하는 것으로 나타났다. 경기동행지수의 경우는 실물경기의 흐름이 주식시장의 변동성에 영향을 미치고 있음을 나타내며, 옵션 시장에서 도출되는 내재변동성(implied volatility)으로 공포지수라 불리는 VKOSPI지수는 금융시장의 여건을 나타내는데 이 또한 한국 주식시장 변동성에 영향을 미침을 드러내고 있다. 마지막으로 주택매매가격지수의 경우 주식시장과 함께 대표적인 자산시장으로 간주되는 부동산시장의 상황 또한 주식시장의 변동성에 영향을 미침을 나타내고 있다.

본 논문은 실증분석에서 2004년 1월 5일부터 2014년 8월 29일까지의 일별 KOSPI 지수 수익률을 이용하여 표본 기간 내 및 기간 외 변동성 예측력을 비교/평가하고 있으며, 변동성 예측력의 평가에 대한 최근의 관련 연구들을 반영한 평가기준을 채택하고 있다. 실제 변동성에 대한 대용치(proxy)로는 고빈도 데이터(high frequency data)를 이용해 만들어진 실현 변동성 중 시장 미시구조 잡음(market microstructure noise)에 강건(robust)한 실현 커널(realized kernel)을 사용하였고, 예측 평가시 손실함수(loss function)로는 변동성 측정 오차에 강건한 손실함수들 중 가장 검증력(power)이 높다고 알려진 QLIKE 손실함수를 이용하였고, 마지막으로 Diebold-Mariano(1995)와 West(1996)의 검정(DMW test)을 시행하여 손실함수 간 차이의 유의성을 검정했다. 본 논문은 경기동행지수, VKOSPI지수, 주택매매가격지수 세 가지 변수를 사용한 반모수 단일지표 모형이 기존의 변동성 모형들에 비해 우월한 표본 기간 내 및 기간 외 예측력을 나타냄을 보이고 있다.

본 논문은 다음과 같이 구성되어 있다. 2장에서 모형을 소개하고 추정 방법을 설명하며 3장에서는 실증분석에 사용한 데이터, 모형의 추정 결과 그리고 모형의 예측력 평가 결과를 보고한다. 그리고 4장에서 연구의 결론을 제시한다.

---

수익률의 실현된 변동성(realized variance)과 산업 생산(industrial production)의 증가율, 생산자 물가 상승률, CD 금리와 원/달러 환율의 변화율을 이용하였다.

## 2 모형과 추정

### 2.1 모형

최근 Han and Kristensen(2015)은 SEMIX(Semiparametric multiplicative GARCH-X) 변동성 모형을 소개하였는데, 이는 다음과 같이 정의된다.

$$y_t = \sigma_t \epsilon_t, \quad \epsilon_t \stackrel{\text{iid}}{\sim} (0, 1)$$

$$\sigma_t^2 = h_t \times f(x_{t-1}) \quad (1)$$

여기서  $y_t$ 는 평균을 제거(demean)한 금융시계열의 수익률이고  $\sigma_t^2$ 은  $y_t$ 의 조건부 분산(conditional variance), 즉 변동성이다.  $\sigma_t^2$ 을 구성하는 첫 번째 요소인  $h_t$ 는 변동성의 단기변동을 설명하며 이는 GARCH(1,1) 또는 GJR-GARCH(1,1) 등의 GARCH 형태로 모형화한다.  $\sigma_t^2$ 을 구성하는 두 번째 요소인  $f(x_{t-1})$ 는 변동성의 장기변동을 설명하며 이는 단위근 또는 준단위근을 가진 외생적인 공변량  $x_{t-1}$ 의 비모수 함수로 모형화한다. Han and Kristensen(2015)은 SEMIX 모형의 시계열 특성, 추정 방법 그리고 추정치의 일치성(consistency)과 점근 분포(asymptotic distribution)에 대한 이론 등을 제시하고 있으며 또한 SEMIX 모형의 다변량(multivariate) 확장형 모형으로 반모수 단일지표 모형을 간략히 소개하였다.

구체적으로 반모수 단일지표 모형은 변동성의 장기변동을 설명할 때 하나 이상의 경제/금융 변수들을 이용할 수 있도록 위 식 (1)에서  $f(x_{t-1})$  대신 단일지표 형태의  $f(X'_{t-1}\beta)$ 을 채택하는 것이다. 이 경우  $X_{t-1}$ 은 외생적인 공변량들로 이루어진 벡터  $X'_{t-1} = [x_{1,t-1}, \dots, x_{k,t-1}]$ 이다. 본 논문에서는 주가 수익률의 변동성 모형을 고려하고 있으므로 수익률과 변동성의 비대칭적 관계를 설명할 수 있는 GJR-GARCH 모형을  $h_t$ 로 정의하였다.

$$h_t(\theta) = \omega + \alpha h_{t-1} \epsilon_{t-1}^2 + \beta h_{t-1} + \gamma h_{t-1} \epsilon_{t-1}^2 \mathbb{I}(\sqrt{h_{t-1}} \epsilon_{t-1} < 0)$$

여기서  $\theta$ 는  $h_t$ 의 모수들의 벡터이다. 모형에 대한 세부적인 가정들은 다음과 같다.

- (i)  $\alpha, \beta, \gamma \geq 0, \quad \alpha + \gamma/2 + \beta < 1, \quad \omega = 1 - \alpha - \gamma/2 - \beta$
- (ii)  $x_{i,t} = (1 - c/n)x_{i,t-1} + v_t$  for some  $c \geq 0, \quad v_t$ 와  $\epsilon_t$ 는 독립
- (iii)  $f(X'_{t-1}\beta) > 0 \quad \forall t$

첫 번째 가정에서  $\alpha + \gamma/2 + \beta < 1$ 의 가정은 본 논문에서 사용한 GJR-GARCH가 비정상성 모형(nonstationary model)인 IGARCH가 아닌 정상성에 기반한 모형(stationary model)임을 가정하는 것이

다. 그리고  $\omega = 1 - \alpha - \gamma/2 - \beta$ 의 가정은  $E[h_t] = 1$ 로 정규화(normalize)하기 위해서이며, 결과적으로 수익률 제곱의 장기적인 평균(즉 수익률  $y_t$ 의 비조건부 분산)은  $f(X'_{t-1}\beta)$ 에 의해 결정된다. 두 번째 가정은 경제/금융 변수인 외생적인 공변량 중 하나 이상의 공변량이 단위근 또는 준단위근을 가진다는 것이다. Han and Kristensen(2015)에서 밝혀졌듯이 지속성이 높은 외생 공변량을 포함한 경우 금융시계열의 비조건부 분산이 시간에 따라 변화하는 비정상성이 설명될 수 있다. 만약 모형에 포함된 외생 공변량이 모두 안정적이라면  $y_t$ 의 비조건부 분산은 일정(constant)할 것이다. 세 번째 가정은 변동성은 음이 아니기에 당연한 가정이라 할 수 있다.

IGARCH가 아닌 정상성을 가지는 GARCH 형태의 모형의 특징은 조건부 분산은 시간에 따라 변하지만 비조건부 분산은 일정하다는 것이다. 그러나 현실의 많은 경우에서 주식을 비롯한 자산의 수익률의 비조건부 분산은 경기 상황과 금융시장 여건의 변화에 영향을 받아 일정하지 않고 시간에 따라 변화하는 모습을 보임을 쉽게 확인할 수 있다. 따라서 반모수 단일지표 모형에서는 이러한 특징을 잘 설명할 수 있다. 모형에서  $\theta$ 와  $\beta$ 는 모수적으로  $f(\cdot)$ 는 비모수적으로 추정하므로 이 모형을 반모수 단일지표 모형이라고 부른다.

반모수 단일지표 모형을 선행연구와 비교하자면, 우선 spline-GARCH 모형 및 time-varying GARCH 모형 등은 변동성의 장기변동부분을 시간의 함수로 정의하여 실물경제 또는 금융시장의 상황을 반영하는 경제/금융지표를 외생 공변량으로 직접 사용하지 않고 있다. 이 중 Engle and Rangel(2008), Hafner and Linton(2010), Feng and Sun(2013), Koo and Linton(2013)은 변동성을  $\sigma_t^2 = h_t \times f(t/n)$ 로 정의하고  $f(t/n)$  부분을 비모수적으로 추정하고 있으며, Amado and Teräsvirta(2013)은 함수  $f(\cdot)$ 를 logistic 함수로 정의하고 있다. 따라서 이들 모형들은 구체적으로 어떠한 경제/금융지표들이 해당 금융시계열의 변동성에 영향을 미치는지에 대한 정보를 제공하지 못하는 한계가 있다. 또한 이들 모형들은 이용 가능한 표본 기간 내 예측(in-sample fitting)에만 적용되며 특히 변동성 관련 연구의 중요한 목적인 표본 기간 외 예측(out-of-sample forecasting)에는 사용될 수 없는 한계를 지니고 있다. 이에 반해 반모수 단일지표 모형은 경제/금융지표를 외생변수로 직접 이용하여 변동성의 장기변동부분을 모형화하기 때문에 이러한 한계를 극복할 수 있다. 또한 spline-GARCH 모형의 경우 Hafner and Linton(2010)의 방법(kernel based method)으로 추정할 때 과대적합(overfitting)이 발생하는 문제가 있다.

그리고 time-varying GARCH 모형과 GARCH-MIDAS 모형은 각각 logistic 함수 및 Beta weighting 을 이용한 함수를 이용하여 장기변동요인을 모수적으로 추정하는데, 이에 따라 모형 설정 오류의 우려가 있다. 그리고 time-varying GARCH 모형의 경우 상태(regime)의 개수를 검정해야하는 번거로움이 있으며 추정시간이 오래 걸리는 한계가 있다. 이에 반해 반모수 단일지표 모형은 장기변동요인을 비모수적으로 추정하기 때문에 모형 설정 오류의 우려를 극복하고 있으며 추정에도 시간이 오래 걸리지 않는 장점을 지니고 있다.

마지막으로, 관련 선행 연구 중 (additive) GARCH-X 모델을 간단히 평가한 후 선행연구와의 비교를 마무리하고자 한다. 변동성 모형화에 있어 외생 공변량을 이용하고자 하는 시도는 오래전부터 이어져 왔다. 구체적으로는 GARCH 형태의 변동성 모형에 외생 공변량을 더하기(additive) 형태로 결합한 GARCH-X 모형이 주로 이용되어 왔다(Han and Kristensen(2014)에 인용된 논문들 참조). 그런데, GARCH-X 모형의 경우 GARCH 모형 내에 포함된 외생 공변량의 함수형태를 조금만 복잡하게 설정하여도 추정이 어려워지는 단점이 있으며 또한 외생 공변량의 함수를 비모수적으로 추정하는 것은 너무도 복잡하여 사실상 불가능한 점 등의 한계가 있다. 본 논문에서 제안하는 모형은 외생변수가 곱하는 형태로 포함되어 이를 비모수적으로 추정하는 것이 가능하므로 기존의 GARCH-X 모형의 한계를 극복할 수 있다.

## 2.2 추정 방법

반모수 단일지표 모형에서 추정해야 할 것은 GJR-GARCH의 모수들인  $\theta = [\alpha, \beta, \gamma]$ 와 단일지표((pure) single index) 모형에서의  $\beta$ 와  $f(\cdot)$ 이다. 전체적인 추정 방법은 Hafner and Linton(2010)과 Han and Kristensen(2015)의 반모수 추정 방법(kernel based method)과 유사하다. 즉, 다음의 로그 우도 함수(log likelihood function)

$$L_n(\theta, \beta, f) = \frac{1}{n} \sum_{t=1}^n \left[ \log[f(X'_{t-1}\beta)h_t(\theta)] + \frac{y_t^2}{f(X'_{t-1}\beta)h_t(\theta)} \right]$$

를 극대화하는  $\theta$ 와  $\beta$  그리고  $f(\cdot)$ 을 찾는 것이다. 전체적인 추정 방법은 다음과 같은 방법을 따라 순차적으로  $f(X'_{t-1}\beta)$ 와  $h_t$ 를 추정한다.

1. 1단계에서는 일반적인 단일지표 모형을 추정하는 방법과 동일한 방법으로  $f(X'_{t-1}\beta)$ 을 먼저 추정한다. 즉, 위 식 (1)으로부터 도출된 다음의

$$y_t^2 = f(X'_{t-1}\beta) + u_t, \quad u_t := f(X'_{t-1}\beta)\{h_t e_t^2 - 1\}$$

비모수 회귀 함수(nonparametric regression function)를 추정하는 것에서 출발한다.

이 회귀 함수에서 먼저  $\beta$ 부터 추정하는데,  $\beta$ 를 추정하는 방법은 일반적인 최소자승법(least square)과 유사하게 추정하며 이는 비모수적인 기법을 통해 해결할 수 있다. 본 논문에서는 Ichimura (1993)의 반모수 최소자승법(semiparametric least square)을 통해  $\beta$ 를 도출하였다.  $f(\cdot)$ 에 대해서는 leave-one-out estimator를 사용하여 반모수 최소자승법을 통해  $\beta$ 를 추정하는 식은 다음과 같다.

$$\tilde{\beta} = \operatorname{argmin}_{\beta} \frac{1}{n} \sum_{t=1}^n [y_t^2 - \check{f}(X'_{t-1}\beta)]^2$$

$$\check{f}(x_i\beta) = \frac{\sum_{j \neq i} y_j^2 K_h(x_i\beta - x_j\beta)}{\sum_{j \neq i} K_h(x_i\beta - x_j\beta)} \text{ where } K_h(s) = h^{-1}K(s/h)$$

$x_i$ 과  $x_j$ 는 각각  $X'_{t-1}$ 의  $i$ 번째,  $j$ 번째 표본이다. 이와 같은 방법으로  $\tilde{\beta}$ 를 추정하고 추정된  $\tilde{\beta}$ 를 이용하여 leave-one-out estimator로  $\tilde{f}(X'_{t-1}\tilde{\beta})$ 를 추정하면 다음과 같다.

$$\tilde{f}(x_i\tilde{\beta}) = \frac{\sum_{j \neq i} y_j^2 K_h(x_i\tilde{\beta} - x_j\tilde{\beta})}{\sum_{j \neq i} K_h(x_i\tilde{\beta} - x_j\tilde{\beta})}$$

2. 2단계에서는 앞서 1단계에서 추정한  $\tilde{f}(X'_{t-1}\tilde{\beta})$ 를 이용하여  $h_t(\theta)$ 를 추정한다. 1단계에서  $\tilde{\beta}$ 와  $\tilde{f}(\cdot)$ 를 추정하였으므로, 이를 바탕으로 하여  $\tilde{\beta}$ 와  $\tilde{f}(\cdot)$  하에서 로그 우도 함수의 최적화 문제의 해인  $\tilde{\theta}$ 를 추정한다.

$$\tilde{\theta} = \underset{\theta}{\operatorname{argmax}} L_n(\theta, \tilde{\beta}, \tilde{f})$$

추정된  $\tilde{\theta}$ 을 바탕으로  $\tilde{h}_t$ 를 도출한다.

3. 3단계에서는 2단계에서 도출한  $\tilde{h}_t$ 를 이용하여 1단계의 방법과 동일하게 추정을 한다. 다음과 같이

$$\tilde{y}_t^2 = \frac{y_t^2}{\tilde{h}_t}$$

$y_t^2$ 을 2단계에서 추정한  $\tilde{h}_t$ 로 나눠줌으로써 새로운  $\tilde{y}_t^2$ 를 정의하고  $y_t^2$ 을 대신하여 추정에 사용한다. 앞서 제시한 1단계에서의 모든 추정 과정을  $y_t^2$  대신  $\tilde{y}_t^2$ 에 대하여 다시 추정하고 새롭게  $\hat{\beta}$ 과  $\hat{f}(\cdot)$ 을 도출한다.

$$\tilde{y}_t^2 = f(X'_{t-1}\beta) + \tilde{u}_t, \quad \tilde{u}_t := f(X'_{t-1}\beta)\{\epsilon_t^2 - 1\}$$

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \frac{1}{n} \sum_{t=1}^n [\tilde{y}_t^2 - \check{f}(X'_{t-1}\beta)]^2 \text{ where } \check{f}(x_i\beta) = \frac{\sum_{j \neq i} \tilde{y}_j^2 K_h(x_i\beta - x_j\beta)}{\sum_{j \neq i} K_h(x_i\beta - x_j\beta)}$$

$$\hat{f}(x_i\hat{\beta}) = \frac{\sum_{j \neq i} \tilde{y}_j^2 K_h(x_i\hat{\beta} - x_j\hat{\beta})}{\sum_{j \neq i} K_h(x_i\hat{\beta} - x_j\hat{\beta})}$$

4. 4단계에서도 앞서와 유사하게 2단계의 과정을 똑같이 시행한다. 앞서 3단계에서 추정한  $\hat{\beta}$ 과  $\hat{f}(\cdot)$ 을 이용하여

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} L_n(\theta, \hat{\beta}, \hat{f})$$

위 식을 만족하는  $\hat{\theta}$ 을 추정한다.

위와 같은 방법으로 단일지표 모형에서의  $\beta$ 와  $f(\cdot)$ , GARCH 형태의 모형에서의  $\theta$ 를 순서대로 지속적으로 추정하고, 위 과정을 통해 추정된  $\theta$ 가 적절한 수준에서 수렴할 때까지 위의 과정을 지속적으로

반복하여 최적의 추정치들을 찾는다.<sup>3</sup> 본 논문에서의  $K$ 는 가우시안 커널 함수(Gaussian kernel function)를 사용하였다. 띠폭(bandwidth)  $h$ 의 선택 방법으로는 교차타당성(cross-validation) 방법을 이용하여 QLIK 손실함수,  $1/n \sum_{t=1}^n [y_t^2 / \hat{f}(X'_{t-1}\beta) + \log(\hat{f}(X'_{t-1}\beta))]$ 를 최소화하는  $h$ 를 선택하였다.

### 3 실증 적용

#### 3.1 데이터, 모형들과 추정 방법

본 논문에서 설정한 표본기간(sample period)은 2004년 1월 5일부터 2014년 8월 29일까지로, 표본 수는 2,643이다. 본 연구에서는 한국 주식시장의 KOSPI 지수의 일별 로그 수익률을 사용하였고, 각 수익률에서 수익률의 표본 평균을 뺀으로써 평균이 제거된 KOSPI의 수익률을  $y_t$ 로 사용하였다. 표본 기간 동안의 KOSPI의 수익률과 그 제곱의 그래프는 다음과 같다.

Figure 1

두 그래프에서 음영 처리된 부분은 통계청에서 발표하는 경기 순환 주기표를 기준으로 하여 한국의 경기 침체기를 표시한 것이다.<sup>4</sup> 두 그래프를 통해서 봤을 때, 경기 침체기에는 다른 기간에 비해 변동성이 더 큰 경향이 있는 것을 확인할 수 있다. 또한 시간에 따라 변동성의 수준이 크게 변하는 것으로 볼 때 변동성에 대한 더 적합한 분석을 위해서는 시간에 따라 변화하는 비조건부 분산을 반영해야 함을 알 수 있다. 본 논문에서는 이를 위해 경기 상황 및 금융시장 여건을 반영하는 경제/금융 변수들을 도입하여 시간에 따라 변하는 비조건부 분산(즉 변동성의 장기적 변동 부분)을 설명하고자 하였다.

시간에 따라 변화하는 비조건부 분산을 설명하기 위해서 어떠한 변수들을 외생적인 공변량  $X_{t-1}$ 으로 모형에 사용할 수 있고 또한 설명력을 더 높여줄 수 있을지에 대한 답을 찾기 위해 여러 경제/금융 변수들을 고려해 보았다. 본 논문에서는 경제/금융 변수로 경기동행지수, 경기선행지수, VKOSPI지수, VIX지수, 국민은행 주택매매가격지수, 원/달러 환율, 국제 금 가격 지수(London gold fixing price 3pm), 단기 금리로 콜금리(1일), CD금리(91일)과 장기 금리로 국고채 금리(3년, 5년), 장·단기 금리 차(국고채 금리(3년,5년)-콜금리(1일), 국고채 금리(3년, 5년)-CD금리(91일)), 총 거래량(trading volume)을 고려해 보았다. 이 변수들은 모두 주식시장 수익률의 변동성과 관련이 있는 것으로 알려져 있으며 실제 실증 분석에서도 널리 활용되고 있다. 이와 같은 다양한 변수들 중에 어떠한 변수들이 변동성을 예측하는 데 가장

<sup>3</sup> 이러한 모형의 추정치에 대한 일치성과 접근 분포에 대한 계량경제학 이론을 확립하는 것이 필요한데, 이는 외생 공변량이 비정상 계열을 포함할 경우 기존의 이론들을 적용하기에 쉽지 않은 일이다. 따라서 저자들의 후속 연구에서 관련된 계량경제학 이론을 확립하고자 한다.

<sup>4</sup> 제시된 경기 침체기는 통계청에서 발표하는 경기 순환 주기표를 기준으로 제8순환기인 2002년 12월에서 2005년 4월, 제9순환기의 침체기인 2008년 1월에서 2009년 2월까지를 표시하였다. 제10순환기의 경우 경기 순환의 정점과 저점이 확정되지 않았다. 가장 최근의 발표 자료인 2014년 6월의 발표 자료에 따르면 잠정적으로 정점은 2011년 8월로 설정되었고 저점은 설정되지 않아서 2011년 8월 이후에 경기동행지수, 경기선행지수가 가장 작은 값을 나타내는 2012년 10월까지를 경기 침체기로 설정하였다.

큰 도움을 줄 수 있는지에 대한 선택 기준으로는 3.3장에서 설명되는 QLIKE 손실함수를 기준으로 삼았다. 위에서 제시한 변수들 중 여러 조합<sup>5</sup>을 고려하였고, 각각의 조합들마다 QLIKE 손실함수를 도출하여 가장 낮은 QLIKE 손실함수 값을 가지는 변수들의 조합을 찾았다. 그 결과 경기동행지수, VKOSPI지수, 주택매매가격지수 이 세 개의 변수를 사용한 경우가 고려한 조합들 중 가장 낮은 QLIKE 값을 나타냈다.<sup>6</sup>

이와 같이 채택된 경제/금융 변수들인 경기동행지수, VKOSPI지수, 주택매매가격지수는 이미 많은 연구에서 주식시장을 설명할 수 있는 중요한 변수들로 인식되어 사용되고 있다. 경기동행지수와 주택매매가격지수는 실물 경제를 나타내는 변수들이다. 이들 변수가 낮은 값을 나타내는 것은 실물 경제가 침체에 접어들었다는 것을 의미하며 실물 경제의 침체는 소비의 감소와 투자 심리 위축으로 이어져 기업들에게 악영향을 미친다. 이는 금융경제에도 부정적 영향을 끼치게 되어 기업의 주가가 하락하고 이로 인한 변동성의 증대를 야기한다. 먼저 경기동행지수에 관련해서는 Hamilton and Lin (1996)의 연구를 비롯하여 많은 연구에서 경기 침체가 주식시장 수익률의 변동성이 더 심해지는 것을 야기하는 주된 요인 중 하나라고 밝히고 있다. 다음으로 VKOSPI지수는 한국거래소(Korea Exchange)에서 산출한 KOSPI 200 옵션 시장의 변동성 지수이다. VIX지수와 VKOSPI지수와 같은 변동성 지수는 시장의 옵션 만기까지의 향후 변동성에 대한 기대치로 시장의 미래 변동성에 대해서 가장 좋은 추정치로 알려져 있다. 마지막으로 주택매매가격지수의 경우, 주식시장과 함께 우리나라에서 일반인이 투자하는 대표적인 두 자산 시장 중 하나인 부동산 시장의 흐름을 반영하는 변수이다. 주식시장과 부동산 시장은 널리 알려진 것처럼 모두 경기 상황에 민감하게 반응하고 서로 관련이 있음이 알려져 있다. Conrad and Loch(2014)에 따르면, 주택 시장을 반영할 수 있는 변수는 경기 상황과 주식시장의 선행 지수로서 작용할 수 있다고 밝히고 있다. 우리나라에서의 주식시장과 부동산 시장 각각의 연구에 대한 높은 관심과 달리 주식시장과 부동산 시장에 대한 연구는 상대적으로 많이 이루어지지 않았는데, 최근의 임병진·한성운(2009)의 연구에서는 주택매매가격지수와 KOSPI가 공적분(cointegration)관계가 있음을 밝히고 벡터오차수정모형(vector error correction model)을 이용하여 서로 유의한 관계가 있음을 밝혔다.

위 변수들 중 경기동행지수, 주택매매가격지수의 경우 월별 데이터로 KOSPI의 수익률과 비교했을 때 저빈도 데이터(low frequency data)이다. 따라서 이들 데이터를 함께 사용하였을 때 발생하는 데이터에서의 혼합 빈도(mixed frequency) 문제를 해결하기 위하여 월별 데이터를 선형 보간법(linear interpolation)을 이용<sup>7</sup>하여 일별 데이터로 변환하였다. 선형 보간법을 이용하여 일별 데이터로 변환시킨 경제/금융 변수들의 그래프는 아래와 같다.

**Figure 2**

<sup>5</sup>경제/금융 변수가 1개인 경우와 변수가 2개인 경우, 변수가 3개인 경우까지의 가능한 모든 조합을 고려해보았다.

<sup>6</sup>이러한 방법으로 특정 변수들을 선택하는 대신 변수들의 common factor를 추출하여 이를 외생변수로 사용하는 것도 가능할 것이다. Common factor를 이용하는 방법이 실증적으로 유용한지를 살펴보는 것은 후속 연구에서 다룰 계획이다.

<sup>7</sup>데이터에 존재하는 혼합 빈도 문제를 해결하기 위해 월별 데이터를 해당 달의 수준(level)을 그대로 이용하는 방법과 선형 보간, 선형추세(linear trend), 선형 스플라인(linear spline), 커널 평탄(kernal smoothing)을 이용하여 일별 데이터로 변환하는 방법들을 이용해보았고, 이용한 방법들에 관계없이 강건성을 보였다.

2.1장에서 설명하였듯이 반모수 단일지표 모형이 금융시계열의 비조건부 분산이 시간에 따라 변화하는 비정상성을 설명하기 위해서는 사용하는 외생적 공변량 중 최소 하나 이상이 단위근 또는 준단위근을 가져야 한다. 따라서 위에서 언급한 경제/금융 변수들에 대한 단위근 검정을 실시하였으며 그 결과는 다음과 같다.

**Table 1**

단위근 검정은 절편만 고려한 경우와 절편과 시간에 따른 선형 추세를 고려한 경우 두 가지에 대해 검정하였다. 우선 검정에 앞서 자기회귀모형(autoregressive model)인 AR(1) 모형을 통해 자기회귀계수(autoregressive coefficient)가 1에 가까운지를 확인해 보았다. 고려한 모든 경우에서 자기회귀계수는 1에 가깝게 나타났고, 잠정적으로 단위근이 존재할 가능성이 있음을 확인하였다. 변수가 단위근을 가지는지를 확인하기 위해 두 가지 검정을 사용하였다. 첫 번째 검정 방법은 Augmented Dickey-Fuller(ADF) 검정이고, 두 번째 검정 방법은 Kwiatkowski-Phillips-Schmidt-Shin (KPSS) 검정이다. 먼저 경기동행지수의 경우 절편만 고려한 경우와 절편과 시간에 따른 추세를 고려한 경우 모두에 대해 두 가지 검정 방법 모두 단위근을 가진다는 결과를 도출해냈다. 다음으로 VKOSPI지수의 경우 두 가지 경우 모두에서 적어도 준단위근을 가진다고 해석할 수 있다. 비록 ADF 검정은 단위근이 존재하지 않는다는 결과를 보이지만, KPSS 검정은 시계열이 정상적이지 않음을 나타내고 있으며 또한 자기회귀계수가 1에 가까운 점을 고려할 때 준단위근을 가진다고 볼 수 있다. 마지막으로 주택매매가격지수는 절편만을 고려했을 때는 두 검정 모두 단위근을 가진다는 동일한 검정 결과를 제시하는 등 VKOSPI지수와 마찬가지로 적어도 준단위근을 가진다고 결론지었다.

주식시장의 변동성 분석에 도움을 줄 수 있는 경제/금융 변수를 선택한 후에 다음으로 확인해 보아야 하는 것은 과연 반모수 단일지표 모형의 변동성 예측력이 기존의 다른 모형들과 비교했을 때 더 나은 예측력을 가지는지에 대해서이다. 이를 확인하기 위해서 기존의 변동성을 분석하는 데 이용되는 모형들인 GJR-GARCH 모형, GARCH-X 모형, GARCH-MIDAS 모형, 단일지표 모형과 본 논문에서 사용한 반모수 단일지표 모형까지 다섯 가지 모형의 변동성 예측력을 비교해 보았다. 변동성을 설명할 수 있는 위의 모형들은 다음과 같이 정의된다.

$$y_t = \sigma_t \epsilon_t, \quad \epsilon_t \stackrel{iid}{\sim} (0, 1)$$

$$x_1 : \text{경기동행지수}, x_2 : \text{VKOSPI지수}, x_3 : \text{주택매매가격지수}, X' = [x_1, x_2, x_3]$$

---

### 1. GJR-GARCH

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma y_{t-1}^2 \mathbf{I}(y_{t-1} < 0)$$

---

### 2. GARCH-X

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma y_{t-1}^2 \mathbf{I}(y_{t-1} < 0) + \pi_1 \frac{1}{x_{1,t-1}} + \pi_2 x_{2,t-1} + \pi_3 \frac{1}{x_{3,t-1}}$$

---

### 3. GARCH-MIDAS

$$\sigma_{i,t}^2 = h_{i,t} \times f_t$$

where  $h_{i,t} = 1 - \alpha - \gamma/2 - \beta + \alpha h_{i-1,t} \epsilon_{i-1,t}^2 + \beta h_{i-1,t} + \gamma h_{i-1,t} \epsilon_{i-1,t}^2 \mathbf{I}(\sqrt{h_{i-1,t}} \epsilon_{i-1,t} < 0)$

$$f_t = m + \theta_1 \sum_{k=1}^K \phi_{1,k}(\omega) \Delta x_{1,t-k} + \theta_2 \sum_{k=1}^K \phi_{2,k}(\omega) \Delta x_{2,t-k} + \theta_3 \sum_{k=1}^K \phi_{3,k}(\omega) \Delta x_{3,t-k}$$

---

### 4. 단일지표

$$\sigma_t^2 = f(X'_{t-1} \beta)$$

---

### 5. 반모수 단일지표

$$\sigma_t^2 = h_t \times f(X'_{t-1} \beta)$$

where  $h_t(\theta) = 1 - \alpha - \gamma/2 - \beta + \alpha h_{t-1} \epsilon_{t-1}^2 + \beta h_{t-1} + \gamma h_{t-1} \epsilon_{t-1}^2 \mathbf{I}(\sqrt{h_{t-1}} \epsilon_{t-1} < 0)$

---

GJR-GARCH 모형의 경우는 앞서 밝힌 바와 같이 표준적인 GARCH 형태의 모형 중 하나로 비대칭적인 레버리지 효과를 반영할 수 있다. GARCH-X 모형의 경우는 추가적인 공변량을 도입하여 표본 내 예측력과 표본 외 예측력을 개선할 수 있는데, 본 논문에서의  $x_{i,t-1}$ 을 모형에 사용할 때  $\pi$ 의 부호를 양으로 설정하기 위해서 VKOSPI지수의 경우는  $x_{i,t-1}$  그 자체를 사용하고 경기동행지수와 주택매매가격지수의 경우는 역수  $1/x_{i,t-1}$ 를 사용하였다. GARCH-MIDAS 모형의 경우는 다른 모형들에서 시간 변수  $t$ 가 각각의 일별을 지칭하는 것과 다르게 GARCH-MIDAS 모형에서는 시간 변수가  $t$ 와  $i$ 가 존재한다.  $t$ 는 특정 기간을 지칭한다. 본 논문에서는  $t$ 를 1개월로 할당하였고  $i$ 는 특정 기간  $t$ 에 해당하는 1일로 할당하였다. 즉,  $h_{i,t}$ 로 표기된 GJR-GARCH 모형은 일별로 변화하는 반면, 곱해지는  $f_t$ 의 경우 월별로 변화하는데, 이는 GARCH 형태의 모형으로 단기 변동성을 설명하고 MIDAS 방법을 사용한  $f_t$ 로 변화하는 장기 변동성을 설명하고자 한 것이다. 장기 변동 부분인  $f_t$ 는 월별 경제/금융 변수의 차분(difference),  $\Delta x_{i,t-k}$ 의 함수이며, 베타가중 함수(beta weight function),  $\phi_k(\omega) = [(k/K)^{\omega_1-1}(1-k/K)^{\omega_2-1}]/[\sum_{k=1}^K (j/K)^{\omega_1-1}(1-j/K)^{\omega_2-1}]$ 를 통해 도출된다. 이는 경제/금융 변수를 차분한 것의 가중합을 구하는 것과 유사하다. 단일지표 모형은 반모수 단일지표 모형에서 GARCH 형태의 모형을 제외한 모형으로서 추정하는 방법은 앞에서 제시한 반모수 단일지표 모형의 추정 방법의 1단계와 동일하게 추정한다. 즉 경제/금융 변수들만을 이용하여 주식시장의 변동성을 분석하는 모형이다.

### 3.2 추정 결과

먼저 반모수 단일지표 모형을 통해 추정된  $\hat{f}(X'_{t-1}\hat{\beta})$ 과  $\hat{h}_t$ 의 그래프는 다음과 같다.

Figure 3

그래프를 살펴보았을 때 시간에 따라 변화하는 비조건부 분산을 장기 변동 부분인  $f(X'_{t-1}\beta)$ 를 통해 분석하고 그 이외의 단기에 변화하는 조건부 분산을  $h_t$ 가 분석할 수 있도록 모형화한 것과 일치하게  $\hat{f}(X'_{t-1}\hat{\beta})$ 은 시간에 따라 변화하는 비조건부 분산을 설명하고 있고 그로 인해  $\hat{h}_t$ 은 정상성을 가지는 것으로 보인다. 특히 음영 처리된 부분인 경기 침체기 때,  $\hat{f}(X'_{t-1}\hat{\beta})$ 은 다른 기간에 비해 더 변동성이 높은 모습을 보이면서 경기 상황에 따라 변화하는 장기 변동 부분을 설명하고 있고 반면  $\hat{h}_t$ 는 시간의 변화에 영향을 받지 않는 안정적 시계열인 것처럼 보인다.

$\hat{f}(X'_{t-1}\hat{\beta})$ 과  $\hat{h}_t$ 에 대해 더 자세하게 해석하고 그 결과가 모형을 설정할 때 기대한 바와 일치하는지 확인하기 위해서는 본 연구에서 사용한 반모수 단일지표 모형에서 추정된  $\hat{h}_t$ 과  $\hat{f}(X'_{t-1}\hat{\beta})$ 의 여러 성질들을 살펴보는 것이 필요하다.  $y_t^2$ 에서 일반적인 GJR-GARCH 모형으로 추정한 추정치들과  $y_t^2$ 을  $\hat{f}(X'_{t-1}\hat{\beta})$ 로 나눈  $y_t^2/\hat{f}(X'_{t-1}\hat{\beta})$ 을 GJR-GARCH 모형으로 추정한 추정치들을 도출한 후에 두 모형에서의 추정치들에 어떠한 차이가 있는지를 확인한다. 즉, 일반적인 GJR-GARCH 모형에서의 추정된 모수와 본 논문에서의 반모수 단일지표 모형에서 단기 변동성 부분을 나타내는  $h_t$ 에서의 추정된 모수를 비교하는 것이다. 앞서 제시한 변동성 모형들의 추정 결과는 다음과 같다.

Table 2

각 모형에서 단기변동부분을 GJR-GARCH 모형으로 추정한 결과를 보면 반모수 단일지표 모형의  $\alpha$ 를 제외하고는 모두 1% 유의수준에서 유의한 것으로 나타났다. GARCH-X 모형에서  $\hat{\pi}_i$ 는 모두 양수로 나타났다. 경기동행지수 역수의 계수인  $\pi_1$ 과 주택매매가격지수 역수의 계수인  $\pi_3$ 가 모두 양수로 추정되었다는 것은 변동성과 경기동행지수, 주택매매가격지수는 음의 관계를 가지고 있음을 의미하고 마찬가지로 VKOSPI지수의 계수인  $\pi_2$ 도 양수로 추정되었다는 것은 VKOSPI지수와 변동성은 양의 관계가 있음을 의미한다. GARCH-MIDAS의 경우에  $\hat{\theta}_i$ 가 모두 음수로 나타났다. 일반적으로 변수의 차분을 통해 증가/감소하는 추세를 알 수 있다. GARCH-MIDAS에서 장기변동부분에서 사용한 변수의 차분의 가중합이 변수의 증가/감소하는 추세를 나타낸다고 했을 때 경기동행지수와 주택매매가격지수의 추정된 계수가 음수인 것은 이들 변수의 변화와 변동성이 음의 관계를 가지고 있음을 나타내기 때문에 경제적 사실과 일치한다. 그러나 VKOSPI의 변화도 변동성과 음의 관계를 나타내고 있는데 이는 VKOSPI와 변동성이 정의 관계를 가지고 있다는 일반적인 견해에 반하는 것이다.

추정된 결과를 바탕으로 변동성의 지속성 정도(persistence measure)를 나타내는  $\alpha + 2/\gamma + \beta$ 를 계산하면 GJR-GARCH 모형의 경우 0.9814로 1에 굉장히 가깝게 나타났고 GARCH-X 모형의 경우 0.9153,

GARCH-MIDAS 모형의 경우 0.8876, 반모수 단일지표 모형의 경우 0.8449로 GJR-GARCH 모형보다 지속성 정도의 값이 1보다 크게 작아진 것을 알 수 있다. 경제/금융 변수를 도입한 모형들의  $h_t$ 에서 추정된 지속성 정도가 GJR-GARCH 모형보다 1에서 더 멀다는 것은 경제/금융 변수를 도입하여 변동성의 지속적인 특성을 잘 분석할 수 있다는 것을 의미한다. 특히 반모수 단일지표 모형과 GJR-GARCH 모형과 비교했을 때 이와 같은 추정치의 결과는 두 경우에서 각각 도출한 자기상관함수(autocorrelation function)를 통해서도 확인할 수 있는 결과와 일치한다.  $y_t^2$ 과 이를  $\hat{f}(X'_{t-1}\hat{\beta})$ 로 나눈  $y_t^2/\hat{f}(X'_{t-1}\hat{\beta})$ 를 이용하여 각각 자기상관함수를 도출하였고, 이는 아래와 같다.

**Figure 4**

그래프 (a)와 그래프 (b)를 비교해보면 그래프 (a)에서는 시차(lag)가 40까지 상당히 커져도  $y_t^2$ 의 자기상관계수(autocorrelation)이 유의하게 유지되는 반면 그래프 (b)에서는  $y_t^2$ 을  $\hat{f}(X'_{t-1}\hat{\beta})$ 로 나뉜 경우 자기상관계수가 빠른 속도로 줄어드는 것을 확인할 수 있다. 이는 수익률  $y_t^2$ 의 장기 기억 특성(long memory property)이  $\hat{f}(X'_{t-1}\hat{\beta})$ 로 나뉘춤으로써 사라진 것인데, 변동성의 장기적 변동 부분이 변동성의 장기 기억 특성을 만들어내는 요인임을 의미한다고 볼 수 있다. Han and Kristensen(2015)은 SEMIX 모형을 이용하여 이를 이론적으로 밝히고 있다.  $y_t$ 와  $y_t/\sqrt{\hat{f}(X'_{t-1}\hat{\beta})}$ 에서 각각 GJR-GARCH 모형을 추정하여 얻어진 변동성,  $\hat{h}_t$ 의 그래프 비교에서도 유사한 성질을 확인할 수 있다.

**Figure 5**

Figure 5에서 그래프 (a)는 GJR-GARCH 모형의  $\hat{h}_t < y_t$ 에서 GJR-GARCH 모형을 추정한 것이고 그래프 (b)의 경우는 반모수 단일지표 모형에서의  $\hat{h}_t$ 으로 이는  $y_t/\sqrt{\hat{f}(X'_{t-1}\hat{\beta})}$ 에서 GJR-GARCH 모형을 추정한 것이다. 그래프 (a)에 제시된  $y_t$ 로부터 추정한  $\hat{h}_t$ 는 그래프 (b)에 나타난  $y_t/\sqrt{\hat{f}(X'_{t-1}\hat{\beta})}$ 로부터 추정한  $\hat{h}_t$ 과 비교했을 때 더 비정상적인 그래프에 가까운 것을 확인할 수 있다. 이는 KOSPI 지수의 수익률에 내재하는 비정상성(시간에 따라 변화하는 비조건부 분산)을  $\hat{f}(X'_{t-1}\hat{\beta})$ 이 잘 설명한 결과라고 할 수 있다

다음으로는 단일지표 모형과 반모수 단일지표 모형에서 추정된  $\hat{f}(X'_{t-1}\hat{\beta})$ 에서 각각의 경제/금융 변수  $x_i$ 의 추정된 계수에 해당하는  $\hat{\beta}$ 의 부호를 확인해 본다. 만약 추정된 계수  $\hat{\beta}$ 의 부호가 양수라면 해당 변수는 변동성과 같은 방향으로 움직이는 변수임을 의미하고  $\hat{\beta}$ 의 부호가 음수라면 반대 방향으로 움직이는 변수이다. 반모수 단일지표 모형에서 사용한 외생적인 공변량인 경기동행지수, VKOSPI지수와 주택매매가격지수의 추정된 계수는 다음과 같다.

**Table 3**

추정 결과에 따르면 두 모형에서 추정된 계수들은 차이가 있지만, 크게 경기동행지수와 주택매매가격지수의 계수는 음수로 VKOSPI지수의 계수는 양수로 나타났다. GARCH-X 모형에서 추정한 결과와 마찬가지로 이는 우리가 알고 있는 일반적인 경제적 사실과 일치한다. 주식시장 수익률의 변동성은 일반적으로 불경기에 더 변동이 심한 것으로 알려져 있다. 경기동행지수와 주택매매가격지수는 경기순행적인 변수들로 값이 높으면 경기가 호황이고 낮으면 불황이라고 할 수 있다. 따라서 경기동행지수와 주택매매가격지수가 높을 때는 호황이기 때문에 변동이 덜하고 낮을 때는 변동이 심한데 이는 반모수 단일지표 모형에서 추정된 결과와 일치한다. 마찬가지로 VKOSPI지수의 경우 옵션 시장의 변동성 지수로 VKOSPI가 높다면 KOSPI 수익률의 변동성도 높게 나타나고 낮다면 변동성도 낮게 나타날 것이다. 따라서 VKOSPI 지수의 추정된 계수도 일반적인 경제적 사실과 일치한다고 할 수 있다.

### 3.3 예측 평가기준 및 결과

앞서 여러 모형들 중 어떤 모형이 가장 주식시장 수익률의 변동성을 잘 설명하는지에 대한 평가를 하기 위해서 앞에서의 경제/금융 변수의 최적의 조합을 고려할 때와 같은 방법으로, 모형에서 추정된 변동성과 실제 변동성과의 손실함수로서 QLIKE 손실함수를 평가기준으로 삼았다. 일반적으로 변동성 예측에서의 손실함수는 실제 변동성  $\sigma_t^2$ 과 모형에서 추정한 변동성  $\hat{\sigma}_t^2$ 의 함수로 정의된다. 그러나 현실에서 실제 변동성  $\sigma_t^2$ 을 안다는 것은 불가능하다. 따라서 실제 값 대신 이에 대한 대응치를 사용하여야 한다. QLIKE 손실함수는 실제 변동성을 알 수 없기에 사용해야 하는 잡음(noise)이 있는 변동성 대응치(volatility proxy)에 대해 강건하고 동차(homogeneous) 손실함수 중 DMW 검정을 시행하였을 때 가장 검증력이 좋은 것으로 알려져 있다.(Patton and Sheppard(2009)) 따라서 본 논문에서는 QLIKE 손실함수를 평가기준으로 정하였고 QLIKE 손실함수는 다음과 같이 정의된다.

$$L(\hat{\sigma}_t^2, \sigma_t^2) = \frac{\sigma_t^2}{\hat{\sigma}_t^2} - \log \frac{\sigma_t^2}{\hat{\sigma}_t^2} - 1$$

앞에서 밝힌 바와 같이  $\sigma_t^2$ 은 실제 변동성이지만 관측이 불가능하기 때문에 이에 대한 대응치를 사용하였고 본 논문에서는 여러 대응치 중  $\sigma_t^2$ 로 실현 커널을 사용하였다. 기존에 변동성에 대한 대응치로 많이 사용되었던 실현 분산은 고빈도 데이터를 사용할 경우 시장 미시구조(market microstructure)에 대해 잡음이 있다는 단점이 있다. 하지만 이와는 다르게 Barndorff-Nielsen, et al.(2008)에 따르면 틱 자료(tick data)를 바탕으로 만들어진 실현 커널이 시장 미시구조 잡음에 대해 강건성이 있다고 밝히고 있다. 따라서 실현 커널을 이용함으로써 실현 분산을 사용하였을 때의 단점을 극복할 수 있다.  $\hat{\sigma}_t^2$ 은 모형에서 추정한 변동성의 추정치 또는 예측치이다. 따라서 앞서 제시한 다섯 모형들로부터 각각 변동성을 추정한 후 추정된 변동성과 실현 커널을 이용하여 QLIKE 값을 도출하여 이를 비교한다. 위의 모형들 중 가장 낮은 QLIKE 값을 가지는 모형을 더 예측력이 좋은 모형이라 판단한다.

다양한 모형들에서 QLIKE 손실함수들을 구한 다음의 과정으로 도출한 QLIKE 손실함수의 값들 간의 차이가 유의한지를 보기 위해서 변동성 예측에서 쌍별 비교(pairwise comparison)의 한 방법으로 DMW 검정을 사용한다. DMW 검정은 두 모형에서 각각 도출한 값의 차이를

$$d_t = L(\hat{\sigma}_{1t}^2, \sigma_t^2) - L(\hat{\sigma}_{2t}^2, \sigma_t^2)$$

로 정의한다. 본 논문에서는 손실함수  $L(\cdot)$ 로 앞에서 밝힌 QLIKE 손실함수를 사용하였다. DMW 검정의 귀무가설과 대립가설은 다음과 같다.

$$H_0 : E(d_t) = 0$$

$$H_1 : E(d_t) \neq 0$$

$E(d_t)$  대신  $d_t$ 의 표본 평균  $\bar{d}_T$ 을 사용하고,  $\bar{d}_T$ 는 예측치의 개수  $T$ 에 대해  $\bar{d}_T = 1/T \sum_{t=1}^T d_t$  와 같이 정의된다. 이를 바탕으로 한 DMW 통계량은 점근적으로 표준 정규 분포를 따름이 알려져 있다.

$$\text{Test statistic DMW} = \frac{\sqrt{T} \bar{d}_T}{\sqrt{\widehat{AVar}[\sqrt{T} \bar{d}_T]}} \xrightarrow{d} N(0, 1)$$

위 식에서  $AVar$ 은 점근 분산(asymptotic variance)를 의미한다. DMW 검정의 점근 분산은 Newey-West 추정치를 통해 추정되고 시차는  $[T^{1/3}]$ 으로 설정한다.

이와 같은 QLIKE 손실함수와 DMW 검정을 통해서 모형들 간의 예측력을 평가하기 위해서 먼저 다섯 모형들로부터 각각 QLIKE 손실함수의 값을 구하고 다음으로 손실함수의 값의 차이를 바탕으로 DMW 검정을 시행하고 검정 통계량이 유의한지를 확인한다. 본 논문에서는 표본 내 예측력 검정과 표본 외 예측력 검정에서 모두 QLIKE 값을 각각 구해 비교하고 손실함수 간의 차이를 DMW 검정을 통해 검정하였다.

먼저 표본 내 예측력 검정의 경우 전체 표본 기간 동안 앞에서 밝힌 다섯 모형들을 통해 변동성을 추정하여 QLIKE 손실함수를 도출한 후 DMW 검정을 시행한다. 전체 표본 기간은 2004년 1월 5일에서 2014년 8월 29일까지이지만 분석에 사용한 모형 중 GARCH-MIDAS 모형의 경우 과거 경제/금융 변수들의 가중합을 통해 장기변동부분을 분석한다. 이를 위해 MIDAS 시차(MIDAS lag)를 설정하는 것이 필요하다. MIDAS 시차는 기존의 연구들을 참고하여 6개월로 정하였다. GARCH-MIDAS 모형과 다른 모형들의 예측 기간을 동일하게 정하기 위하여 표본 내 예측력 검정에서의 표본 기간은 2004년 8월 2일부터 2014년 8월 29일까지 2502일 간의 데이터로 정하였다. 표본 내 예측력 검정에서 구한 모형 별 QLIKE 값과 DMW 검정 통계량은 다음과 같다.

**Table 4**

표본 내 예측력 검정의 결과를 비교해보면 먼저 QLIKE 손실함수의 값에 있어서 반모수 단일지표 모형이 가장 낮은 것으로 나타났다. 기존의 GJR-GARCH 모형과 GARCH-X 모형, GARCH-MIDAS 모형의 경우 QLIKE 값이 약 0.19 정도이고 단일지표 모형은 QLIKE 값이 약 0.17 정도인 것과 비교했을 때 반모수 단일지표 모형은 약 0.15 정도의 낮은 값을 가진다. GJR-GARCH 모형과 비교했을 때 GARCH-X 모형과 GARCH-MIDAS 모형, 반모수 단일지표 모형이 모두 낮은 QLIKE 값을 가지는 것은 분석에 이용한 세 가지의 경제/금융 변수들이 변동성 분석에 도움을 주기 때문이다. 동일하게 세 가지의 경제/금융 변수를 추가한 모형들이 서로 다른 예측력을 보이는 것은 함수의 형태가 다르기 때문이다. 이 세 모형들 중 반모수 단일지표 모형의 QLIKE 값이 가장 낮은 것으로 보았을 때, 경제/금융 변수를 더하거나 가중합으로 곱하는 형태의 모수적 추정인 반모수 단일지표 모형과 같이 경제/금융 변수가 변동성에 미치는 영향에 대한 함수의 형태를 비모수적인 방법으로 유연성 있게 추정하는 것이 변동성을 예측하는 데 가장 적합하다고 할 수 있다. 모형들 중 가장 낮은 QLIKE 값을 가지는 반모수 단일지표 모형을 기준으로 삼아 반모수 단일지표 모형의 QLIKE 값과 다른 모형과의 QLIKE 손실함수의 값의 차이의 유의성을 검정하기 위해 DMW 검정을 통해 검정하였다. 검정 결과는 표의 두 번째 행에 제시되어 있는 것처럼 모두 1% 유의수준에서 유의한 것으로 나타났다. 즉, 분석한 모형들 중 가장 낮은 QLIKE 손실함수의 값을 가지고 DMW 검정 통계량이 모두 유의하게 나타났으므로 표본 내 예측력 검정에서 반모수 단일지표 모형이 기존의 다른 모형들보다 주식시장 수익률의 변동성을 더욱 잘 설명하고 있는 것을 알 수 있다.

다음으로는 표본 외 예측력 검정을 살펴보고자 한다. 표본 외 예측 평가를 시행하는 방법으로 1008 일을 시간창(time-window)으로 정한 후 표본 이동 분석법(rolling window approach)을 통해 예측 평가를 시행하였다. 이 방법은 예를 들어 1부터 1008번째까지의 관측치, 1008개를 시간창으로 설정하여 이 데이터를 통해 1009번째 관측치를 예측하고 그 다음 단계로 2부터 1009번째까지의 1008개의 관측치를 사용하여 1010번째 관측치를 예측한다. 이와 같은 단계를 반복하여 표본 외 예측을 시행하는 방법을 말한다. 이 방법을 통해 표본 외 예측력 검정을 하였을 때 예측한 기간은 2010년 1월 4일에서 2014년 8월 29일까지 1155일 간의 데이터이다. 표본 외 예측력 검정에서 구한 모형 별 QLIKE 값과 DMW 검정 통계량은 다음과 같다.

**Table 5**

표본 내 예측력 검정과 마찬가지로 반모수 단일지표 모형의 QLIKE 값이 다섯 개의 모형들 중 가장 낮은 것으로 나타났다. 특히 표본 외 예측력 검정에서는 다른 모형들의 QLIKE 값이 0.2 이상이거나 약 0.2인 것에 비해 반모수 단일지표 모형에서의 QLIKE 값은 약 0.18로 다른 모형들에 비해 현저히 낮게 나타난 것을 알 수 있다. 표본 내 예측력 검정과 마찬가지로 가장 낮은 QLIKE 값을 가지는 반모수 단일지표 모형을 기준으로 삼아 반모수 단일지표 모형과 다른 모형들 간의 QLIKE 값의 차이에 대하여 DMW 검정을 시행하였고, DMW 검정 통계량도 모두 1% 수준에서 유의한 것으로 나타났다. 표본 내

예측력 검정과 표본 외 예측력 검정에서의 비교를 통해서 반모수 단일지표 모형이 표본 내 예측력 검정과 표본 외 예측력 검정 모두에서 기존의 다른 모형들보다 더 예측력이 뛰어남을 확인할 수 있다. 따라서 반모수 단일지표 모형을 통해 변동성을 분석했을 때 시간에 따라 변화하는 장기간의 비조건부 분산을 잘 설명하기 때문에 표본 내 예측력 검정과 표본 외 예측력 검정에서 모두 기존의 모형들보다 더 좋은 예측력을 보인다고 할 수 있겠다.

## 4 결론

본 논문에서는 반모수 단일지표 모형을 이용하여 한국 주식시장의 수익률의 변동성을 분석하였다. 반모수 단일지표 모형은 기존의 일반적인 GARCH 형태의 모형이 정상성을 가지기 때문에 시간에 따라 변화하는 비조건부 분산을 설명하지 못하는 것을 보완하여 변동성을 GARCH 형태의 모형과 단일지표 모형의 곱한 형태로 정의한다. 이와 같은 형태로 인해 단기간에 변화하는 조건부 분산은 GARCH 형태의 모형으로 설명하고 GARCH 형태의 모형이 반영하지 못하는 시간에 따라 변화하는 장기간의 비조건부 분산을 경제/금융 변수들을 이용한 단일지표 모형으로 설명할 수 있다. 실증 분석에서는 2004년 1월 5일부터 2014년 8월 29일까지, 2643일 간의 KOSPI 수익률 데이터와 경기동행지수, VKOSPI, 국민은행 주택매매가격지수와 같은 경제/금융 변수 데이터를 사용하였다. 표본 내 예측력 검정과 표본 외 예측력 검정을 하여 기존의 여러 모형들과 비교한 결과, 반모수 단일지표 모형이 기존 모형들보다 변동성의 설명 및 예측 시에 보다 우월함을 나타냈다. 본 논문이 보이는 것처럼 비정상 시계열이 포함된 반모수 단일지표 모형은 현실적으로 유용성이 높을 것으로 기대된다. 따라서 그 추정치들의 일치성과 점근 분포에 대한 계량경제이론을 확립하는 것이 필요하며 이는 저자들이 후속 연구에서 다룰 계획이다. 또한 Amado and Teräsvirta(2014)가 보였듯이 본 모형을 다변량(multivariate) GARCH 모형으로 확장할 때에도 유용성이 높음을 보일 수 있을 것으로 기대되므로, 이러한 다변량 확장 모형에 대해서도 저자들이 후속 연구에서 다룰 계획이다.

# Appendix

## 1. Figure

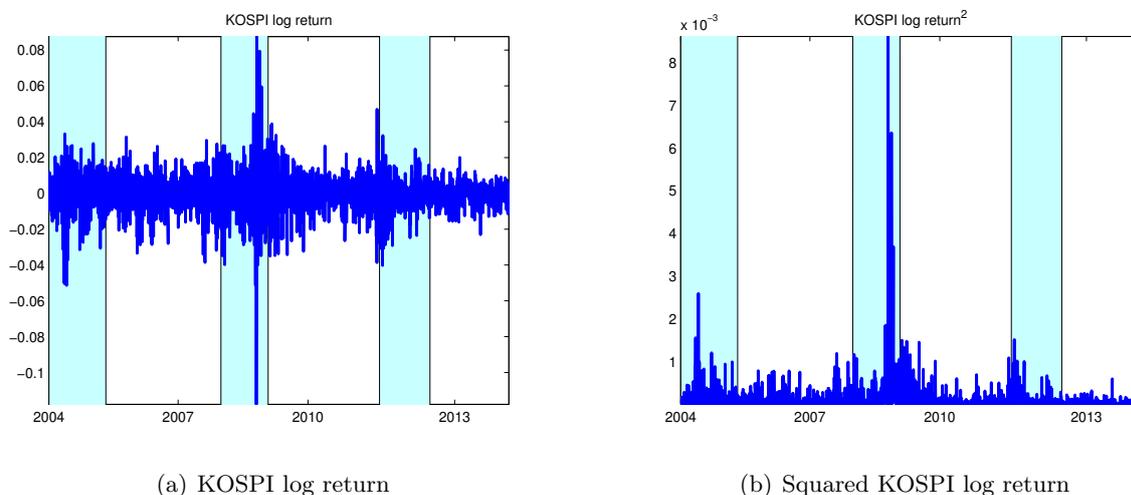


Figure 1: Daily KOSPI log return and squared return from January 5, 2004 to August 29, 2014. Shaded areas represent recession periods announced by the department of statistics in Korea.

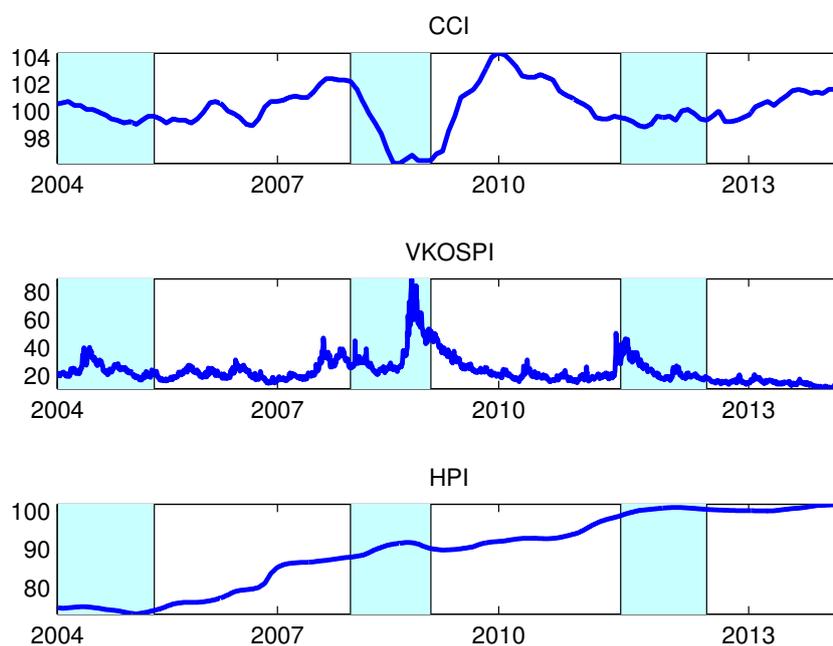


Figure 2: Coincident composite index (CCI), VKOSPI, and housing market price index (HPI) from January 5, 2004 to August 29, 2014. Shaded areas represent recession periods.

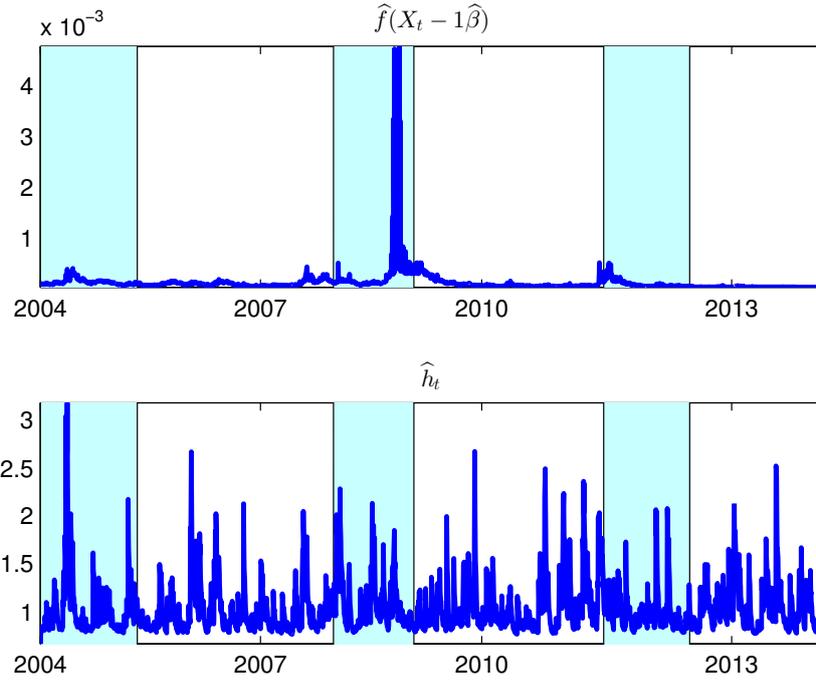
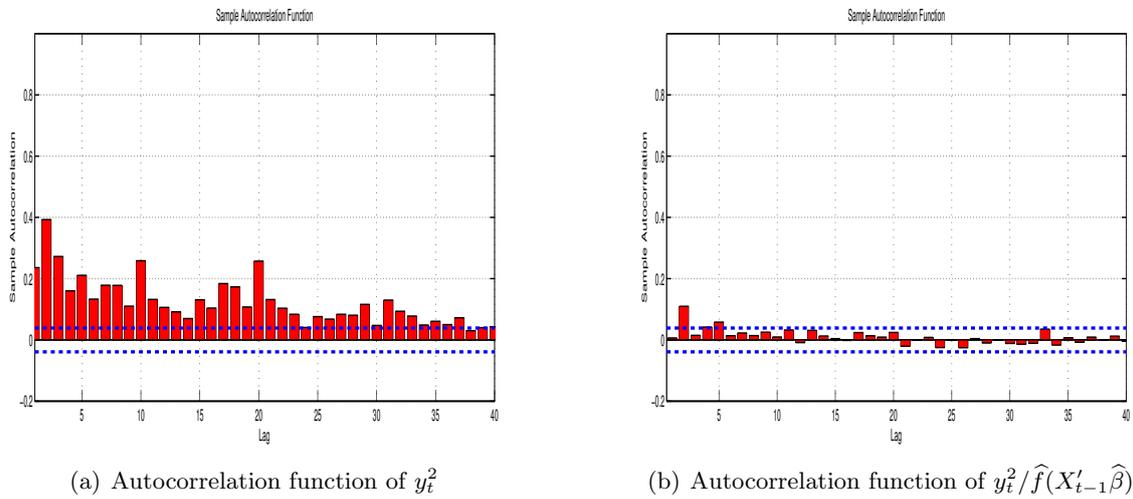


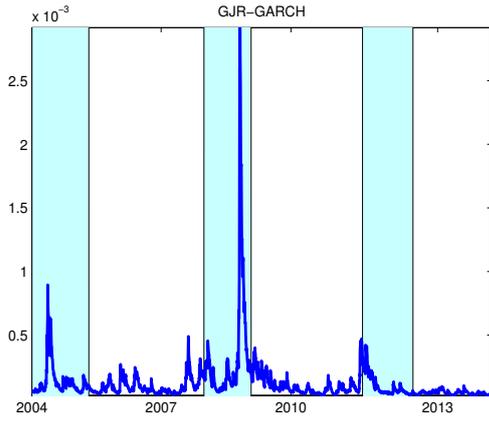
Figure 3:  $\hat{f}(X'_t - 1\hat{\beta})$  and  $\hat{h}_t$  estimated from the semiparametric single index model. Sample period is from January 5, 2004 to August 29, 2014, and shaded areas represent recession periods.



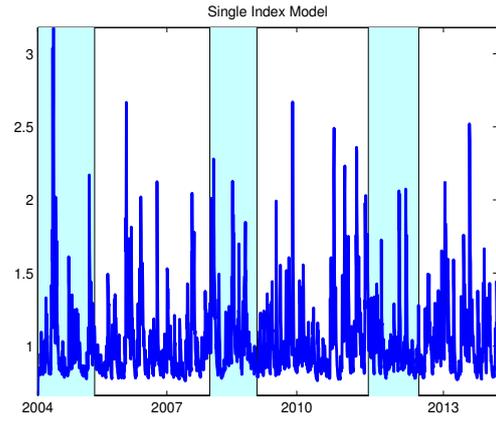
(a) Autocorrelation function of  $y_t^2$

(b) Autocorrelation function of  $y_t^2 / \hat{f}(X'_t - 1\hat{\beta})$

Figure 4: Autocorrelation functions of squared KOSPI return ( $y_t^2$ ) and  $y_t^2 / \hat{f}(X'_t - 1\hat{\beta})$ .  $\hat{f}(X'_t - 1\hat{\beta})$  is estimated from the semiparametric single index model.



(a)  $\hat{h}_t$  from GJR-GARCH model



(b)  $\hat{h}_t$  from semiparametric single index model

Figure 5:  $\hat{h}_t$  from GJR-GARCH model and semiparametric single index model from January 5, 2004 to August 29, 2014. Shaded areas represent recession periods.

## 2. Table

Table 1: Unit root test results for economic and financial variables

|                            | With intercept | With intercept and trend |
|----------------------------|----------------|--------------------------|
| coincidenc composite index |                |                          |
| (1)AR coefficient          | 1.0002         | 1.0001                   |
| (2)ADF test                | -0.3403        | -0.3405                  |
| (3)KPSS test               | 11.5167***     | 5.9928***                |
| VKOSPI                     |                |                          |
| (1)AR coefficient          | 0.9850         | 0.9837                   |
| (2)ADF test                | -4.0145**      | -4.0175***               |
| (3)KPSS test               | 32.9212***     | 17.5121***               |
| housing market price index |                |                          |
| (1)AR coefficient          | 0.9999         | 1.0004                   |
| (2)ADF test                | -1.7421        | -1.7434**                |
| (3)KPSS test               | 253.8291***    | 33.1127***               |

Notes : (1) The table gives an overview over unit root test results. The followings should be checked; 1) AR coefficient is clost to unity, 2) the null hypothesis of ADF test is not rejected, 3) the null hypothesis of KPSS test is rejected.

(2) \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 2: Parameter estimates of volatility models

|                             | GJR                                  | GARCH-X  | MIDAS                 | semiparametric single index                            |
|-----------------------------|--------------------------------------|--|-----------------------|--|
| $\alpha$                    | 0.0368***<br>(1.0892 <sup>-4</sup> ) | $7.8208 \times 10^{-6}$ ***<br>( $2.7147 \times 10^{-6}$ ) | 0.0426***<br>(0.1101) | $1.2942 \times 10^{-8}$<br>( $2.9257 \times 10^{-8}$ ) |
| $\gamma$                    | 0.0906***<br>(3.5188 <sup>-4</sup> ) | 0.1741***<br>(0.0305)                                      | 0.0797***<br>(0.0192) | 0.0993***<br>(0.0220)                                  |
| $\beta$                     | 0.8993***<br>(1.6826 <sup>-4</sup> ) | 0.8283***<br>(0.0502)                                      | 0.8990***<br>(0.0134) | 0.7953***<br>(0.0444)                                  |
| $\alpha + \gamma/2 + \beta$ | 0.9814                               | 0.9153   | 0.8876                | 0.8449   |

|         | GARCH-X  |                | MIDAS                                |
|---------|--|----------------|--------------------------------------|
| $\pi_1$ | $3.5433 \times 10^{-8}$ ***<br>( $9.3523 \times 10^{-9}$ ) | $\theta_1$     | $-2.9094 \times 10^{-7}$<br>(0.0028) |
|         |  | $\omega_{2,1}$ | 28.1103***<br>(0.1618)               |
| $\pi_2$ | $3.5000 \times 10^{-7}$ **<br>( $1.4801 \times 10^{-7}$ )  | $\theta_2$     | $-1.0847 \times 10^{-7}$<br>(0.0006) |
|         |  | $\omega_{2,2}$ | 5.0036***<br>(0.0956)                |
| $\pi_3$ | $5.0926 \times 10^{-8}$ ***<br>( $6.4716 \times 10^{-9}$ ) | $\theta_3$     | $-1.5288 \times 10^{-7}$<br>(0.0020) |
|         |  | $\omega_{2,3}$ | 6.0028<br>(8.3573)                   |

Notes : (1)  $\alpha + \gamma/2 + \beta$  indicates persistence measure. (2) The numbers in parentheses are standard errors (3) \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . (4) In GARCH-X model,  $\pi_1$  indicates coefficient of the reciprocal of coincident composite index,  $\pi_2$  refers to coefficient of VKOSPI, and  $\pi_3$  stands for coefficient of the reciprocal of housing market price index. (5) GARCH-MIDAS model is estimated under restriction of  $\omega_1 = 1$  in beta weighting function.  $\theta_1$  indicates coefficient of difference of coincident composite index in  $f_t$ ,  $\theta_2$  refers to coefficient of difference of VKOSPI, and  $\theta_3$  stands for coefficient of difference of housing market price index in  $f_t$ .  $\omega_{2,i}$  is parameter in beta weighting function.

Table 3:  $\hat{\beta}$  of single index model and semiparametric single index model

|           | single index | semiparametric single index |
|-----------|--------------|-----------------------------|
| $\beta_1$ | -0.2588      | -0.2226                     |
| $\beta_2$ | 0.9545       | 0.9574                      |
| $\beta_3$ | -0.1479      | -0.1840                     |

Notes :  $\beta_1$  indicates coefficient of coincident composite index,  $\beta_2$  refers to coefficient of VKOSPI, and  $\beta_3$  stands for coefficient of housing market price index.

Table 4: Comparison of in-sample fitting

|       | GJR       | GARCH-X   | MIDAS     | single index | semiparametric single index |
|-------|-----------|-----------|-----------|--------------|-----------------------------|
| QLIKE | 0.1951    | 0.1917    | 0.1947    | 0.1753       | <b>0.1518</b>               |
| DMW   | 5.4387*** | 6.3547*** | 5.3951*** | 3.9919***    |                             |

Notes : (1) The lowest QLIKE loss is presented in bold. (2) \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 5: Comparison of out-of-sample forecasting

|       | GJR       | GARCH-X   | MIDAS     | single index | semiparametric single index |
|-------|-----------|-----------|-----------|--------------|-----------------------------|
| QLIKE | 0.2272    | 0.2128    | 0.3643    | 0.2062       | <b>0.1764</b>               |
| DMW   | 5.8313*** | 5.7440*** | 7.7809*** | 3.6760***    |                             |

Notes : (1) The lowest QLIKE loss is presented in bold. (2) \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

## Reference

김세완. “경기변동을 고려한 주식수익률과 변동성 관계의 변화: 비대칭 GARCH 모형을 이용하여.” 한국금융학회 23.2 (2009): 1-28.

(Translated in English) Kim, S. “Stock returns and its volatility under business cycles changes : Using asymmetric GARCH Model.” Journal of Money and Finance 23.2 (2009): 1-28.

김영일. “한국 주식시장의 지속적 변동성과 거시경제적 관련성 분석.” 한국개발연구 35.4 (2013): 63-94.

(Translated in English) Kim, Y. “The long-lived volatility of Korean stock market and its relation to macroeconomic conditions.” KDI Journal of Economic Policy 35.4 (2013): 63-94.

오현탁 · 이현상 · 이치송. “한국주식시장의 시장상황별 비대칭적 변동성에 관한 실증연구.” 재무관리연구 17.1 (2000): 45-65.

(Translated in English) Oh, H. and H. Lee, C. Lee “An empirical study on asymmetric stock market volatility in Korea.” The Korean Journal of Financial Management 17.1 (2000): 45-65.

임병진 · 한성윤. “주식시장 지수와 부동산시장 지수의 시계열 특성비교와 관계에 관한 실증적 연구.” 산업경제연구 22.4 (2009): 2065-2083.

(Translated in English) Lim, B. and S. Han “A study on the relationship between return on real estate and Korea composite stock price index.” Korea Industrial Economics Association 22.4 (2009): 2065-2083.

Amado, C. and T. Teräsvirta “Modelling volatility by variance decomposition.” Journal of Econometrics 175.2 (2013): 142-153.

Amado, C. and T. Teräsvirta “Conditional correlation models of autoregressive conditional heteroscedasticity with nonstationary garch equations.” Journal of Business & Economic Statistics 32.1 (2014): 69-87.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde and N. Shephard “Designing realized kernels to measure the ex post variation of equity prices in the presence of noise.” Econometrica (2008): 1481-1536.

Conrad, C. and K. Loch “Anticipating long-term stock market volatility.” Journal of Applied Econometrics (2014).

Corsi, F. "A Simple Approximate Long-Memory Model of Realized Volatility." *Journal of Financial Econometrics* 7.2 (2009): 174-196.

Diebold, F.X. and R.S. Mariano "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics* 13.3 (1995): 253-63.

Engle, R.F. "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation." *Econometrica: Journal of the Econometric Society* (1982): 987-1007.

Engle, R.F., E. Ghysels and B. Sohn "Stock market volatility and macroeconomic fundamentals." *Review of Economics and Statistics* 95.3 (2013): 776-797.

Engle, R.F. and J.G. Rangel "The spline-GARCH model for low-frequency volatility and its global macroeconomic causes." *Review of Financial Studies* 21.3 (2008): 1187-1222.

Feng, Y. and L. Sun "A semi-APARCH approach for comparing long-term and short-term risk in Chinese financial market and in mature financial markets." No. 69. University of Paderborn, CIE Center for International Economics, 2013.

Hafner, C.M. and O. Linton "Efficient estimation of a multivariate multiplicative volatility model." *Journal of Econometrics* 159.1 (2010): 55-73.

Hamilton, J.D. and G. Lin "Stock market volatility and the business cycle." *Journal of Applied Econometrics* 11.5 (1996): 573-593.

Han, H. and D. Kristensen "Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates." *Journal of Business & Economic Statistics* 32.3 (2014): 416-429.

Han, H. and D. Kristensen "Semiparametric Multiplicative GARCH-X Model: Adopting Economic Variables to Explain Volatility." Working paper (2015).

Ichimura, H. "Semiparametric least squares (SLS) and weighted SLS estimation of single-index models." *Journal of Econometrics* 58.1 (1993): 71-120.

Koo, B. and O. Linton "Let's get LADE: Robust estimation of semiparametric multiplicative volatility models." *Econometric Theory* (2013): 1-32.

Patton, A.J. and K. Sheppard “Evaluating volatility and correlation forecasts.” Handbook of financial time series. Springer Berlin Heidelberg, 2009. 801-838.

Schwert, G.W. “Why does stock market volatility change over time?.” The journal of finance 44.5 (1989): 1115-1153.

West, K.D. “Asymptotic inference about predictive ability.” Econometrica 64.5 (1996): 1067-1084.

# Adopting economic and financial variables to explain stock market volatility in Korea \*

Seunghee Lee <sup>†</sup>

Heejoon Han<sup>‡</sup>

## Abstract

This paper studies stock market volatility in Korea using a semiparametric single index volatility model, in which a single index long run component induced by exogenous covariates is multiplied to a GARCH short run component. When a covariate is nonstationary, i.e. integrated or near-integrated, the model can account for time-varying unconditional variance of financial time series. Among various economic and financial indicators, it is found that the coincident composite index, VKOSPI, and housing price index are helpful in fitting and forecasting stock market volatility in Korea. It is shown that the model using these three variables outperforms standard models both in terms of in-sample fitting and out-of-sample forecasting.

This version: Dec 4, 2015

*JEL Classification:* C22, C50, G12

*Key words and phrases:* volatility, semiparametric single index model, GARCH-X, nonstationarity

---

\*We would like to thank Chang Sik Kim(Sungkyunkwan University), Joon Y. Park(Indiana University & Sungkyunkwan University) and participants of 2015 Korean Econometric Society Summer Meeting, 90th Annual conference of the West Economic Association International(Honolulu), 2015 Workshop of Sungkyun Institute of Economic Research for helpful comments and suggestions.

<sup>†</sup>M.A. Student, Department of Economics, Sungkyunkwan University, 25-2, Sungkyunkwan-ro, Jongno-gu, Seoul 03063, Korea, Phone: +82-2-760-1236, e-mail: prious@skku.edu

<sup>‡</sup>Corresponding Author, Associate Professor, Department of Economics, Sungkyunkwan University, 25-2, Sungkyunkwan-ro, Jongno-gu, Seoul 03063, Korea, Phone: +82-2-760-0428, e-mail: heejoonhan@skku.edu