

An Endogenous Regime Switching Model for Realized Volatility *

Sejung Kim [†] Heejoon Han[‡]

Abstract This paper introduces and analyzes a new model for realized volatility that accommodates endogenous regime switching. The model is based on the heterogeneous autoregressive model and allows for two-state regime switching. Importantly, a current shock to the realized volatility affects the regime switching in the next period. We apply the model to the realized volatility of the daily S&P 500 Index return series. The estimation result shows that the short-term volatility component is the most influential in the high volatility regime while the long-term volatility component is dominant in the low volatility regime. Moreover, the model significantly outperforms existing models in within-sample fitting.

Keywords realized volatility, endogenous regime switching, heterogeneous autoregressive model

JEL Classification C22, C50, G12

*We would like to thank Chang Sik Kim (Sungkyunkwan University) and Joon Y. Park (Indiana University and Sungkyunkwan University) for helpful comments and suggestions.

[†]Researcher, Korea Institute of Finance, 19, Myeong-Dong 11 Gil, Jung-gu, Seoul 04538, Korea, Phone: +82-2-3705-6265, e-mail: kimsj@kif.re.kr

[‡]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

내생적 국면 전환을 이용한 실현 변동성 모형 *

김세정 †

한희준 ‡

Abstract

본 논문은 금융 시계열의 실현 변동성을 설명하기 위해 내생적 국면 전환을 도입한 새로운 모형을 제안하고 분석하고 있다. 모형은 이질적 자기회귀(heterogenous autoregressive) 모형을 기반으로 한 두 국면(two-state) 전환 모형이다. 모형에서 가장 중요한 특징은 현재 기의 실현 변동성에 대한 충격이 다음 기 국면 전환에 영향을 미치는 내생적 국면 전환 구조를 도입한 점이다. 본 모형을 일별 S&P 500 지수 수익률의 실현 변동성에 적용한 결과, 변동성이 높은 국면에서는 단기 변동성 요인이 다음 기 변동성에 지배적인 영향을 미치는 반면 변동성이 낮은 국면에서는 장기 변동성 요인이 지배적인 영향을 미치는 것으로 나타났다. 또한 내생적 국면 전환을 도입함에 따라 기존 모형들에 비해 표본기간 내 설명력이 유의하게 개선되는 것으로 나타났다.

Keywords realized volatility, endogenous regime switching, heterogeneous autoregressive model

JEL Classification C22, C50, G12

*본 논문에 대해 유익한 논평과 조언을 해주신 김창식 교수님(성균관대학교)과 박준용 교수님(Indiana University 및 성균관대학교)께 감사드립니다.

†금융연구원 연구원, e-mail: kimsj@kif.re.kr

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

1. 서론

금융 시장의 변동성(volatility)은 리스크(risk)에 대한 측정 지표로 사용될 수 있어 금융 기관들의 자산 가격 결정(asset pricing), 리스크 관리(risk management), 자산 배분(portfolio allocation) 등의 의사 결정에 중요한 영향을 끼치는 것으로 알려져 있다. 따라서 변동성을 설명/예측하기 위한 시계열 모형들이 활발하게 연구되어 왔으며, 최근 2008년 세계 금융위기 이후 금융시장의 급격한 변화에 따른 세계 경제의 침체를 겪으며 관련 연구의 중요성이 더해지고 있다. 주가 수익률이나 환율 변동률 등 금융 시계열의 시변(time-varying) 변동성은 관측이 불가능하기 때문에 일반적으로 조건부 분산(conditional variance)을 이용한 GARCH 형태의 모형 또는 확률 변동성(stochastic volatility) 모형 등의 구조형(structural form) 모형을 이용하여 추정 및 예측하여 왔다.

최근 2000년대 들어서는 고빈도 데이터(high-frequency data)의 활용이 가능해지면서 이를 이용하여 관측되지 않는 실제 변동성을 추정하는 연구들이 활발해졌다. 예를 들어 Andersen et al. (2003)과 Barndorff-Nielsen and Shephard (2002) 등은 실제 변동성에 대한 일치 추정량을 제공하는 다양한 실현 변동성(realized volatility) 측정치들을 제시하였다. 이와 함께 실현 변동성 데이터를 직접 이용함으로써 GARCH 형태의 구조형 모형과 차별화되는 축약형(reduced form) 변동성 모형들이 발전되고 있는데, Corsi (2009)의 이질적 자기회귀 모형(heterogeneous autoregressive model, HAR 모형)을 그 예로 들 수 있다.

변동성 관련 문헌에서는 Ding et al. (1993) 등에 의해 금융 시계열의 변동성이 장기 기억(long memory) 특성을 가지는 지속적(persistent)인 시계열임이 알려져 왔고, Banerjee and Urga (2005)이 기술하였듯이 이러한 변동성의 높은 지속성을 설명하려는 연구가 활발히 진행되어 왔다. Andersen et al. (2003)이 실현 변동성의 모델링에 ARFIMA(Autoregressive Fractionally Integrated Moving Average) 모형을 이용한 이유도 ARFIMA 모형이 장기 기억 특성을 설명할 수 있기 때문이었다. 그러나 ARFIMA 모형은 추정이 비교적 복잡하고 경제학적 해석이 명료하지 않은 단점이 있는데, 이에 대한 대안으로 제시된 HAR 모형은 Müller et al. (1997)의 이질적 시장 가설(heterogeneous market hypothesis)에 근거하여 투자자 특성별 해석을 가능하게 하고 또한 단순한 선형 구조를 가지므로 최소자승법을 이용하여 쉽게 추정할 수 있는 특성이 있다.

HAR 모형은 추정이 쉬운 장점에다 장기 기억 특성 또한 설명할 수 있는 것으로 알려져 있기 때문에 실현 변동성의 모델링과 예측에 널리 이용되어 왔는데, 최근 HAR 모형을 더욱 발전시키는 연구들이 활발히 진행되어 왔다.

이는 Corsi et al. (2008) 등이 밝힌 바와 같이 HAR 모형의 잔차항(residual)에 자기상관(autocorrelation) 또는 변동성 군집(volatility clustering) 현상이 관측되어 단순한 HAR 모형만으로는 실현 변동성 시계열을 충분히 설명하는 데 부족함이 있다는 것이 알려졌기 때문이다. 이러한 이유로 최근 Andersen et al. (2007), Liu and Maheu (2008), McAleer and Medeiros (2008), Pypko (2015), Wang et al. (2016) 등은 HAR 모형에 비선형성, 점프(jump), 구조 변화(structural break), 순조로운 이행 함수(smooth transition function)를 이용한 국면 전환(regime switching) 등을 결합한 모형들을 발전시켜 왔다.

이러한 맥락에서 본 연구는 새로운 실현 변동성 모형을 제안하고 분석하는데, 이는 구체적으로 최근 Chang et al. (2016)에 의해 소개된 내생적 국면 전환 모형(endogenous regime switching model)을 HAR 모형에 결합한 것으로 내생적 국면 전환에 따라 HAR 모형의 계수들이 국면별로 다른 값을 가지게 한다. 기존의 마코프 전환 모형(Markov regime switching model)에서는 상태 변수(state variable)가 기저 시계열(underlying time series)과 독립적으로 결정되는 한계가 있는 반면, Chang et al. (2016)의 내생적 국면 전환 모형은 이를 보완하여 다음 기의 상태 변수가 현재 기의 기저 시계열에 의해 영향을 받도록 하는 특징이 있다. 본 연구의 목적은 내생적으로 두 상태(two-state)의 국면 전환이 이루어지는 HAR 모형을 분석하여 이러한 모형의 확장이 실현 변동성 모형의 설명력을 개선할 수 있는지 살펴보는 것이다.

본 논문은 실증분석에서 일별 S&P 500 지수 수익률의 실현 변동성을 이용하며, 2006년 1월 3일부터 2015년 12월 31일까지의 전체 표본기간과 금융위기 기간을 제외한 2010년 1월 4일부터 2015년 12월 31일까지의 하위 표본기간을 고려하고 있다. 모형의 추정 결과 중 가장 큰 특징은 변동성이 높고 낮은 각 국면별로 다음 기 변동성에 가장 크게 영향을 미치는 요인들이 상이하게 드러났다는 점이다. 즉, 변동성이 높은 국면에서는 일(day) 단위의 단기 변동성 요인이 다음 기 변동성에 지배적인 영향을 미치고, 변동성이 낮은 국면에서는 주(week) 또는 월(month) 단위의 중기와 장기 변동성 요인들이 지배적인 영향을 미치는 것으로 나타났다. 이처럼 각 국면별로 단기와 장기 변동성 요인들이 대조적으로 추정되는 결과는 기존 연구들에서 주목하지 못한 부분으로, 본 논문에서는 단기 투자자와 장기 투자자의 행태별 특징을 고려하여 이에 대한 해석을 제공하고 있다.

또한 본 논문은 내생적 국면 전환을 이용한 HAR 모형이 기존의 모형들에 비해 유의하게 우월한 표본기간 내 설명력을 지니고 있음을 보이고 있다. 기존 모형들과 비교하여 모형 적합도를 평가할 때, 손실함수(loss function)로는 변동성 측정 오차에 강건(robust)한 손실함수들로 알려진 MSE(Mean Squared

Error) 손실함수와 QLIKE 손실함수를 이용하였고 Diebold-Mariano(1995)와 West(1996)의 검정을 시행하여 손실함수 간 차이의 유의성을 검정하였다. 특히 국면 전환의 내생성을 허용하지 않을 경우에 비해 이를 허용할 때 표본기간 내 설명력이 유의하게 개선됨을 보이고 있는데, 이는 내생적 국면 전환을 실현 변동성 모형에 사용하는 본 모형의 유용성을 입증하는 것이다.

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

2. 모형 및 추정 방법

2.1. 내생적 국면 전환 HAR 모형

본 연구에서 제안하고 분석하는 모형은 HAR 모형에 새로운 국면 전환 구조를 결합하여 자기회귀 계수들이 상태 변수에 따라 전환하도록 한 내생적 국면 전환 HAR 모형(Endogenous Regime Switching HAR Model, 이하 ERS-HAR 모형)이다. 본 연구에서는 이 모형을 통해 보다 유연하게 변동성을 추정하고, 변동성 수준에 따른 주식시장 상황과 거래자들의 특성을 살펴보고자 한다. 우선 ERS-HAR 모형을 소개한 후, 모형의 두 가지 구성 요소인 HAR 모형과 내생적 국면 전환 모형에 대해 각각 설명하도록 하겠다. 본 논문 전체에서 $RV_t^{(d)}$ 은 주가 수익률이나 환율 변화율 등 금융 시계열의 일별(daily) 실현 변동성을 나타낸다. 또한 $RV_t^{(w)}$ 과 $RV_t^{(m)}$ 은 각각 주별(weekly)과 월별(monthly) 실현 변동성을 나타내며, 다음과 같이 일별 실현 변동성의 단순 평균으로 계산된다.

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

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

실현 변동성 추정을 위한 ERS-HAR 모형의 기본적인 구조는 다음과 같다.

$$\log RV_t^{(d)} = c(s_t) + \beta_d(s_t) \log RV_{t-1}^{(d)} + \beta_w(s_t) \log RV_{t-1}^{(w)} + \beta_m(s_t) \log RV_{t-1}^{(m)} + \sigma u_t, \quad (1)$$

$$\begin{pmatrix} u_t \\ v_{t+1} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right), \quad (2)$$

$$s_t = I(w_t \geq \tau), \quad (3)$$

$$w_{t+1} = \alpha w_t + v_{t+1}, \quad (4)$$

$$\alpha \in [-1, 1], v_t \sim iid(0, 1).$$

ERS-HAR 모형은 식 (1)에서 볼 수 있듯이 서로 다른 주기의 실현 변동성들을 이용한 선형 회귀 모형인 HAR모형을 기반으로 하고 있다. ERS-HAR 모형은 HAR 모형에 국면 전환 구조를 이용하여 계수들 $(c(s_t), \beta_d(s_t), \beta_w(s_t), \beta_m(s_t))'$ 이 각각 상태 변수 s_t 가 0 혹은 1의 값을 가짐에 따라 두 국면(two-state)에서 다른 값들을 가지는 형태이다. 이 때 상태 변수 s_t 는 식 (3)과 같이 잠재 변수(latent variable) w_t 의 값에 따라 결정된다. 두 국면 (two-state) 모형을 가정하여 상태 변수가 0의 값을 가질 때 변동성이 낮은 국면 그리고 1의 값을 가질 때 변동이 높은 국면으로 구체화하고 있다. 상태 변수는 식 (4)와 같이 1차 자기회귀(AR(1)) 과정을 따르는 잠재 변수 w_t 와 역치 수준(threshold) τ 에 의해 결정되는 지시 함수(indicator function)로 정의된다. 이 때 ERS-HAR 모형에서 $t+1$ 시점의 잠재 변수의 오차항 v_{t+1} 과 t 시점의 관측 시계열의 오차항 u_t 이 상관관계(correlation)를 가지도록 하는 것이 특징인데, 이는 다음 기 국면을 현재 기의 기저 시계열(underlying time series)의 함수로 정의하여 내생적으로 결정되도록 한 것이다. 즉, 현재의 실현 변동성에 대한 충격이 미래의 상태 변수 결정에 영향을 미쳐 국면이 보다 현실적으로 결정되도록 한 것이다. 다음 절에서 HAR 모형과 내생적 국면 전환 모형에 대해 차례로 설명하고, 본 연구에서 제시한 모형의 구성과 추정 방법에 대하여 논의하도록 하겠다.

2.1.1 HAR 모형

주가 수익률이 연속 확산 과정(continuous diffusion process)를 따를 때, 자산의 로그 가격 $p(t)$ 은 다음과 같이 나타낼 수 있다.

$$dp(t) = \mu(t)dt + \sigma(t)dW(t).$$

이 경우 $\sigma(t)$ 는 순간 변동성(instantaneous volatility)이며 t 일(day)의 적분 변동성(integrated volatility)은 식 (5)와 같이 정의된다.

$$IV_t^d = \int_{t-1}^t \sigma^2(\omega) d\omega. \quad (5)$$

Andersen et al. (2003)은 표본 추출 빈도(sampling frequency)가 높아지면 일중(intraday) 수익률의 제곱의 합으로 정의되는 실현 분산(realized variance)이 적분 변동성에 대한 일치추정량이 됨을 보였다. 그러나 시장 미시구조 잡음(market microstructure noise)이 존재할 경우에는 이러한 실현 변동성 추정치에 편의가 발생하게 되는데, Barndorff-Nielsen et al. (2008)은 아래와 같이 정의되는 실현 커널(realized kernel)이 시장 미시구조 잡음에 대하여 강건(robust)한 것을 보였다.

$$RK_t = \sum_{h=-H}^H K\left(\frac{h}{H+1}\right) \gamma_h, \quad \gamma_h = \sum_{j=|h|+1}^n r_{j,t} r_{j-|h|,t}.$$

여기서 $K(\cdot)$ 는 Parzen 커널 함수이며 $r_{j,t}$ 는 t 일의 j 번째 고빈도 수익률이다.¹

이러한 실현 분산, 실현 커널 등의 다양한 실현 변동성 추정치들은 지속성(persistence)이 강하게 나타나는 것으로 알려져 있는데, HAR 모형은 주식시장 거래자들의 이질성에 주목하여 이러한 특성을 설명한다. 주식시장의 주체들이 서로 다른 거래 주기를 가지고 있기 때문에 각각 다른 변동성 요소를 구성한다는 소위 이질적 시장 가설(Mieller et al., 1997)에 근거하여 거래 주체들을 단기, 중기와 장기 행위자로 규정한다. 이와 같은 기간 분류는 간단하게 하루, 일주일, 한 달로 구체화될 수 있다.² HAR 모형은 아래와 같이 세 가지 단위 실현 변동성의 이전 기($t-1$ 기) 값들이 더해진 자기회귀 형태로 실현 변동성을 모델링한다.

$$RV_t^{(d)} = c + \beta^{(d)} RV_{t-1}^{(d)} + \beta^{(w)} RV_{t-1}^{(w)} + \beta^{(m)} RV_{t-1}^{(m)} + \varepsilon_t.$$

¹ Bandwidth H 의 선택 등 자세한 사항들을 위해서는 Barndorff-Nielsen et al. (2009)과 Heber et al. (2009)를 참조하기 바란다.

² HAR 모형의 예측력 성과가 주목받으며 모형의 구조에 대한 논의도 최근 들어 증가하고 있다. Craioveanu and Hillebrand (2012)는 주 단위, 월 단위 평균 변동성을 중기와 장기 변동성 요소로 모형화하는 것이 이론적으로 적합한지에 대한 연구를 진행하였고, Hwang and Shin (2014)은 다른 시차 구조를 고려한 HAR 모형들이 표본 내 적합성과 표본 외 예측력을 주목할 만큼 변화시키지 못한다는 결론을 도출하였다.

동 모형은 단순 선형함수 형태로 일반적인 최소자승법을 적용할 수 있기 때문에 추정이 간편한데 반해 변동성의 장기 기억 특성을 설명할 수 있고 예측력이 뛰어난 것으로 알려져 있다. 그런데, Deo et al. (2006)과 Martens et al. (2004) 등에 의해 각 실현 변동성 변수에 로그를 취한 형태를 사용하는 것이 모델링 관점에서 유용하다고 알려져 있다. 그 이유는 첫째, 실현 변동성에 로그변환을 취했을 때 정규분포에 근사하게 되어 표준적인 시계열 모델링에 더 적합하다는 점이고, 둘째, 로그 형태의 변수를 고려하였을 경우 추정된 변동성 값이 양의 값(positive value)을 가져야 한다는 제약을 추가적으로 가할 필요가 없다는 점이다. 따라서 본 연구에서도 다음과 같이 로그 변환을 취한 HAR 모형 형태를 식 (1)과 같이 모형설계에 차용하였다.

$$\log RV_t^{(d)} = c + \beta^{(d)} \log RV_{t-1}^{(d)} + \beta^{(w)} \log RV_{t-1}^{(w)} + \beta^{(m)} \log RV_{t-1}^{(m)} + u_t.$$

2.1.2 내생적 국면 전환 모형

관측되지 않는 상태 변수가 마코프 체인(Markov chain)을 따르는 국면 전환 모형은 Goldfeld and Quandt (1973)에 의해 소개된 이후, Hamilton (1989)에 의해 시계열 모형에 1차 마코프 과정(first order Markov process)을 따르는 상태변수를 결합한 마코프 전환 모형(Markov Regime Switching model, MRS)이 발전되었고, Kim (1994)이 상태 공간 모형(State-Space model)의 개념을 이용하여 모형을 더욱 일반화하였다. Hamilton and Susmel (1994), Hamilton and Gang (1996), Gray (1996) 등에 의하여 국면 전환 구조가 변동성과 관련한 연구에도 활발하게 적용되었다.

일반적으로 많이 사용되는 두 상태 ($s_t = 0, 1$) 국면 전환모형은 경제의 호황과 불황, 저변동기와 고변동기 등 상반되는 경제 상황을 상태 변수를 이용해 모형에 반영하게 되는데, 이러한 두 상태 국면 전환 모형의 전이확률(transition probability)은 $p = \mathbb{P}(s_t = 0 | s_{t-1} = 0)$, $q = \mathbb{P}(s_t = 1 | s_{t-1} = 1)$ 와 같이 표현되고, 이 모형의 전이 행렬(transition matrix)은 잘 알려진 바와 같이 다음과 같다.

$$P = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}.$$

여기에서 p 는 경제 상황이 저국면인 국면 0에 머무를 확률을 나타내고 q 는 고국면인 국면 1에 머무를 확률을 나타낸다. 전통적인 마코프 국면 전환

모형³은 상태 변수 s_t 가 마코프 체인(Markov chain)을 따름으로써 오로지 전기의 상태 변수 s_{t-1} 에 의해서만 영향을 받고, 기저 시계열과는 독립적으로 결정된다고 가정한다. 동질적(homogeneous)이고 정상적(stationary)인 마코프 체인을 가정하면 상태 변수 (s_t)가 위의 전이행렬로 완전히 설명될 수 있고, 따라서 관측 불가능한 상태 변수를 p 와 q 두 모수로 명료하게 모형에 반영할 수 있다.

그러나 경제 변수의 구조적인 움직임을 반영하기 위한 국면 전환이 기저 시계열과 독립적인 상태 변수에 의해 외생적(exogenous)으로 결정되는 것은 굉장히 제한적인 가정이라고 할 수 있다. 이러한 한계점을 극복하기 위한 시도로 기저 시계열이 상태 변수에 영향을 미치는 ‘내생적(endogenous)’인 국면 전환 모형들이 제안되었다. Kim et al. (2008)에서는 독립동일분포(i.i.d.)를 따르는 잠재 변수의 t 기 현재의 오차항과 관측 시계열의 현재 기 오차항 간의 동시발생적인(contemporaneous) 상관 관계를 허용하였다. 최근 소개된 Chang et al. (2016)의 국면 전환 모형은 Kim et al. (2008)의 모형을 더욱 발전시킨 것으로, 상태 변수 s_t 를 식 (3)과 식 (4)와 같이 1차 자기회귀 과정을 따르는 잠재 변수가 역치 수준보다 클 때 1의 값을 가지도록 정의하였다. 이 때 잠재 변수의 $t + 1$ 기의 오차항과 관측 시계열의 현재 기의 오차항 간의 상관관계를 ρ 라는 모수로 두고 상태 변수가 전기의 충격에 종속적으로 결정될 수 있도록 모델링함으로써, 현재의 충격이 미래의 상태 변수 결정에 관여하여 국면이 보다 현실적으로 결정되도록 하였다. Kim et al. (2008)의 모형과 Chang et al. (2016)의 모형의 자세한 차이점에 대해서는 Chang et al. (2016)의 2.1절을 참조하기 바란다.

Chang et al. (2016)의 내생적 국면 전환 모형에서 AR(1)을 따르는 잠재 변수 (w_t)의 자기회귀 계수 α 는 모형의 국면이 얼마나 지속적으로 결정되는가를 추론할 수 있게 한다. α 값이 클수록 국면이 보다 오래 지속됨을 의미하며, 이 모형에서는 $\alpha = 1$ 인 단위근(unit root)을 가지는 비정상적(nonstationary)인 경우까지 고려할 수 있게 하였다. Chang et al. (2016)의 모형의 경우 모형 추정과 함께 잠재 변수를 추출하는 것이 가능하므로 관측 불가능한 상태변수의 움직임을 추론할 수 있는 장점이 있다. Chang et al. (2016)의 내생적 국면 전환 모형은 잠재 변수가 정상적(stationary)이고 ($|\alpha| < 1$) 내생성을 나타내는 모수 ρ 가 0인 경우에 상태 변수가 외생적으로 결정되는 기존의 마코프 전환 모형과 관측적으로 동일해지기 때문에(observationally equivalent) 전통적인 국면 전환 모형을 포괄하는 좀더 유연한 형태의 국면 전환 모형이라고 할 수 있다.

³여기서 말하는 전통적인 마코프 전환 모형은 베이지안(Bayesian) 등의 추정법을 제외한 최우추정을 이용한 고전적인 마코프 전환 모형을 일컫는다.

다음 절에서 내생적 국면 전환 모형의 추정 방법을 서술하도록 하겠다.

2.2. 추정 방법

전통적인 마코프 전환 모형의 추정은 칼만 필터(Kalman filter)와 같은 순차적 정보처리 알고리즘에 기반한 마코프 전환 필터를 이용한 최우추정 방법을 사용한다. 최근 Bauwens et al. (2010) 등에 의해 베이저안 추정(Bayesian estimation) 방법을 이용한 국면 전환 모형의 추정도 이루어지고 있으나, 본 논문에서는 Chang et al. (2016)을 따라 최우추정법을 이용하여 ERS-HAR 모형을 추정하도록 하겠다. 본 모형의 추정은 기본적으로 식 (2) 와 같이 $(u_t, v_{t+1})'$ 이 정규성(normality)을 만족시킨다는 가정 하에 최우추정을 진행하되, 관측되지 않는 상태 변수의 움직임을 반영하기 위해 수정된 마코프 전환 필터(modified Markov switching filter) (Chang et al., 2016)를 이용하여 로그 우도함수(log likelihood function)를 도출한다. Chang et al. (2016)의 내생적 국면 전환 모형에서는 특수한 경우인 $\rho = 0$ 인 경우를 제외하고는 s_t 가 마코프 체인을 따르지 않아 전통적인 마코프 전환 필터(Markov switching filter)를 사용한 추정이 불가능하기 때문에 수정된 마코프 전환 필터를 이용한다. 기존의 마코프 전환 모형에서 사용되던 마코프 전환 필터와 마찬가지로 예측(prediction)과 갱신(updating)의 두 단계를 반복하며 로그 우도함수를 최대화하는 수치해를 계산해낼 수 있다. 모형의 추정에 관한 보다 자세한 사항은 Chang et al. (2016)를 참고하기 바란다.

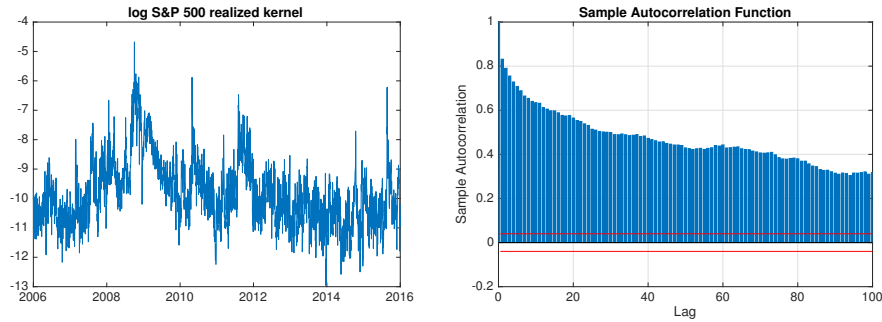
3. 실증 분석

3.1. 데이터 및 추정 모형

본 연구에서 사용한 데이터는 Oxford Man Institute Realized library에서 제공하는 일별 S&P 500 실현 커널(realized kernel)이다. Realized library에서 제공하는 실현 변동성 추정치들 중 실현 커널은 2장에서 언급했듯이 시장 미시구조 잡음에 대하여 강건하도록 설계된 추정량이다. 표본(sample)은 2006년 1월 3일 - 2015년 12월 31일의 전체 표본기간(2,510일)과 금융위기 기간을 제외한 2010년 1월 4일 - 2015년 12월 31일의 하위 표본기간(1,510일)을 고려하였다. 그림 (1)은 전체 표본기간의 로그를 취한 실현 커널과 이의 자기상관함수(autocorrelation function)를 나타낸 것이다. 그림 (1)의 우측 패널을 살펴보면 본 논문에서 고려하는 실현 변동성도 관련 연구들에서 알려진 바와 같이

시간 간격이 멀어져도 자기상관이 유의하게 남아있는 장기 기억 특성을 지님을 확인할 수 있다.

1: 실현 변동성 시계열과 자기상관함수



Notes: 왼쪽의 그림은 추정에 사용한 일별 S&P 500 지수 수익률의 실현 커널(realized kernel)에 로그를 취한 시계열을 나타낸다. 표본기간은 2006년 1월 3일 - 2015년 12월 31일의 전체 표본(2,510 obs)과 변동성이 비정상적으로 상승하였던 금융 위기 기간을 제외한 2010년 1월 4일 - 2015년 12월 31일의 하위 표본(1,510 obs)을 고려하였다. 오른쪽 그림은 로그를 취한 실현 커널의 자기상관함수이다.

2장에서 언급하였듯 로그 변환을 취하면 실현 변동성 시계열이 지니는 큰 왜도와 첨도를 보정할 수 있는데, 표 (1)에 나타난 실현 커널의 기초통계량을 통해 이를 확인해볼 수 있다. 로그변환을 취하기 전 초과첨도(excess kurtosis)가 전체 표본과 하위표본기간에서 각각 308과 161에 달했던 것에 비해 변환 후 두 표본기간 모두에서 3에 가깝게 줄어들었으며 왜도는 변환 후 0에 근접하였다. 즉, 실현 변동성에 로그를 취할 경우 보다 대칭적이고 첨도가 작은 분포를 띠어 정규분포에 근사한 형태가 됨을 알 수 있다. 각 표본기간의 평균 변동성 수준을 살펴보면, 하위 표본기간의 평균 변동성 수준이 전체 표본기간에 비해 좀더 낮게 나타나고 있는데 이는 변동성이 높은 금융위기 기간이 제외되었기 때문이다.

추정 모형으로는 일반적인 HAR 모형과 상태 변수의 내생성을 고려한 ERS-HAR 모형 그리고 ERS-HAR 모형에서 내생성을 제거한 모형($\rho = 0$) 등 세 모형들을 고려하였다. 본 장에서는 모형들의 추정 결과를 분석하고 모형들간 적합성 평가를 실시하도록 하겠다. ERS-HAR 모형에서 $\rho = 0$ 라는 제약을 주어 내생성을 제거한 모형(이하 외생적 RS-HAR 모형)은 2.1.2절에서 언급하였듯

1: 기초 통계량

	전체 표본		하위 표본	
	RV_t	$\log RV_t$	RV_t	$\log RV_t$
Mean	0.0001	-9.8391	0.0001	-10.1067
Median	0.0000	-9.9718	0.0000	-10.2059
Minimum	0.0000	-12.9611	0.0000	-12.9611
Maximum	0.0093	-4.6764	0.0028	-5.8804
Std.dev	0.0003	1.1132	0.0001	0.9451
Skewness	13.2540	0.7097	10.3131	0.4873
Kurtosis	308.871	3.7352	161.551	3.6422
N	2,510		1,510	

Notes: 전체 표본기간과 하위 표본기간의 실현 커널(realized kernel) 데이터에 대한 기초 통계량을 나타낸 표이다. 각 표본의 좌측 열은 실현 커널, 우측 열은 로그 변환을 취한 실현 커널에 대한 정보를 나타낸다.

이 전통적인 국면 전환 모형으로 분류되는 마코프 전환을 사용한 HAR 모형 (이하 MRS-HAR 모형)과 관측적으로 동일해지기 때문에 유사한 추정결과를 얻을 수 있음이 밝혀져 있다. HAR 모형의 경우에는 최소자승법으로 추정이 가능하지만, 여기서는 모형 간 우도함수의 최대값 비교를 위해 최우추정법으로 세 모형을 추정하도록 하겠다.⁴ 국면 전환을 이용한 모형들을 추정할 경우에는 $s_t = 1$ 일 때 변동성이 높은 국면을 나타내도록 판별하기 위하여 각 국면의 절편 값에 대해 $c_0 < c_1$ 을 가정하였고, 이러한 제약 하의 마코프 전환 필터를 이용해 최우추정치를 도출하였다.

3.2. 추정 결과

표 (2)는 전체 표본기간과 하위 표본기간에 대한 HAR 모형과 ERS-HAR 모형의 추정 결과를 나타낸다. 우선 HAR 모형의 추정 결과는 전체 표본기간에서 일 단위 변동성 $\log RV_{t-1}^{(d)}$ 의 계수가 0.431로 가장 크고, 주 단위 변동성 $\log RV_{t-1}^{(w)}$ 의 계수가 0.362, 월 단위 변동성 $\log RV_{t-1}^{(m)}$ 의 계수가 0.156로 가장 낮

⁴ 최우추정법을 이용한 경우에도 최소자승법을 이용한 HAR 모형과 동일한 추정결과가 도출되었다.

은 값을 가진다. 하위 표본기간에도 계수들의 추정치들은 전체 표본기간의 값들과 유사한 것으로 나타났다. 전체와 하위 표본기간 모두에서 단기·중기·장기 세 단위 변동성이 모두 유의미한 영향을 가지는 것으로 나타나 하루, 일주일, 한 달의 변동성 요소들로 구성된 HAR 모형이 실현 변동성 데이터를 적절하게 설명하고 있음을 알 수 있다. 그러나 HAR 모형의 경우 잔차항의 자기상관관계에 대하여 Ljung-Box Q 검정을 실시하였을 때 시차 5, 10, 20에 대하여 모두 자기상관이 남아있는 것으로 관측되었다. 표 (2)의 최하단 부분은 Ljung-Box Q 검정에 대한 p-value를 나타낸다. HAR 모형은 두 표본기간에 대하여 잔차항에 자기상관이 존재하지 않는다는 Q 검정의 귀무가설을 시차 5, 10, 20에 대해 유의수준 1% 또는 5%에서 모두 기각하는 것으로 나타났다. 이는 기존 연구결과들에서 알려진 것처럼 HAR 모형의 단순한 구조로는 실현 변동성을 충분히 설명하지 못한다는 것을 의미한다.

반면 국면 전환 구조를 HAR 모형에 추가하였을 경우에는 HAR 모형과는 달리 잔차항에 자기상관이 존재하지 않는다는 검정 결과가 도출되었다. ERS-HAR 모형의 잔차항에 Ljung-Box Q 검정을 실시한 결과 모든 경우에서 유의수준 5%에서 자기상관이 없는 것으로 나타났다. 이는 단순한 구조의 HAR 모형으로는 충분히 설명하지 못하는 실현 변동성의 특징을 국면 전환 구조를 이용하여 적절하게 모델링할 수 있음을 의미한다.

ERS-HAR 모형을 추정한 결과 각 국면별 변동성의 움직임에서 흥미로운 결과가 관찰되었다. $s_t = 1$ 인 국면에서는 전기의 단기 변동성이 현재 변동성에 가장 큰 영향을 미치는 반면, $s_t = 0$ 인 국면에서는 전기의 중기 또는 장기 변동성이 가장 큰 영향을 미치는 것으로 관측되었다. 각 국면의 ERS-HAR 모형의 추정치들을 살펴보면, 우선 절편 값 $c(s_t)$ 는 상태변수 값이 $s_t = 0$ 일 때에 $s_t = 1$ 일 때 보다 약 2배 정도 작게 나타났다. 이에 따라 국면 1($s_t = 1$)을 고변동성 국면 그리고 국면 0($s_t = 0$)을 저변동성 국면으로 파악할 수 있다.

전체 표본기간의 경우를 먼저 살펴보면, 일 단위 변동성의 계수는 저변동성 국면($s_t = 0$)에서 유의하지 않은 낮은 값(0.032)으로 추정되었고 고변동성 국면($s_t = 1$)에서는 계수들중 가장 큰 값(0.662)으로 유의하게 추정되었다. 주 단위 변동성의 계수와 월 단위 변동성의 계수는 저변동성 국면에서 각각 0.505와 0.381의 큰 값들로 유의하게 추정되었고 고변동성 국면에서는 이에 비해서 낮은 값들이 낮아져 각각 0.191과 0.103으로 추정되었다.

2: 모형 추정결과

Sample Periods	전체 표본 (2006-2015)		하위 표본 (2010-2015)	
Model	HAR	ERS-HAR	HAR	ERS-HAR
	$s_t = 0$	$s_t = 1$	$s_t = 0$	$s_t = 1$
α		0.978 *** (0.02)		0.965 *** (0.02)
ρ		-0.106 * (0.05)		-0.175 ** (0.08)
τ		-0.728 *** (0.06)		-2.122 *** (0.13)
c	-0.507 *** (0.09)	-0.840 *** (0.07)	-0.416 *** (0.03)	-1.753 *** (0.26)
β_d	0.431 *** (0.01)	0.032 (0.05)	0.662 *** (0.03)	-0.120 *** (0.02)
β_w	0.362 *** (0.03)	0.505 *** (0.05)	0.191 *** (0.02)	0.342 *** (0.02)
β_m	0.156 *** (0.02)	0.381 *** (0.03)	0.103 *** (0.02)	0.597 *** (0.03)
σ	28.722 *** (0.39)	0.550 *** (0.01)	23.165 *** (0.04)	0.565 *** (0.05)
log likelihood	-11989.249	-2133.369	-6887.955	-1339.442
Q(5)	0.0000	0.0812	0.0016	0.4894
Q(10)	0.0002	0.2802	0.0069	0.7090
Q(20)	0.0011	0.2464	0.0390	0.7700

Notes: 1) *, **, ***는 각각 10%, 5%, 1% 유의수준 하에서 유의함을 나타낸다.
 2) Q(k)는 시차 k 까지의 Q 검정통계량의 p-value를 의미한다.
 3) 괄호 안의 수치는 표준오차를 나타낸다.

하위 표본기간의 경우도 전반적으로는 전체 표본기간의 결과와 유사하다. 일 단위 변동성의 계수는 고변동성 국면에서 가장 큰 값인 0.623으로 추정된 반면, 저변동성 국면에서는 -0.120으로 추정되었다. 주 단위 변동성과 월 단위 변동성의 계수는 하위 표본기간에서도 저변동성 국면에서 고변동성 국면보다 더욱 높은 값들로 추정되었다. 다만, 전체 표본기간에서는 주 단위 변동성의 계수가 월 단위 변동성의 계수보다 큰 값으로 추정되었으나, 하위 표본기간에서는 저변동성 국면에서 월 단위 변동성의 계수가 주 단위 변동성의 계수보다 큰 값인 0.597로 추정되었다.

고변동성 국면의 월 단위 변동성의 계수는 전체 표본기간과 하위 표본기간에서 가장 낮은 값으로 추정되었는데, 이는 변동성이 높은 기간에는 장기 변동성 요인의 영향이 미미하다는 것을 의미한다. 반면 저변동성 국면에서는 월 단위 변동성의 영향에 대한 추정치가 전체 표본기간과 하위 표본기간에서 모두 큰 값들로 추정되었는데, 이는 변동성이 낮은 기간에는 장기 변동성 요인의 영향력이 크다는 것을 의미한다. 이에 반하여 단기 변동성 요소를 나타내는 일 단위 변동성의 계수는 정반대의 패턴을 보인다. 변동성이 낮은 국면에서는 유의하지 않거나 미미한 영향을 미치는 반면, 변동성이 높은 국면에서는 가장 지배적인 영향을 미치는 요소로 나타났다. 이러한 경향은 전체 표본기간과 하위 표본기간에서 공통적으로 나타나 위기 기간의 존재 유무에 상관없이 일관성 있는 추정 결과로 드러났다.

이렇듯 각 국면에서 대조적으로 추정된 단기·장기 계수들의 패턴은 앞서 언급한 HAR 모형의 이론적 배경인 이질적 시장 가설을 통해 해석할 수 있다. 장기 투자자들이 장기 변동성 요소를 구성하고 하루 이하의 주기로 투자 계획을 변경하는 단기 투자자들이 단기 변동성 요소를 구성한다고 할 때, ERS-HAR 모형의 추정 결과는 변동성이 높은 기간에는 단기 투자자들의 영향을 크게 받고 변동성이 낮은 기간에는 장기 투자자들의 영향이 지배적인 것으로 해석할 수 있다. 일반적으로 주식시장의 장기 투자자들은 투자 계획 수립시 자산의 펀더멘탈 요소들이나 시장의 장기적인 전망을 고려하여 안정적인 수익을 얻는 것을 목표로 한다고 알려져 있다. 펀더멘탈 트레이딩(fundamental trading)이라고도 불리는 장기 투자는 안정적인 투자를 선호하며, 단기적인 시장상황의 변화에 의해 투자 계획을 수정하는 경우가 적기 때문에 주식시장의 급격한 변동을 유도하기 보다는 장기적인 시장 변동성의 추세를 결정하는 역할을 한다. 반면, 데이 트레이딩(day trading) 등 단기적인 투자를 하는 거래자들은 하루에도 여러번 투자 계획을 변경하며 단기적인 차익을 내는 것을 목표로 하는 것으로 알려져 있다. 이들은 시장의 비효율적인 부분에 투자하여 시장 가격이 균형가격에 벗어났을 때, 즉 어떤 기업의 주가가 기업의 본래

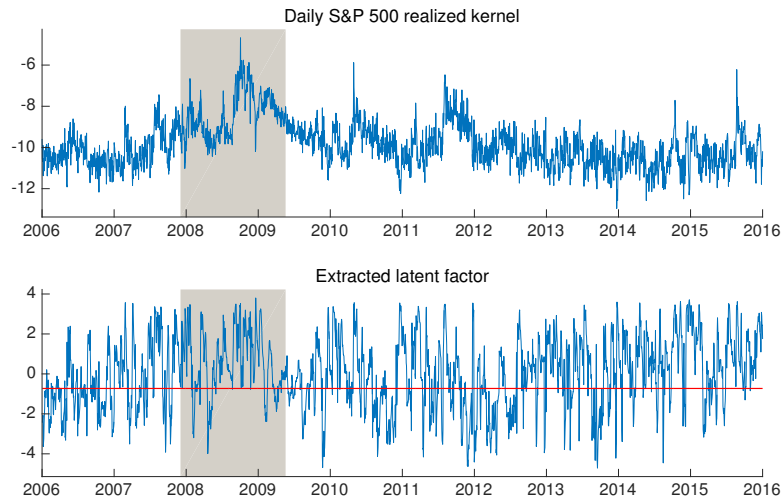
특성(fundamental characteristic)보다 저평가된 시점에 투자하여 기업에 대한 평가가 긍정적으로 변화하게 되면 단기간에 자산을 매도함으로써 균형 가격으로부터의 차이 만큼의 수익을 얻는 경향이 있다. 이러한 단기 투자는 시장이 비효율적일 때 그리고 주식시장의 변동성이 높을 때에 더욱 활발하게 일어나는 경향이 있다. 따라서 변동성이 높은 시기에는 단기 투자자의 영향력이 크고 변동성이 낮은 시기에는 장기 투자자의 영향이 크게 드러나는 것이다.

ERS-HAR 모형에서 국면 전환의 내생성을 설명하는 모수 ρ 는 전체 표본기간에서 -0.106으로 추정되었고 10% 수준에서 유의한 것으로 나타났다. 하위 표본기간에서는 -0.175로 추정되었으며 5% 수준에서 유의한 것으로 나타났다. ρ 에 대한 이러한 추정 결과는 전기의 충격이 현재 기의 국면 결정에 아주 큰 영향은 아니더라도 유의미한 영향을 미치고 있다는 것을 의미한다. 또한 ρ 값이 음(-)의 값으로 추정되었다는 것은 변동성의 평균 회귀 성향(mean reversion behavior)을 나타내는 것으로 해석할 수 있다. 즉, 음의 충격(negative shock)이 발생해 전기의 변동성이 작아졌을 경우에는 음(-)의 상관관계에 따라 현재 기의 잠재변수 값이 커지게 되므로 고변동성 국면에 놓일 확률이 커진다. 반대로 양의 충격(positive shock)이 발생해 전기의 변동성이 커졌을 경우 잠재변수 값이 작아지면서 저변동 국면에 놓일 확률이 커지는 것을 알 수 있다. 이 때 ρ 의 추정치의 절대값이 0에 가까운 점을 고려하여 $\rho = 0$ 이라는 제약을 두고(내생성을 제거하고) ERS-HAR 모형을 추정해 보았는데, 표 (3)의 추정 결과를 보면 표 (2)에서 나타난 ERS-HAR 모형의 결과들과 전체적으로 크게 달라지지 않는다는 것을 알 수 있다. 또한 표 (3)은 기존의 마코프 전환 HAR 모형의 추정 결과를 제시하고 있는데, 앞서 모형 설명에서 밝힌 바와 같이 외생적 ER-HAR 모형($\rho = 0$)과 추정 결과가 거의 일치하는 것을 알 수 있다.

내생적 국면 전환 모형의 각 국면은 식 (3)과 식 (4)에서 보였듯 1차 자기회귀과정을 따르는 잠재 변수가 역치 τ 값보다 클 때 국면 1로, 작을 때 국면 0으로 결정된다. 내생적 국면 전환 모형의 특징 중 하나는 추정치들을 이용하여 잠재 변수 w_t 를 추출하여 그 움직임을 살펴볼 수 있다는 점이다. 그림 (2)는 전체 표본기간 중 ERS-HAR 모형의 잠재 변수와 역치 수준을 나타내고 있다. 그림 (2)의 상단 패널은 추정에 사용한 종속 변수인 로그 형태의 일별 실현 변동성 데이터이고, 하단 패널은 추정 결과 추출된 잠재 변수이다. 하단 패널 중앙의 붉은 선이 국면을 식별하는 역치 수준을 나타낸다. 전반적으로 실제 실현 변동성 수준이 증가하는 구간이 변동성이 높은 국면으로 판별되는 경우가 많으며, 특히 세계 금융위기로 분류되는 2007년 12월부터 2009년 5월까지의 기간(그래프의 음영 처리된 부분)중에 고변동 국면이 오래 지속되는 것을 살

퍼볼 수 있다. 금융위기 이후의 기간에 대해서는 변동성 국면의 전환이 굉장히 활발하게 일어나고 있다.

2: 추출된 잠재 변수(latent factor)



Notes: 아래 그림은 상태 변수를 결정하는 잠재 변수를 나타낸다. 추출된 잠재 변수는 AR(1) 과정을 따르며, 역치(threshold) 수준을 나타내는 중앙의 직선보다 크면 높은 변동성 국면을 나타내는 $s_t = 1$, 작으면 낮은 변동성을 보이는 $s_t = 0$ 인 국면으로 결정된다. 음영처리된 부분은 2008년 세계 금융위기 기간을 나타낸다.

3: 추정 결과: MRS-HAR 모형과 제약 하의 ERS-HAR 모형 비교

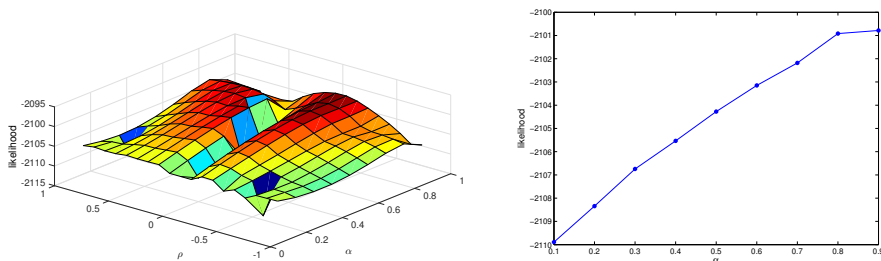
Sample Periods	전체 표본(2006-2015)			하위 표본(2010-2015)		
	MRS-HAR	RS-HAR($\rho = 0$)	MRS-HAR	RS-HAR($\rho = 0$)	MRS-HAR	RS-HAR($\rho = 0$)
Model	$s_t = 0$	$s_t = 1$	$s_t = 0$	$s_t = 1$	$s_t = 0$	$s_t = 1$
α			0.978 *** (0.01)			0.975 *** (0.05)
τ			-0.689 *** (0.09)			-2.285 *** (0.18)
c	-0.926** (0.43)	-0.407** (0.17)	-0.863*** (0.12)	-0.412 *** (0.06)	-1.693 *** (0.85)	-1.637 * (0.81)
β_d	0.032 (0.09)	0.655 *** (0.06)	0.035 (0.06)	0.662 *** (0.02)	-0.099 (0.20)	-0.082 *** (0.03)
β_w	0.476 *** (0.13)	0.200 *** (0.07)	0.494 *** (0.14)	0.192 *** (0.04)	0.325 (0.17)	0.327 (0.25)
β_m	0.402 *** (0.11)	0.103 ** (0.04)	0.387 *** (0.08)	0.103 *** (0.03)	0.599 (0.17)	0.587 *** (0.14)
σ	0.550 *** (0.01)		0.550 *** (0.02)		0.566 *** (0.01)	0.566 *** (0.02)
p	0.943 *** (0.03)		0.943 *** (0.03)		0.954 *** (0.02)	
q	0.076 ** (0.03)		0.076 ** (0.03)		0.110 (0.08)	
log likelihood	-2135.954		-2133.479		-1340.773	
					-1339.670	

Notes: 1) *, **, ***는 각각 10%, 5%, 1% 유의수준 하에서 유의함을 나타낸다.

2) 괄호 안의 수치는 표준오차를 나타낸다.

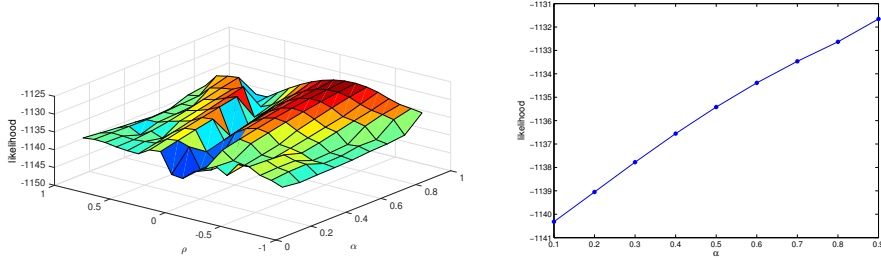
국면 전환과 같은 고도의 비선형 모형의 최우추정은 수치적 최적해(numerical solution)를 찾는 과정에서 국지적 최적치(local maximum)로 수렴하는 경우가 빈번하게 나타나는 어려움이 있다. 따라서 표 (2)에 나타난 추정 결과가 전역적 최적치(global maximum)를 달성하고 있는지를 확인하기 위해 모형에서 민감한 모수들인 α 와 ρ 에 대한 로그 프로파일 우도함수(profile likelihood)를 살펴보았다. 그림 (3)과 그림 (4)는 α 와 ρ 의 값을 각각 0.1 간격으로 조정하면서 다른 모수들을 추정하여 도출한 로그 우도함수의 값들을 나타내고 있다. 그림 (3)은 전체 표본기간, 그림 (4)는 하위 표본기간에 대한 프로파일 우도함수를 나타낸다. 전체표본의 경우 $(\alpha, \rho) = (0.9, -0.1)$ 일 때 전역적 최댓값으로 수렴함을 확인할 수 있는데, 이는 표 (2)의 ERS-HAR 모형의 추정 결과 $(\hat{\alpha}, \hat{\rho}) = (0.978, -0.106)$ 에 대응하고 있으므로 표 (2)의 ERS-HAR 모형의 추정 결과가 전역적 최적치로 수렴하였음을 의미한다. 하위 표본기간의 경우 $(\alpha, \rho) = (0.9, -0.2)$ 일 때 전역적 최적치로 수렴함을 확인할 수 있는데, 이는 표 (2)의 ERS-HAR 모형의 추정 결과 $(\hat{\alpha}, \hat{\rho}) = (0.965, -0.175)$ 에 대응하고 있으므로 표 (2)의 ERS-HAR 모형의 추정 결과가 하위 표본기간에도 전역적 최적치로 수렴하였음을 의미한다.

3: 로그 우도함수에 대한 전역적 최적해 확인(전체 표본)



Notes: 로그 우도함수에 대해 민감한 모수 α 와 ρ 에 대한 로그 프로파일 우도함수(profile likelihood)를 나타내고 있다. α @ ρ 의 값을 각각 0.1 간격으로 조정하면서 다른 모수들을 추정하여 도출한 로그 우도함수의 값들을 나타내고 있다. 전체표본에서는 국면 전환의 내생성이 허용된 경우(왼쪽 그림) $(\alpha, \rho) = (0.9, -0.1)$ 일 때, 그리고 내생성이 제거된($\rho = 0$) 경우(오른쪽 그림) $\alpha = 0.9$ 일 때 전역적 최적치를 확인하였다.

4: 로그 우도함수에 대한 전역적 최적해 확인(하위 표본)



Notes: 로그 우도함수에 대해 민감한 모수 α 와 ρ 에 대한 로그 프로파일 우도함수(profile likelihood)를 나타내고 있다. α @ ρ 의 값을 각각 0.1 간격으로 조정하면서 다른 모수들을 추정하여 도출한 로그 우도함수의 값들을 나타내고 있다. 하위표본에서는 국면 전환의 내생성이 허용된 경우(왼쪽 그림) $(\alpha, \rho) = (0.9, -0.2)$ 일 때, 그리고 내생성이 제거된($\rho = 0$) 경우(오른쪽 그림) $\alpha = 0.9$ 일 때 전역적 최적치를 확인하였다.

3.3. 모형의 적합성 비교

3.3.1 적합성 평가 방법

변동성 모형의 적합성을 평가하기 위하여 표본 내 예측치 또는 표본 외 예측치들의 비교를 할 때 손실함수(loss function)를 선택해야 하는데, 변동성 모형 관련 연구에서는 변동성 대용치(volatility proxy)의 잡음(noise) 대해 강건한 손실함수를 사용하여야 한다는 점이 잘 알려져 있다. 본 연구에서는 Hansen and Lunde (2006)와 Patton (2011)의 연구에 따라 강건한 손실함수로 알려진 MSE(Mean Squared Error)와 QLIKE 손실함수를 채택하였다. 두 손실함수는 다음과 같이 정의된다.

$$MSE_t = (RV_t - \widehat{RV}_t)^2$$

$$QLIKE_t = \frac{RV_t}{\widehat{RV}_t} - \log \frac{RV_t}{\widehat{RV}_t} - 1$$

모형들의 손실함수 값의 크기를 비교하고 이 손실함수의 값들 간의 차이가 유의한지를 보기 위하여 DMW 검정(Diebold-Mariano-West test) (Diebold and Mariano, 1995; West, 1996)을 실시하였다. DMW 검정은 표본 외 예측에 대한 평가를 위하여 설계되었으나(Diebold, 2015), 두 모형 간 비교척도 역할을 하기 때문에 표본 내 적합성 평가에도 활발히 적용되고 있다. 본 연구에서는 ERS-HAR 모형의 표본 내 적합성을 기존의 HAR 모형 등에 비교하기 위해 DMW

검정을 실시하였다. DMW 검정은 두 모형 간 예측력에 차이가 있는지를 비교하는 쌍별 비교(pairwise comparison) 방식으로 시행되는데, 우선 두 모형에서 각각 도출한 손실함수 값의 차이를

$$d_t = L(\widehat{RV}_{1t}, RV_t) - L(\widehat{RV}_{2t}, RV_t)$$

로 정의한다. 손실함수 $L(\cdot)$ 은 앞에서 밝힌 MSE와 QLIKE 손실함수들을 각각 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.

3.3.2 적합성 평가 결과

표 (4)은 HAR 모형, 상태 변수의 내생성을 고려한 ERS-HAR 모형, ERS-HAR 모형에서 내생성을 제거한 모형($\rho = 0$) 등 세 모형의 표본기간 내 적합성을 평가한 결과를 보여주고 있다. 첫째, 각 모형의 MSE와 QLIKE 손실함수 값들을 비교해 보면, 전체 표본과 하위 표본 모두에서 ERS-HAR 모형의 MSE와 QLIKE 손실함수 값이 가장 낮은 것으로 나타났고 HAR 모형이 가장 높은 손실함수 값을 나타냈다. ERS-HAR 모형과 내생성을 제거한 RS-HAR 모형의 손실함수 값들은 미세한 차이를 보이는 반면, 국면 전환을 허용하지 않는 HAR 모형의 경우 손실함수 값들이 더욱 크게 나타나고 있다. 이 결과는 3.2절의 모형 추정결과에서 살펴본 결과(HAR 모형의 잔차항에는 시차 의존성이 여전히 남아있으나 국면 전환을 허용할 경우 잔차항의 시차 의존성이 사라지는 결과)와 같이 HAR 모형 그 자체로는 실현 변동성을 설명하기에 충분하지 않다는 사실을 의미한다.

4: 표본기간 내 적합성(in-sample fit) 비교

	HAR	RS-HAR($\rho = 0$)	ERS-HAR
전체 표본: Jan 03, 2006 - Dec 31, 2015 (2,510 obs)			
MSE	0.3287	0.2845	0.2842
DMW _{MSE}	12.506 ***	2.415 **	
QLIKE	1.7959×10^{-3}	1.5538×10^{-3}	1.5507×10^{-3}
DMW _{QLIKE}	11.644 ***	2.626 ***	
하위 표본: Jan 04, 2010 - Dec 31, 2015 (1,510 obs)			
MSE	0.3554	0.2982	0.2957
DMW _{MSE}	9.559 ***	5.399 ***	
QLIKE	1.8280×10^{-3}	1.5372×10^{-3}	1.5237×10^{-3}
DMW _{QLIKE}	9.149 ***	5.718 ***	

Notes: 1) DMW 통계량들은 다음과 같은 손실함수의 차이를 통해 계산되었다.

$$d_t = L(\widehat{RV}_{benchmark,t}, RV_t) - L(\widehat{RV}_{ERS-HAR,t}, RV_t).$$

여기서 $\widehat{RV}_{benchmark,t}$ 는 HAR 모형 또는 RS-HAR($\rho = 0$) 모형의 표본기간 내 예측치이고 $\widehat{RV}_{ERS-HAR,t}$ 는 ERS-HAR 모형의 표본기간 내 예측치이다.

2) *, **, ***는 각각 10%, 5%, 1% 유의수준 하에서 유의함을 나타낸다.

둘째, DMW 검정을 통해 ERS-HAR 모형과 나머지 두 모형들의 표본기간 내 적합도에 통계적으로 유의한 차이가 있는지를 살펴보았다. ERS-HAR 모형을 기준으로 다른 두 모형과 각각 DMW 검정을 실시하였을 때 모든 경우에서 1% 또는 5% 유의수준에서 귀무가설을 기각하고 있다. 이는 ERS-HAR 모형이 다른 두 모형보다 통계적으로 유의하게 우월한 적합성을 나타내고 있음을 의미한다. 특히 내생성을 제거한 RS-HAR 모형에 비해서도 ERS-HAR 모형이 통계적으로 유의하게 낮은 손실함수 값을 나타내는 점은 국면 전환 HAR 모형에서 내생성을 허용하는 것이 모형 적합성 측면에서 중요함을 의미한다.

4. 결론

본 논문에서는 내생적 국면 전환을 허용하는 HAR 모형을 이용하여 S&P 500 지수 수익률의 변동성을 분석하였다. 2006년 1월 3일부터 2015년 12월 31일까지의 전체 표본기간과 금융위기 기간을 제외한 2010년 1월 4일부터 2015년 12월 31일까지의 하위 표본기간 모두에서 내생적 국면 전환을 도입하는 것이 기존 모형들에 비해 변동성에 대한 설명력을 크게 향상시키는 것으로 나타났다. 기존의 HAR 모형의 경우에는 잔차항에 자기상관이 남아있는 것으로 나타나 실현 변동성을 충분히 설명하지 못하는 것으로 보였으나, 내생적 국면 전환을 도입한 HAR 모형의 경우에는 잔차항에 자기상관이 존재하지 않는 것으로 나타났다. 또한 MSE 손실함수와 QLIKE 손실함수를 고려하여 Diebold-Mariano(1995)와 West(1996)의 검정을 시행한 결과, 내생적 국면 전환을 도입한 HAR 모형이 기존 모형들에 비해 통계적으로 유의하게 표본기간 내 설명력이 높다고 나타났다.

내생적 국면 전환을 도입한 HAR 모형을 추정한 결과중 가장 주목할 점은 변동성이 높고 낮은 각 국면별로 다음 기 변동성에 가장 크게 영향을 미치는 요인들이 상이하게 드러났다는 점이다. 즉, 변동성이 높은 국면에서는 단기 변동성 요인이 다음 기 변동성에 지배적인 영향을 미치고, 변동성이 낮은 국면에서는 장기 변동성 요인이 지배적인 영향을 미치는 것으로 나타났다. 이는 변동성이 높은 기간에는 주식시장에서 단기 투자자들의 영향력이 큰 반면, 변동성이 낮은 기간에는 장기 투자자들의 영향이 크게 드러나는 것으로 해석할 수 있다. 본 논문에서 제안하고 분석한 실현 변동성 모형은 이처럼 변동성이 높고 낮은 국면별로 상이한 투자자들의 행태를 효과적으로 반영하고 있다. 본 논문에서는 두 상태(two-state) 국면만을 허용하는 모형을 고려하고 있는데, 향후 거시경제 변수들과의 결합, 세 상태(three-state) 국면 이상의 국면 전환을 허용하는 등의 방법으로 모형을 확장할 수 있을 것으로 기대된다.

References

- Andersen, T., T. Bollerslev, and F. Diebold (2007) Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility, *The Review of Economics and Statistics*, 89, 701-720.
- Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003) Modeling and forecasting realized volatility, *Econometrica*, 71, 579-625
- Banerjee, A. and G. Urga (2005) Modeling structural breaks, long memory and stock market volatility: an overview, *Journal of Econometrics*, 129, 1-34.
- Barndorff-Nielsen, O.E. and N. Shephard (2002) Estimating quadratic variation using realized variance, *Journal of Applied Econometrics*, 17, 457-477.
- Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde and N. Shephard (2008) Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise, *Econometrica*, 76, 1481-1536.
- Bauwens, L., A. Preminger, and J.V. Rombouts (2010) Theory and inference for a Markov switching GARCH model, *The Econometrics Journal*, 13, 218-244.
- Chang, Y., Y. Choi, and J.Y. Park (2016) A new approach to model regime switching, *Journal of Econometrics*, forthcoming.
- Corsi, F. (2009) A simple approximate long-memory model of realized volatility, *Journal of Financial Econometrics*, 7, 174-196.
- Corsi, F., S. Mittnik, C. Pigorsch, and U. Pigorsch (2008) The volatility of realized volatility, *Econometric Reviews*, 27, 46-78.
- Craioveanu, M., and E. Hillebrand (2012) Level changes in volatility models, *Annals of Finance*, 8, 277-308.
- Deo, R., C. Hurvich, and Y. Lu (2006) Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment, *Journal of Econometrics*, 131, 29-58.

- Diebold, F.X. (2015) Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests, *Journal of Business & Economic Statistics*, 33, 1-1.
- Diebold, F.X., and R.S. Mariano (1995) Comparing predictive accuracy, *Journal of Business & Economic Statistics*, 13, 253-263.
- Ding, Z., C.W.J. Granger, and R.F. Engle (1993) A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, 1, 83–106.
- Goldfeld, S.M. and R.E. Quandt (1973) A Markov model for switching regressions, *Journal of Econometrics*, 1, 3-16.
- Gray, S.F. (1996) Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics*, 42, 27–62.
- Hamilton, J. (1989) A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.
- Hamilton, J., and L. Gang (1996) Stock market volatility and the business cycle, *Journal of Applied Econometrics*, 11, 573-593.
- Hamilton, J., and R. Susmel (1994) Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics*, 64, 307-333.
- Hansen, P.R., and A. Lunde (2006) Consistent ranking of volatility models, *Journal of Econometrics*, 131, 97-121.
- Heber, G., A. Lunde, N. Shephard, and K. Sheppard (2009) *OMI's realised library*, Version 0.1, Oxford-Man Institute, University of Oxford.
- Hwang, E., and D. W. Shin (2014) Infinite-order, long-memory heterogeneous autoregressive models, *Computational Statistics and Data Analysis*, 76, 339-358
- Kim, C.J. (1994) Dynamic linear models with Markov-switching, *Journal of Econometrics*, 60, 1-22.
- Kim, C.J., and C.R. Nelson (1999) *State-Space models with regime switching*, MIT Press.

- Kim, C.J., J. Piger, and R. Startz (2008) Estimation of Markov regime-switching regression models with endogenous switching, *Journal of Econometrics*, 143, 263-273.
- Liu, C., and J.M. Maheu (2008) Are there structural breaks in realized volatility?, *Journal of Financial Econometrics*, 6, 326-360.
- Martens, M., M. De Pooter, and D.J. Van Dijk (2004) Modeling and forecasting S&P 500 volatility: Long memory, structural breaks and nonlinearity, Tinbergen Institute Discussion Paper No. 04-067.
- McAleer, M., and M.C. Medeiros (2008) A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries, *Journal of Econometrics*, 147, 104-119.
- Müller, U.A., M.M. Dacorogna, R.D. Davé, R.B. Olsen, O.V. Pictet, and J.E. Von Weizsäcker (1997) Volatilities of different time resolutions—analyzing the dynamics of market components, *Journal of Empirical Finance*, 4, 213-239.
- Patton, A.J. (2011) Volatility forecast comparison using imperfect volatility proxies, *Journal of Econometrics*, 160, 246-256.
- Pypko, S. (2015) Volatility forecast in crises and expansions. *Journal of Risk and Financial Management*, 8, 311-336.
- Wang, Y., F. Ma, Y. Wei, and C. Wu (2016) Forecasting realized volatility in a changing world: A dynamic model averaging approach, *Journal of Banking & Finance*, 64, 136-149.
- West, K.D. (1996) Asymptotic inference about predictive ability, *Econometrica*, 64, 1067-1084.