

混合週期資料 VAR模型을 이용한 分期 GDP成長率의 最適豫測*

김 윤 영** . 박 준 용***

논 문 초 록

본고에서 우리는 좀 더 정확한 분기GDP성장을 예측을 위해서 분기별 GDP성장을 외에 다른 월별 경제변수도 함께 이용할 수 있는 방법을 개발하였다. 이를 위하여 우선 GDP 성장률과 기타 관련된 경제변수 모두를 포함하는 월별 VAR(vector autoregressive) 모델을 상정한 후, 이로부터 오직 분기별로만 관측되는 GDP성장률을 위한 예측식을 유도하였다. 유도된 예측식은 유한한 시차를 가지는 이동평균 오차를 갖게 되며, 따라서 차분변수를 도구변수로 하는 GMM 방법으로 쉽게 추정할 수 있다. 제시된 방법이 현실적으로 잘 적용될 수 있는가 알아보기 위하여, 실제로 한국의 분기GDP성장을 예측을 시도해 보았다. 분기별 GDP성장률 외에 월별로 관측되는 소비자물가지수(CPI) 및 여타 금융변수 등을 고려한 바, 제 금융변수들 중 월별 통화량 M1과 회사채의 신용스프레드를 각각 추가 이용할 경우 기존의 분기VAR모형이나 단일변수 AR(autoregressive) 모형 등에 비해 예측 평균자승오차(forecasting MSE) 또는 중위수 오차(median bias)를 현저히 개선시킬 수 있는 것으로 나타났다.

핵심 주제어: 혼합주기자료, VAR모형, GDP성장을 최적예측

경제학문헌목록 주제분류: C3

* 본 연구를 위해 좋은 지적을 해주신 남상호, 김기호, 유병학, 홍승제, 장동구, 오금화 박사와 익명의 두 심사위원께 깊이 감사드린다. 본 연구는 학술진흥재단의 연구비 지원에 의해 수행되었다.

** 제1저자, 한국은행 금융경제연구원, e-mail: yunyeongkim@hotmail.com

*** 제2저자, 성균관대학교 및 Texas A&M University, e-mail: jpark@neo.tamu.edu

I. 머리말

단기 GDP성장률을 예측하는 데는 통상적인 구조적 거시계량모형이외에 다양한 시계열모형들이 사용되고 있다. 특히 Sims(1972, 1980)로부터 시작되어 거시경제학의 실증적 연구에 널리 쓰이고 있는 VAR모형은 실제 예측에 있어서 매우 폭넓게 사용되고 있으며 대체적으로 그로부터 유용한 예측결과를 얻을 수 있는 것으로 판단되고 있다. 일찍이 Nelson(1972)은 경제성장률에 대한 단기예측을 하는 데는 ARIMA 등 시계열모형이 대규모 구조적 거시계량모형에 못지않게 유용함을 보여주었다. 또한 Litterman(1986)은 단일변수모형을 확장한 베이지안 VAR모형을 통해 보다 정확한 경제예측치를 얻을 수 있음을 보여주고 있다.

하지만 실제로 사용되는 분기별 GDP성장률 예측을 위한 대부분의 VAR모형들은 시차분기자료만을 이용할 뿐 중요한 분기내 다른 월별 경제자료들을 전혀 활용하지 않고 있다. 그러나 월별로 관측되는 귀중한 정보들을 무시한다면 필연적으로 GDP 성장률 분기예측결과의 정확도는 떨어지게 될 것이다. 이와 같은 가능성은 기존의 많은 연구 결과들에서 찾아볼 수 있다. 실제로 Abeyasinghe(1998), Lupoletti and Webb(1986), Miller-Chin(1996), Nunes(2005), Ingentino-Trehan(1996), Webb(1999), Zheng and Rossiter(2006) 등은 VAR모형을 이용하여 분기 경제성장률을 예측할 때 분기내 월별 정보변수를 반영할 경우 예측오차가 줄어들 여지가 있음을 보이고 있다.

그러나 이들이 사용하고 있는 예측모형은 모두 일관된 논리 없이 매우 자의적(ad hoc)으로 설정되어 있다.¹⁾ 월별 및 분기별로 관측되는 변수들의 자료가 섞여있는 혼합주기자료(mixed-frequency data)를 보다 효과적으로 다루기 위해서는, 이들 모두를 포함하는 하나의 일관된 VAR모형을 설정한 후 이로부터 분기별로 관측되는 GDP성장률을 위한 적절한 예측모형을 도출하는 것이 바람직하다고 할 수 있다. 이와 같은 접근방식은 기존의 몇몇 연구에서도 시도되었다. 하지만 그들 중 어떤 것도 일반적으로 유용한 통계학적 분석방법을 제시하지 못하였을 뿐 만 아니라, 여러

1) 가령 Abeyasinghe는 분기GDP와 월별 무역액으로 구성된 2변수의 자의적인 모형에서 예측치를 유도하고 있다. 그런데 이 모형은 월별 VAR 모형에 계수 제약(coefficient restriction)을 가했을 경우 얻어지는 특수한 경우임을 쉽게 보일 수 있다. 여타 다른 연구들도 유사한 단점을 가지고 있다.

한계점들을 가지고 있다.

첫째, Zadrozny(1990)은 모형의 일부변수가 부분적으로 일시 합산이 가능한 특수한 경우²⁾에만 적용이 가능한 VAR 모형의 비선형 칼만필터링 추정 방식을 제시하고 있다. 그러나 동 방법은 많은 모수를 지닌 거대모형에서는 적용이 거의 어려운 것으로 보고되고 있다. 더구나 Chen and Zadrozny(1998)에서 밝히고 있듯이 모형의 연산이나 식별 조건을 제시하지 못하는 점이 한계로 지적되고 있다.

둘째, Chen and Zadrozny(1998)는 혼합주기 자료 하에서 VAR모형을 추정하기 위하여 일종의 도구변수법(instrumental variable estimation)을 제안하고 있다. 이 방식은 울-워커 방정식(extended Yule-Walker equation)에서 혼합주기자료 때문에 관측불가능한 2차적률 들을 관측가능한 2차적률의 도구변수 추정치로 대체하는 방식이다. Cadzow(1982) 역시 과다결정(over-determination) 추정 이라는 개념으로 비슷한 방식을 제안하고 있다.

그러나 이들의 방법은 다음과 같은 단점들이 있다. 우선 울-워커 방정식 내에서 고빈도 시계열의 자기공분산 (autocovariance)에 대한 일치추정치를 저빈도 시계열에서는 모두 구할 수 없는 특성 때문에 모형 오차분산의 일치 추정이 불가능하다.³⁾ 이는 VAR 모형을 사용하게 되는 핵심적인 이유 중 하나인 동태분석(예측, 충격반응 함수 추정, 분산 분해)을 어렵게 한다. 또한 Kim and Park(2007)에서 보이는 바와 같이 이들 방법을 비정상성(nonstationarity)을 가지는 자료에 적용할 경우 모형의 자기회귀계수에 대한 일치추정량 조차 얻을 수 없다.

한편 Kim and Park은 위와 같은 선행연구들의 단점을 보완하여 혼합주기자료 하에서 모형의 비정상성 여부와 관계없이 VAR모형의 자기회귀계수와 오차분산의 일치추정이 가능한 방법을 제시하고 있다. 이 방법은 우선 관측가능한 자기회귀계수들만의 함수로 나타내어지는 추정방정식을 VAR모형으로부터 유도한 후, 이 추정방정식의 계수가 VAR 모형의 자기회귀계수로 나타내어지는 점에 착안하여 도구변수를 이용한 일반적률법(GMM)으로 자기회귀계수와 오차분산을 추정하는 것이다. 여기서 유도된 추정방정식은 혼합주기 자료 하에서의 최적예측식을 유도하는데 사

2) 가령 분기GDP는 3개월 월별 GDP의 합으로서 관측가능하다. 이는 월별 GDP를 추정하는 칼만필터링 과정에서 관측방정식(observation equation)으로 사용될 수 있다.

3) 이에 따라 Chen and Zadrozny(1998)는 오차분산의 근사식을 제시하고 있지만 이는 일치 추정량이 아니다.

용할 수 있으며 이의 일치추정은 먼저 추정된 자기회귀계수와 오차분산을 이용하게 된다.

이와 관련 본고에서는 Kim and Park에서 고려한 모형 중 가장 간단한 경우인 VAR(1) 모형을 이용하여 GDP 성장률 예측에 적용하는 방안을 소개하기로 한다. VAR(1) 모형의 경우 예측식이 특별히 간단한 형태로 유도되므로 GDP 예측에 혼합주기자료를 쉽게 이용할 수 있는 장점이 있다.

실증분석의 응용 예로서는 외환위기 이후 우리나라의 자료를 이용하여 3변수(실질GDP, CPI, 금융변수) VAR모형을 통해 월별 금융변수가 포함된 경우 분기 GDP 성장률에 대한 표본 외(out of sample) 예측이 개선되는지를 비교하였다. 분석결과 금융변수 중 월별 통화량 M1과 화사채의 신용스프레드를 각각 추가 이용할 경우 기존의 분기VAR모형이나 단일변수 AR모형 등에 비해 분기성장률에 대한 예측정확도가 높아지는 것으로 나타났다.

본고는 모두 4장으로 구성되어 있다. 먼저 II장에서는 월단위 금융변수를 포함한 혼합주기자료를 사용할 때의 예측식을 유도한다. 다만 예측식은 예측오차를 줄이기 위해 가급적 간결한 모형(parsimonious model)에서 유도하였으며, 예측시계(forecasting horizon)는 단기에 초점을 맞추었다. III장에서는 예측방법별로 GDP성장률 예측결과를 도출한다. 그리고 IV장에서는 결론 및 시사점을 제시한다.

II. 예측모형

먼저 이번 장에서는 분기GDP와 월별 금융변수의 혼합주기 자료를 동시에 고려하는 3변수 VAR(1) 모형의 추정방정식을 유도한다. 다음으로는 추정된 계수와 모형에서 유도된 예측방정식을 바탕으로 최적예측식을 추정하는 방법에 대해 설명하고자 한다.

1. 추정방정식의 유도 및 추정

먼저 추정방정식의 개념 소개를 위하여 x_t (2×1) 및 y_t (1×1)의 변수로 구성된 3변수 $z_t = (x_t', y_t)'$ 의 안정적(covariance stationary)인 월별 VAR(1) 모형을 가정

한다. 다만, 아래 <표 1>에서 보듯이 변수들을 예로 들면, x_t 는 월별금융변수와 소비자물가지수(CPI)를, 그리고 y_t 는 분기별 GDP를 각각 나타낸다고 상정한다.

$$z_t = Az_{t-1} + u_t \quad (1)$$

여기서 A 는 각각 차원이 $\alpha_{11} (2 \times 2)$, $\alpha_{12} (2 \times 1)$, $\alpha_{21} (1 \times 2)$, $\alpha_{22} (1 \times 1)$ 로 이루어진 3×3 의 다음과 같은 계수행렬이며,

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$$

$u_t = (v_t', w_t')'$ 는 $v_t (2 \times 1)$ 및 $w_t (1 \times 1)$ 로 구성된 다음과 같은 분산을 갖는 독립동일분포(independent and identically distributed)의 오차항으로 가정한다.

$$\Sigma = \begin{pmatrix} \Sigma_{vv} & \Sigma_{vw} \\ \Sigma_{vw}' & \Sigma_{ww} \end{pmatrix}$$

여기서 θ 는 계수 A 와 분산 Σ 를 원소로 포함하는 콤팩트 공간으로 정의한다.

변수 z_t 는 아래 <표 1>에서 예시한 자료관측 시점을 가지는 것으로 가정하는데 만일 z_t 가 매월 모두 관측가능하다면, 식 (1)의 계수 A 는 통상최소자승법(OLS)를 통하여 일치추정량을 얻을 수 있고 이를 통하여 예측 등의 동태분석을 수행할 수 있다. 그러나 GDP는 분기 자료로만 이용 가능하기 때문에 이러한 동태분석은 불가능하다. 여기서 <표 1>의 자료 관측 주기는 각각의 수준변수 만이 아니라 전기 대비 증감이나 증감률의 경우에도 동일하게 나타난다는 점에 유의할 필요가 있다.

이와 같은 혼합주기 자료의 단점을 감안하여 본고에서는 먼저 월별 VAR 모형에서 관측불가능한 변수들을 관측가능한 z_t 의 과거 시차변수들로 다시 표현하는 방식으로 ‘추정 방정식’을 유도하고, 이를 통하여 모형계수를 추정하는 방식을 소개하고자 한다. 이러한 추정계수들은 최적예측을 위한 조건부기대(conditional expectation)를 추정하는데 사용하게 된다.

〈표 1〉 월별 및 분기별 자료 관측 예

| 관측시점 (t) | x_t | | y_t |
|----------|-------|-----|-------|
| | 금융변수 | CPI | GDP |
| 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | × |
| 2 | 0 | 0 | × |
| 3 | 0 | 0 | 0 |
| 4 | 0 | 0 | × |
| 5 | 0 | 0 | × |
| 6 | 0 | 0 | 0 |
| 7 | 0 | 0 | × |
| 8 | 0 | 0 | × |
| 9 | 0 | 0 | 0 |
| ⋮ | ⋮ | ⋮ | ⋮ |

주: 0는 관측가능, ×는 관측불능을 표시.

다음은 좀 더 구체적으로 식 (1)을 이용하여 〈표 1〉과 같은 분기·월별 혼합주기 자료를 가지고 어떻게 추정 방정식을 유도하는지를 살펴보기로 한다. 먼저 관측시점이 $t=0, 3, 6, 9\ldots$ 인 경우, 식 (1)과 같은 VAR모형은 자료의 관측가능성에 따라 변수를 구분하여 다음과 같이 다시 쓸 수 있다.

$$z_t = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} x_{t-1} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix} y_{t-1} + u_t \tag{2}$$

여기서 식 (2)의 좌변의 z_t 와 우변의 x_{t-1} 는 〈표 1〉에서 보듯이 관측 가능하지만, 우변의 y_{t-1} 는 관측이 불가능하다. 그러나 관측 불가능한 변수 y_{t-1} 가 식 (1)과 한번 시차를 가진 식을 이용하여 아래와 같이 쓸 수 있음에 유의하자.

$$y_{t-1} = \alpha_{21}x_{t-2} + \alpha_{22}y_{t-2} + w_{t-1} \tag{3}$$

그러나 다시 식 (3)의 우변항에서 우변의 x_{t-2} 는 관측이 가능하지만 y_{t-2} 는 관측이 불가능하다. 따라서 마지막으로 관측 불가능한 변수 y_{t-2} 을 식 (1)과 두 번

시차를 가진 식을 이용하여 다음과 같이 쓰기로 하자.

$$y_{t-2} = (\alpha_{21} \alpha_{22})z_{t-3} + w_{t-2} \quad (4)$$

여기서 식 (4)의 우변항의 자기회귀항 z_{t-3} 는 관측 가능하다.

따라서 추정방정식은 식 (4)에 표현된 변수 y_{t-2} 를 식 (3)에 대입하고 여기서 얻어진 변수 y_{t-1} 의 식을 최종적으로 식 (2)에 대입한 후 변수 z_t 를 다음과 같은 식으로 나타낼 때 얻어진다.

$$\begin{aligned} z_t = & \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} x_{t-1} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix} \alpha_{21} x_{t-2} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix} \alpha_{22} (\alpha_{21} \alpha_{22}) z_{t-3} \\ & + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix} \alpha_{22} w_{t-2} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix} w_{t-1} + u_t \end{aligned} \quad (5)$$

그런데 식 (5)는 오차항들을 제외하고는 양변의 자기회귀항들이 모두 관측 가능한 형태를 띠고 있다. 이와 동일한 방법으로 관측 불가능한 자기회귀 변수의 반복대체를 통하여 변수 x_{t-1} 과 x_{t-2} 도 관측시점 $t=0, 3, 6, 9, \dots$ 에 좌우변의 모든 자기회귀항들이 관측 가능한, 다음과 같은 추정방정식을 유도할 수 있다.

$$x_{t-1} = \alpha_{11}x_{t-2} + \alpha_{12}(\alpha_{21}\alpha_{22})z_{t-3} + \alpha_{12}w_{t-2} + v_{t-1} \quad (6)$$

$$x_{t-2} = (\alpha_{11}\alpha_{12})z_{t-3} + v_{t-2} \quad (7)$$

마지막으로 종속 및 설명변수가 모두 관측 가능한 식 (5), (6) 및 (7)을 모두 모아 놓은 추정방정식 체계는 다음과 같이

$$\begin{pmatrix} z_t \\ x_{t-1} \\ x_{t-2} \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12}\alpha_{21} & \alpha_{12}\alpha_{22}\alpha_{21} & \alpha_{12}\alpha_{22}^2 \\ \alpha_{21} & \alpha_{22}\alpha_{21} & \alpha_{22}^2\alpha_{21} & \alpha_{22}^3 \\ 0 & \alpha_{11} & \alpha_{12}\alpha_{21} & \alpha_{12}\alpha_{22} \\ 0 & 0 & \alpha_{11} & \alpha_{12} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ x_{t-2} \\ z_{t-3} \end{pmatrix} + \begin{pmatrix} \alpha_{12}\alpha_{22}w_{t-2} + \alpha_{12}w_{t-1} + v_t \\ \alpha_{22}^2w_{t-2} + \alpha_{22}w_{t-1} + w_t \\ \alpha_{12}w_{t-2} + v_{t-1} \\ v_{t-2} \end{pmatrix}$$

또는

$$Z_t = B(A)X_t + U_t \tag{8}$$

로 나타낼 수 있다. 여기서 식 (8)은 종속변수 Z_t 와 설명변수 X_t 가 모두 관측 가능하므로 계수 $B(A)$ 의 추정을 위한 회귀분석을 수행할 수 있다.

그런데 여기서 중요한 점은 추정방정식의 설정이 추정방정식의 계수가 자기회귀계수 A 를 모두 포함하도록 이루어져야 한다는 것이다. 가령 위의 식 (7)은 설명변수와 종속변수가 모두 관측가능 하지만 계수가 모형의 자기회귀계수 A 를 모두 포함하고 있지 못하다. 따라서 식 (7)만을 사용하는 추정방정식 구성은 모형을 식별하지 못하게 된다. 그러나 위의 식 (8)은 자기회귀계수 A 를 모두 포함하고 있어 이와 같은 문제점이 발생하지 않는다.⁴⁾

한편 오차항 U_t 의 분산 $\Omega(=\omega_{ij})$ 는 다음과 같이 A 와 Σ 의 함수로 주어진다.

$$\begin{aligned} \omega_{11} &= \alpha_{12}\alpha_{22}\Sigma_{ww}\alpha_{22}'\alpha_{12}' + \alpha_{12}\Sigma_{ww}\alpha_{12}' + \Sigma_{vv} \\ \omega_{12} &= \alpha_{12}\alpha_{22}\Sigma_{ww}\alpha_{22}' + \alpha_{12}\Sigma_{ww}\alpha_{22}' + \Sigma_{vw} \\ \omega_{13} &= \alpha_{12}\alpha_{22}\Sigma_{ww}\alpha_{12}' + \alpha_{12}\Sigma_{wv}, \\ \omega_{14} &= \alpha_{12}\alpha_{22}\Sigma_{wv} \\ \omega_{22} &= \alpha_{22}^2\Sigma_{ww}\alpha_{22}' + \alpha_{22}\Sigma_{ww}\alpha_{22}' + \Sigma_{ww} \\ \omega_{23} &= \alpha_{22}^2\Sigma_{ww}\alpha_{12}' + \alpha_{22}\Sigma_{wv}, \\ \omega_{24} &= \alpha_{22}^2\Sigma_{wv} \\ \omega_{33} &= \alpha_{12}\Sigma_{ww}\alpha_{12}' + \Sigma_{vv}, \\ \omega_{34} &= \alpha_{12}\Sigma_{wv}, \\ \omega_{44} &= \Sigma_{vv} \end{aligned}$$

여기서 위 식 (8)은 모든 $t=0, 3, 6, 9, \dots$ 에 대하여 동일한 설명변수 Z_t 와 X_t 및 계수 $B(A)$ 를 유지하게 되는데, 이는 모형 전체 변수들의 관측주기의 최소공배

4) 추정치의 분산 최소화 등을 위한 최적의 추정방정식 구성 문제는 이론적으로 추가연구가 필요한 부분이다. 이 경우 표본 수와 표본 관측주기 등을 모두 고려하여야 할 것으로 판단된다.

수가 3이기 때문에 가능하다. 그런데 위의 식 (8)에서 설명변수 X_{t-1} 는 이동평균형의 오차항 U_t 와 상관관계를 갖고 있어 축약형 계수 $B(A)$ 는 통상최소자승(ordinary least square, OLS) 추정으로는 일치 추정량을 얻을 수 없다. 그러나 오차항 U_t 와 상관관계를 갖지 않지만 설명변수 X_t 와는 상관관계를 갖는 도구변수 H_t ⁵⁾의 존재를 가정하면 도구변수(instrumental variable) 추정법으로 모형 계수 $B(A)$ 의 일치 추정량을 얻을 수 있다.⁶⁾ 이를 감안하여 식 (8)에서 종속, 설명 및 오차항에서 모두 n 개의 자료가 집적된 행렬을 다음과 같이 각각 나타내기로 한다.

$$\begin{aligned} Z &= (Z_3, Z_6, \dots, Z_{3n})', \\ X &= (X_3, X_6, \dots, X_{3n})', \\ U &= (U_3, U_6, \dots, U_{3n})' \end{aligned}$$

다음은 Hansen(1982)의 최적 GMM(generalized method of moment) 추정을 수행하기 위하여 다음의 점근적 정규분포를 가정하기로 하자.

$$n^{-1/2} \text{vec}(H' U) \rightarrow_d N(0, S) \quad (9)$$

일반적인 경우⁷⁾ 위의 점근분포 (9)에서 분산행렬은 $S = (\Omega \otimes E H_t H_t')^{-1}$ 로 주어지게 된다. 다음은 식별을 위한 차수조건(order condition)을 만족시키기 위해 설명변수 (X_t)의 차원보다 많은 숫자의 VAR 모형 시차변수로 이루어진 도구변수자료 행렬을

$$H = (H_1', H_2', \dots, H_n')'$$

5) 여기서 시차변수들로 이루어진 도구변수의 예로 $H_t = (z'_{t-3}, x'_{t-4}, x'_{t-5}, z'_{t-6})'$ 를 들 수 있다.

6) 필자들의 모의실험 결과에 따르면 이 방법은 모형의 비정상성(nonstationarity) 하에서도 기존의 Chen and Zadrozny(1998)의 접근법과 달리 일치성을 보였다.

7) 가령 $\{\text{vec}(H_t U_t')\} \ t = 3, 6, \dots, n$ 이 마팅게일차분열(martingale difference sequence)일 때가 그 예이다.

와 같이 정의한다. 이 도구변수들로 계수 $B(A)$ 의 도구변수 추정치 $\hat{B}(A)$ 을 구한 후, 위의 추정방정식의 오차항 U 를 $\hat{U}=Z-\hat{B}(A)X$ 와 같이 정의하면, 행렬 S 의 일치추정량은 다음과 같이 나타낼 수 있다.

$$\hat{S}=\left(\frac{\hat{U}'\hat{U}}{n}\otimes\frac{H'H}{n}\right)^{-1}$$

마지막으로 계수 A 의 최적 GMM 추정치는 다음의 최소화 문제를 푸는 해 (solution)로서 주어지게 된다.

$$\min_{A\in\Theta}\text{vec}(H'U)'\hat{S}^{-1}\text{vec}(H'U) \tag{10}$$

다음으로 추정방정식 (8)의 오차항 U_t 의 분산을 나타내는 벡터 $\varpi=\text{vech}[\Omega(A,\Sigma)]$ 의 일치추정치는 오차항의 추정치⁸⁾ $\hat{U}=Z-B(\hat{A})X$ 로부터

$$\hat{\varpi}\equiv\text{vech}(\hat{U}'\hat{U}/n)$$

와 같이 구해진다.

이때 오차항의 분산 $\omega=\text{vech}(\Sigma)$ 의 최적 GMM 추정치는 VAR 계수 A 를 위의 GMM 추정치 \hat{A} 로 고정한 상태에서 다음의 점근적 분포를 가정하면서

$$n^{1/2}(\hat{\varpi}-\varpi(\omega;\hat{A}))\rightarrow_d N(0,\Psi)$$

추정식 (8)의 오차분산항 ϖ 와 그의 추정치 $\hat{\varpi}$ 의 거리에 대한 다음의 최소화 문제를 풀음으로써 얻어진다.

$$\min_{\omega\in\Theta}=\{\hat{\varpi}-\varpi(\omega;\hat{A})\}'\hat{\Psi}^{-1}\{\hat{\varpi}-\varpi(\omega;\hat{A})\}$$

8) 여기서 모수 A 의 추정치 \hat{A} 는 문제 (10)을 푼 GMM 추정치로 정의된다.

여기서 $\hat{\psi}$ 는 분산 ψ 의 일치추정량이며 이를 이용하여 최적 GMM 추정치를 얻을 수 있다.

2. 최적예측식의 유도 및 추정

혼합주기자료 하에서 직접 월별모형 (1)을 통해 최적예측을 위한 조건부 기대를 얻는 것은 월별로 관측되지 않는 자료들 때문에 불가능하다. 가령, $t=0, 3, 6, 9, \dots, n$ 일 때 $t-1$ 까지의 관측가능한 변수들의 집합 $d_{t-1} = (x_{t-1}, x_{t-2}, z_{t-3})$ 을 이용하여 z_t 를 예측하는 것을 상정하여 보자.⁹⁾ 이 경우 조건부 기대는 정의상 식 (2)에서

$$E(z_t | d_{t-1}) = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} x_{t-1} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix} E(y_{t-1} | d_{t-1}) \quad (11)$$

으로 주어지게 된다.

여기서 정상인 모형 (1)로부터 비조건부기대(unconditional expectation)가 $E(d_{t-1}) = 0$ 으로 주어지므로, 일반적인 정규성(normality) 가정 하에서 다음 식이 성립한다.

$$E(y_{t-1} | d_{t-1}) = cov(y_{t-1}, d_{t-1}) var(d_{t-1})^{-1} d_{t-1} \quad (12)$$

식 (12)에서 cov 는 두변수간의 공분산, var 는 개별변수의 분산을 나타낸다.

먼저 문제를 단순화시켜 이해하기 위하여 모수 A 와 Σ 는 별도의 언급이 없으면 주어진 것으로 가정하자. 다음으로 만일 위 식 (12)의 분산 $var(d_{t-1})$ 와 공분산 $cov(y_{t-1}, d_{t-1})$ 의 표본적률을 구할 수 있다면 일치추정된 자기회귀 계수로 조건부 기대 (12)를 최종적으로 일치추정할 수 있다.

이를 위하여 몇 가지 방안을 고려할 수 있다. 첫째, Chen and Zadrozny 에서와

9) 통상의 분기모형을 통한 예측의 경우 d_{t-1} 의 원소 중 관측가능한 분기자료인 z_{t-3} 만을 이용하는 점에 유의하자.

같이 시차 표본적률을 이용하여 직접 이 공분산을 추정하는 것은 불가능하다. 왜냐하면 이와 같은 방법을 사용해서는 $cov(y_{t-1}, d_{t-1})$ 를 구성하는 공분산 항들을 일부 밖에 추정할 수 없는데, 가령 $cov(y_{t-1}, x_{t-1})$ 이나 $cov(y_{t-1}, x_{t-2})$ 는

$$n^{-1} \sum_{t=3,6,9,\dots,3n} y_{t-3} x_{t-3}$$

또는

$$n^{-1} \sum_{t=3,6,9,\dots,3n} y_{t-3} x_{t-4}$$

와 같은 방식의 관측가능한 변수들의 표본평균으로 각각 추정할 수 있어도, ¹⁰⁾ $cov(y_{t-1}, y_{t-3})$ 과 같은 2개월 시차의 것은 <표 1>과 같은 관측례에서 표본평균으로 추정할 수 없다. 이는 GDP가 분기 (3개월) 단위로 관측되기 때문에 나타나는 결과이다.

두 번째로 고려할 수 있는 방법은 Yule-Walker 방정식을 이용하여 z_t 의 자기공분산행렬 (autocovariance matrices) 들을 자기회귀계수 A 와 오차분산 Σ 의 함수로 표현하여 이를 추정하는 방법이다. 가령 모형 (1)에서

$$\text{vec}\{var(z_t)\} = (I - A \otimes A)^{-1} \text{vec}(\Sigma)$$

와 같이 나타낼 수 있으며, 다른 시차의 자기공분산행렬[예: $cov(z_t, z_{t-k})'$, $k=1, 2, 3, \dots$]도 같은 방법으로 나타낼 수 있다. 이 경우 Chen and Zadrozny의 표본평균으로도 추정할 수 없었던 공분산 $cov(y_{t-1}, y_{t-3})$ 도 A 와 Σ 의 함수로 표현한 후, A 와 Σ 의 일치추정량을 대입하면 이 공분산의 일치추정량을 구할 수 있다. 물론 이 방법도 분산 Σ 의 일치추정량을 구할 수 없는 Chen and Zadrozny의 접근법에서는 이용할 수 없다.

세번째 방법은 식 (11)의 조건부 기대항 $E(y_{t-1}|d_{t-1})$ 를 식 (12)와 같이 직접

10) 시계열 $(y_{t-3})_{t=0,3,6,9,\dots}$ 는 관측가능하다는 점에 유의하자.

추정하는 대신 우선 y_{t-1} 를 앞서의 추정방정식에서와 같이 관측가능한 자기회귀변수와 오차항으로 표현한 후 추정하는 방식이다. 이를 위하여 변수 y_{t-1} 를 자기회귀 변수항과 오차항으로 나누는 다음과 같은 두 개의 항

$$p_t \equiv \alpha_{21}x_{t-2} + \alpha_{22}(\alpha_{21}, \alpha_{22})z_{t-3},$$

$$e_t \equiv \alpha_{22}w_{t-2} + w_{t-1}$$

로 정의하자. 이 경우 식 (1)에서 $y_{t-1} = p_t + e_t$ 로 주어지는 것을 쉽게 확인할 수 있으며, y_{t-1} 의 조건부 기대는 정규성 가정하에서

$$E(y_{t-1}|d_{t-1}) = p_t + cov(e_t, d_{t-1})var(d_{t-1})d_{t-1} \quad (13)$$

와 같이 정의된다. 여기서 식 (13)의 우변항을 구성하고 있는 공분산 $cov(e_t, d_{t-1})$ 은 자기회귀계수 A 와 오차분산 Σ 의 함수로 나타낼 수 있다. 따라서 A 와 오차분산 Σ 의 추정치를 대입할 경우 $cov(e_t, d_{t-1})$ 의 추정치는 $\widehat{cov}(e_t, d_{t-1})$ 로 유도된다. 또한 분산항 $var(d_{t-1})$ 은 관측가능한 변수 d_{t-1} 의 표본 2차적률 $\widehat{var}(d_{t-1})$ 로 구해지므로 식 (13)에서 주어지는 조건부 기대치 $E(y_{t-1}|d_{t-1})$ 의 추정치는 다음으로 구해진다.

$$\widehat{E}(y_{t-1}|d_{t-1}) = \widehat{p} + \widehat{cov}(e_t, d_{t-1}) \cdot \widehat{var}(d_{t-1}) \cdot d_{t-1} \quad (14)$$

여기서 최적예측치 (11)의 추정은 자기회귀계수 A 의 추정치와 식 (14)의 추정치를 대입하여 구할 수 있다.

한편 식 (13)에서 조건부 기대치 $E(y_{t-1}|d_{t-1})$ 를 변수 p_t 로 근사하여 최적 예측치 (11)을 다음과 같이 근사하는 경우를 상정하여 보자.

$$E(z_t|d_{t-1}) \simeq \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} x_{t-1} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix} p_t \quad (15)$$

이와 같은 최적예측식의 근사치를 사용하더라도 만일

$$var(p_t|d_{t-1}) \geq -2cov(p_t, e_t|d_{t-1}) \tag{16}$$

라는 조건을 만족할 경우 변수 p_t 를 아주 예측식에서 제외하는 경우에 비해 예측 오차 분산이 줄어들게 됨을 쉽게 보일 수 있다. 그런데 위의 부등 조건 (16) 은 일반적으로 모형이 비정상에 가까울수록 쉽게 만족되는 데 가령 $var(p_t|d_{t-1})$ 의 표본적치를 추정치는 일반적으로 $O_p(n)$ 인 변수로, $cov(p_t, e_t|d_{t-1})$ 는 $O_p(1)$ 인 변수로 근사되기 때문이다. 결론적으로 이 경우 관측가능한 자기회귀항을 제외한 나머지 부분만이 예측오차로 편입됨으로서 통상 예측효율성 제고에 기여하게 된다.

3. 분기모형 대비 예측효율성 비교

한편 위의 조건부 기대치 식 (11)의 경우 정보집합 d_{t-1} 가 분기자료 (z_{t-3}) 만이 아니라 분기내의 월별자료 (x_{t-1}, x_{t-2}) 까지 같이 사용하고 있으므로 분기자료만을 사용하는 통상적인 분기 VAR 모형을 통한 사영예측(projection) 보다 예측평균자승오차(forecasting MSE, FMSE)가 작아지게 된다.

그 차이를 구하기 위하여 일반적인 정규성(normality) 가정 하에서 위의 조건부 기대치 식 (11)이 자료 d_{t-1} 를 이용한 사영예측과 같다는 사실과, 사영예측을 위한 계수는 OLS 추정을 통하여 일치추정량을 얻을 수 있다는 사실을 이용한다. 다음은 $d_{t-1} = (q_{t-1}, m_{t-1})$ 는 분기자료 q_{t-1} 와 월별자료 m_{t-1} 로 구성된 관측가능한 자기회귀항들의 벡터로 정의하기로 한다. 여기서 $M_{ab} \equiv n^{-1}ab'$ 를 임의의 열벡터 변수 a 와 b 의 표본 공분산으로 정의하자.¹¹⁾

그러면 분기 및 월별자료를 모두 사용한 z_t 의 예측식 $f_{t|t+1} \equiv M_{zd}M_{dd}^{-1}d_{t-1}$ 의 FMSE는

11) 여기서 숫자 n 은 개략적으로 ‘자료주기의 최소공배수(LCM)인 3으로 전체 표본의 숫자를 나눈 것’이다. 이는 위에서 유도한 예측방정식의 계수와 자기회귀항이 주기 3을 갖고 동일한 형태를 갖게 되는 사실을 반영한다.

$$V_f \equiv M_{zz} - M_{zd}M_{dd}^{-1}M_{zd}'$$

로 주어지고, 분기 자료만을 쓴 z_t 의 사영예측 $g_{t|t+1} \equiv M_{zq}M_{qq}^{-1}q_t$ 은

$$V_g \equiv M_{zz} - M_{zq}M_{qq}^{-1}M_{zq}'$$

로 정의되는 FMSE를 갖는다. 만일 $M_{mm}^* \equiv M_{mm} - M_{qm}'M_{qq}^{-1}M_{qm} > 0$ 의 조건이 성립한다고 가정하면 두 예측식 $g_{t|t+1}$ 와 $f_{t|t+1}$ 의 FMSE 차이는 다음과 같이 유도될 수 있음을 보일 수 있다.

$$V_g - V_f = M_{zm}^*'(M_{mm}^*)^{-1}M_{zm}^* \geq 0 \quad (17)$$

여기서 $M_{zm}^* \equiv M_{zm} - M_{qz}'M_{qq}^{-1}M_{qm}$ 로 정의¹²⁾된다.

한편 월별모형이 비정상적일 경우 분기모형을 이용한 예측에서 누락된 분기 내 월별자료 역시 비정상성을 가지게 되므로 모형이 정상적인 경우 보다 식 (17)에서 주어진 FMSE 차이는 확대될 것으로 예상된다. 실제로 Kim and Park의 비정상성을 갖는 혼합주기 VAR모형의 점근적 분포이론을 이용하여 위의 식 (17)에서 주어진 차이가 $O_p(n)$ 의 변수임을 쉽게 보일 수 있다. 좀 더 구체적으로 식 (12)에서 FMSE들의 차이를 구성하는 항들인 M_{zm}^* 와 M_{mm}^* 가 모두 $O_p(n)$ 이기 때문에 $V_g - V_f$ 역시 $O_p(n)$ 인 변수가 된다.

여기서 d_{t-1} 를 이용한 최적예측의 FMSE인 V_f 가 $O_p(1)$ 이라는 점을 감안하면, V_g 가 누락된 분기내 월별자료 때문에 $O_p(n)$ 인 변수가 된다고 볼 수 있다.¹³⁾ 한편

12) 여기서 $M_{zm}^* = 0$ 인 경우 양자의 차이($V_g - V_f$)는 사라지게 되는 데, 이는 분기내 월별자료를 분기자료에 회귀시킨 잔차($m - qM_{qq}^{-1}M_{qm}$)가 예측대상변수인 z 와 직교관계라는 사실을 의미한다. 이는 다시 분기내 월별자료(m)은 분기자료(q) 이상으로 예측대상변수에 대한 정보를 갖고 있지 않음을 의미한다. 이러한 사실에 기초하여 우리는 예측대상변수를 분기 및 분기내 월별자료에 회귀시켰을 때 분기내 월별자료에 대한 F 검정이 분기내 월별자료의 예측에 대한 기여도를 측정하는 기준이 될 수 있음을 알 수 있다.

13) 물론 여기서 분기 내 월별자료가 분기자료에 의해 완벽하게 설명(perfect-fitting) 되지 않는다

V_f 가 $O_p(1)$ 이라는 것은 다음과 같이 간단하게 보일 수 있다. 먼저 사영의 성질을 이용하여 V_f 를 다음과 같이 쓴다.

$$V_f = M_{\xi\xi} - M_{\xi d} M_{dd}^{-1} M_{\xi d}' \quad (18)$$

여기서 변수 ξ 는 예측방정식의 오차항 인데, 예를 들어 $t-1$ 기에 식 (5)를 변수 z_t 의 예측식으로 쓸 경우 이는 다음과 같이 정의된다.

$$\xi \equiv \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix} \alpha_{22} w_{t-2} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix} w_{t-1} + u_t$$

그런데 식 (18)의 우변에서 각각 $M_{\xi\xi}$ 는 $O_p(1)$, $M_{\xi d}$ 는 $O_p(1)$, M_{dd} 는 $O_p(n)$ 이므로 V_f 는 $O_p(1)$ 변수임이 자명하다.¹⁴⁾ 한편 이와 같은 논거에 따르면 예측식의 예측효율성 증가 정도는 분기모형에 추가된 월별자료의 유의성에 대한 F 검정을 통하여 이루어 질 수 있다.¹⁵⁾

다음 장에서는 위에서 소개한 이론적 배경을 바탕으로 우리나라 GDP 성장률을 금융변수를 포함한 3-변수 VAR모형을 이용하여 예측하고자 한다. 이 경우 변수설정은 Friedman and Kuttner (1992, 1993)을 따르되 원/달러 환율을 추가하는 등 우리나라 현실에 맞게 일부 조정하였다.

III. 실증분석

1. 금융변수와 GDP성장률의 상관계수 비교

이번 절에서는 먼저 모형에 사용한 변수들의 GDP성장률 대비 상관계수 분석을

는 추가 전제조건이 필요하다.

- 14) 모의실험 결과도 분기모형과 월별 및 분기자료를 모두 사용한 예측식의 예측오차의 차이는 자료가 비정상적일수록 커지는 양상을 보였다.
- 15) 모의실험 결과에 따르면 월별자료의 높은 F 검정통계량은 낮은 예측오차와 밀접하게 관련되어 있다.

하고자 한다. 이는 다음 절들에서 제시될 GDP성장률 예측결과를 이해하는데 도움이 될 것이다. 먼저 사용된 자료를 설명하면, 모형 추정에 사용된 실물경제 변수로는 분기별 GDP 및 월별 소비자물가지수를 각각 로그차분하여 사용하였다. 금융변수로는 월별 신용(credit) 스프레드(3년 만기 회사채수익률 - 5년 만기 국민주택채권 수익률),¹⁶⁾ 기간(term) 스프레드(5년 만기 국민주택채권 수익률 - 1년 만기 통안증권 수익률), 로그차분한 본원통화, M1, M2, M3와 콜금리 및 원/달러 환율을 사용하였다. 분석기간은 금리자료가 최대이용 가능한 1987. 6월 ~ 2005. 12월까지 하였다. 이들 실물·금융변수들에 대한 단위근검정(ADF, Phillips and Perron 및 KPSS)을 실시한 결과 모두 5% 유의수준에서 단위근이 없는 것으로 나타났다.

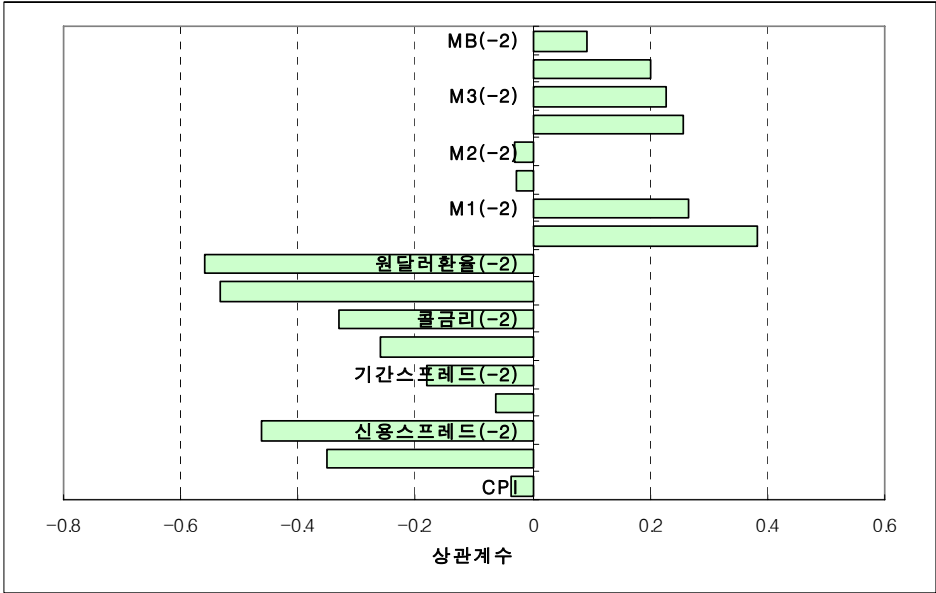
다음으로 금융변수들과 GDP성장률간 시차별 교차상관계수를 살펴본 결과, 〈표 2〉에서 보듯 GDP 성장률은 대체로 통화관련 변수와는 양(+)의 값을, 환율 및 금리와는 부(-)의 값을 보이는 것으로 나타났다. 여기서 상관계수의 크기는 환율, 신용 스프레드, M1 등의 순서로 나타났다. 반면 콜금리, 금리스프레드 및 환율 등은 M1, M3 및 본원통화와는 반대로 1개월 시차에 비해 2개월 시차의 상관관계가 더 높은 것으로 나타났다.

한편 GDP 성장률과 분기 주기로 전환한 1개월 시차 금융변수들을 붙임 〈참고〉와 같은 도표로 단순 비교한 결과 다음과 같은 특징들이 나타났다. 먼저, 외환위기 이후 금융변수들이 GDP 성장률에 대체로 선행하고 있는 것으로 보인다. 특히 금리의 기간 스프레드 및 M2 및 M3와 같은 통화량 등에서 이런 경향이 뚜렷함을 알 수 있다. 이로써 월별 금융변수들이 GDP 성장률에 대한 단기 예측 정보를 어느 정도 가지고 있을 것으로 추측된다.

그러나 위와 같은 단순 시차상관관계 또는 도표를 이용한 특징 도출은 GDP성장률의 예측에 금융변수들이 어떤 역할을 하는지에 대한 엄밀한 통계적 분석이라고 할 수 없다. 따라서 다음 절에서는 표본 내 및 표본 외 예측식을 VAR 모형내에서 도입하여 분기자료에 추가된 월별 금융변수가 예측력 제고에 기여하는지를 구체적으로 살펴보려고 한다.

16) 자료수의 부족으로 만기가 다른 3년 만기 국고채수익률을 이용하였다.

〈표 2〉 금융변수의 GDP 성장률과의 상관계수



주: ()내는 금융변수와 GDP성장률 간의 시차임.

2. 예측결과

(1) 표본내 예측

먼저 1987.6 ~ 2005. 12월중 분기자료를 이용하여 다음의 분기 회귀식을 최소자승법(OLS)으로 추정하였다. 회귀식의 시차는 Friedman and Kuttner(1992, 1993)를 따라 4로 설정하였다. 이는 신용스프레드 등 금융변수가 대부분 2분기 시차 내에서 실물경제에 미치는 영향력을 모두 가진다는 기존 연구결과¹⁷⁾를 감안한 것이다.¹⁸⁾

$$\Delta X_t = \gamma + \sum_{i=1}^4 \beta_{1i} \Delta X_{t-i} + \sum_{i=1}^4 \beta_{2i} \Delta P_{t-i} + \sum_{i=1}^4 \beta_{3i} \Delta Q_{t-i} + \sum_{i=1}^4 \beta_{4i} \Delta IQ_{t-i}^1 + \sum_{i=1}^4 \beta_{5i} \Delta IQ_{t-i}^2 + u_t$$

17) Estrella and Mishkin(1998)를 참조.
18) AIC 또는 BIC 등의 모형설정 기준에 의해서도 여기에서 크게 벗어나지 않았다.

여기서 X 는 실질 GDP의 자연로그, P 는 CPI의 자연로그, Q 는 금융변수의 분기 자료를 나타낸다. 또 IQ^1 와 IQ^2 는 Q 의 해당 분기내 추가된 월별 금융변수(2개)로 전분기내 월별자료와 각각 차분(3개월)한 자료(ΔIQ^1 와 ΔIQ^2)가 이용되었다.¹⁹⁾

다음으로 아래 <표 3>의 GDP 성장률에 대한 회귀분석 결과²⁰⁾를 보면 금융변수의 전반적인 설명력(in sample fitting)을 나타내는 F 값을 기준²¹⁾으로, 분기 자료를 쓸 경우 신용스프레드의 유의성이 가장 컸다. 이는 Friedman and Kuttner (1992, 1993) 등이 미국 자료를 대상으로 수행한 연구결과와도 일치하는 것이다. 그 밖에 원/달러 환율, 회사채수익률, 본원통화 등의 통계적 유의성도 비교적 높은 것으로 나타났다. 한편 본원통화를 제외한 여타 통화지표들은 GDP성장률의 설명력이 거의 없는 것으로 나타났다.

<표 3> 금융변수에 대한 F 통계량

| 금융변수 | 분기자료 | 분기+분기내 월별자료 | 분기내 월별자료 |
|------------------|----------|----------------|----------|
| | | | |
| 신용스프레드 | 10.49*** | 1.97** | 1.48 |
| 기간스프레드 | 0.24 | 0.80 | 2.58** |
| M1 | 0.98 | 0.84 | 2.64** |
| M2 | 1.40 | 1.04 | 1.67 |
| M3 | 1.55 | 0.89 | 1.14 |
| 본원통화 | 4.52*** | 2.33** | 0.94 |
| 국민주택채권 (5년만기) | 1.32 | 1.80* | 2.58** |
| 회사채수익률 (3년만기) | 6.28*** | 1.68* | 3.06*** |
| 원/달러 환율 | 7.51*** | 2.32** | 1.24 |

주: ***는 1%, **는 5%, *는 10% 유의수준에서 각각 유의함을 의미.

19) 예컨대 2분기내의 4월, 5월의 금융데이터는 1분기내 2월, 3월 데이터에서 각각 차분하였다.

20) 외환위기 이후 통화정책 변화 등으로 구조변동의 가능성이 큰 콜금리는 분석대상에서 제외하였다.

21) 귀무가설(H_0)은 모든 i 에 대하여 $\beta_{3i} = \beta_{4i} = \beta_{5i} = 0$ 임.

그러나 국민주택채권, M1 및 기간스프레드 등은 분기자료 만을 썼을 경우에는 유의성이 없었으나, 추가된 분기내 월별 자료를 이용했을 경우 유의성이 높아지는 것으로 나타났다.

(2) 표본 외 예측

표본 외 예측을 위한 예측식은 예측치의 간결성이 중요하다는 기존의 연구결과〔예: Clements and Hendry(1998)〕를 반영하여 상수항을 포함한 식 (11)의 조건부 기대치와 이의 근사치인 식 (15)의 다음과 같은 추정식 두 가지 형태를 모두 사용하였다.²²⁾

$$\hat{E}(z_t|d_{t-1}) = \begin{pmatrix} \hat{\alpha}_{11} \\ \hat{\alpha}_{21} \end{pmatrix} x_{t-1} + \begin{pmatrix} \hat{\alpha}_{12} \\ \hat{\alpha}_{22} \end{pmatrix} \hat{p}_t \quad (19)$$

모형의 시차는 예측식을 이용한 F 검정²³⁾과 AIC 또는 BIC 기준을 병행 사용하여 확정하였는데, 신용스프레드, 국민주택채권 및 회사채수익률 등의 일부변수를 제외하고는 모두 식 (1)과 같은 월별 VAR(1) 형태로 근사할 수 있는 것으로 나타났다.²⁴⁾ 이와 같은 결과는 기존의 이론적인 성과들²⁵⁾과 배치되지 않는 것이다. 이에 따라 최적예측식은 간결한 예측식의 중요성을 감안 월별 VAR(1) 모형에서 유도되는 것으로 가정하였다.

22) 이 근사치는 식 (13)의 예측 방정식에서 자기회귀항(d_t) 들에 의해 설명되지 않는 오차항 e_t 이 그리 크지 않을 경우, 식 (13)의 최적예측식의 $cov(e_t, d_t)var(d_t)d_t$ 부분의 추정오차를 줄여 오히려 예측정확성을 제고할 수도 있다.

23) 가령 VAR(2) 모형은 VAR(1) 모형보다 예측식에서 더 많은 시차설명변수를 갖게 된다. 만일 VAR(1) 모형과 비교할 때 VAR(2) 모형에 추가된 설명변수에 대한 F 통계량이 유의하지 않을 경우에는 VAR(1) 모형을 선택하게 된다.

24) 이와 같은 VAR(1) 모형근사 결과는 미국자료를 이용할 경우에도 유사하게 나타났다. 물론 VAR(p), $p > 1$ 인 경우의 예측 성과 비교는 추후의 중요한 연구과제가 될 것이다.

25) 많은 경제변수들이 임의보행(random walk) 또는 AR(1) 모형을 따른다는 것은 잘 알려져 있다. 가령 Fama(1965)의 주가, Hall(1978)의 소비, Barro(1979)와 Mankiw(1987)의 조세와 통화발행 및 Diebold and Nason(1990)은 환율 등에 대한 분석결과가 그 예이다. Litterman(1986) 등은 이와 같은 인식에 기초하여 VAR(1) 모형을 기초로 한 베이지안 VAR 예측모형을 구축하고 있다.

여기서 예측식의 계수 추정은 GMM 추정 이외에 비교를 위하여 OLS 추정량과 각각 1, 2 및 3분기의 시차를 가진 분기모형들인 VAR(3), VAR(6) 및 VAR(9) 과 예측방법을 평가하는 연구들에서 기준점(bench mark)으로 이용하는 AR(1) 모형을 통한 예측결과를 함께 비교하여 보았다.

각 회귀식의 표본기간은 우선 외환위기 이후 기간인 1997. 10~2005. 12월중(총 33분기)의 자료를 이용하여 위의 표본 외 예측식을 추정한 후, 분석기간을 1분기씩 뒤로 1987년 2/4 분기²⁶⁾까지 이동하면서 추정하는 축차 회귀분석(rolling regression)을 통하여 모두 32회의 예측을 시행하였다.²⁷⁾ 한편 예측능력의 비교 기준으로 1분기 앞의 예측에 대한 평균자승오차와 중위수 오차를 사용하였는데, 위의 축차 회귀분석 결과가 이들 통계치 산출의 표본으로 사용되었다.

추정 결과를 나타낸 아래의 <표 4>를 보면 분기 내 월별 금융자료를 추가로 이용한 예측식 (19)와 같은 근사치²⁸⁾는 분기자료만을 쓴 VAR모형이나 단일변수 AR(1) 모형보다 예측성과(평균자승오차 및 중위수 오차)가 대체로 우수한 것으로 나타났다.²⁹⁾

특히 M1 또는 신용스프레드를 포함한 근사 최적예측식을 이용한 경우 여타 모형에 비해 상대적으로 우월한 예측력을 보였다. 또한 신용 스프레드도 추정 방법에 관계없이 낮은 평균자승오차를 보였는데,³⁰⁾ 이는 신용스프레드의 구성요소인 회사채 및 주택채권수익률을 개별적으로 사용할 때는 나타나지 않는 현상으로 주목의 여지가 있다. 다만 이는 M1을 이용한 경우에 비해 대체로 높은 중위수 오차를 보

26) 콜금리를 제외한 여타 이자율의 이용 가능한 기간은 모두 228개월(76분기)이다.

27) 이와 같은 축차회귀 방식은 Thoma and Gray(1998)가 Friedman and Kuttner(1993)의 추정 결과의 강건성(robustness)을 확인하기 위하여 사용한 방법이다.

28) 이론적으로는 조건부 기대치인 식 (11)을 사용하는 것이 최적이지만 실제 추정결과 자기회귀 항만을 예측에 사용한 근사치가 평균자승 오차 및 중위수 오차에서 우수하였는데, 이는 조건부 기대치 계수의 추가 추정에 따른 오차증가가 이론적인 평균자승오차 감소를 넘어서고 있는 데서 비롯된 것으로 보인다.

29) 물론 이러한 예측의 평균자승오차나 중위수 오차가 예측방법별로 차이가 있는지는 붓스트랩 등 통계적 검증을 통해 확인해 볼 수 있을 것이다. 그러나 이는 모형 소개에 많은 비중을 두고 있는 본고의 논의 범위를 지나치게 확장하는 것이어서 다음 연구 과제로 넘기고자 한다.

30) 신용스프레드의 표본내 실물경제의 예측능력은 일찍이 Friedman and Kuttner(1992, 1993), Estrella and Mishkin(1998) 등 다수의 연구에서 확인된 바 있으나 우리나라의 경우 특히 표본 외 예측에서 이를 확인한 것은 의미가 있는 것으로 판단된다.

이고 있다는 단점이 있다.³¹⁾

〈표 4〉 분기 GDP성장률의 예측평균자승오차

| 예측 방법 | | M11) | M2 | M3 | MB | 신용스프 레드 | 기간스프 레드 | 회사채수 익률 | 주택채권 수익률 | 월/달러 환율 |
|-----------------------------|-------------------------------------|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 혼합 주 기 V A R | GMM(-1) | .026 (-.014) | .033 (-.017) | .026 (-.014) | .036 (-.019) | .027 (-.024) | .026 (-.014) | .029 (-.013) | .025 (-.011) | .027 (-.018) |
| | GMM(-2) | .022 (-.015) | .031 (-.023) | .025 (-.015) | .040 (-.030) | .023 (-.021) | .023 (-.013) | .042 (-.014) | .023 (-.009) | .039 (-.017) |
| | GMM ^a (-1) ²⁾ | .018 (.000) | .023 (-.004) | .021 (-.002) | .027 (-.008) | .016 (-.008) | .022 (-.001) | .022 (.000) | .022 (-.001) | .024 (-.003) |
| | GMM ^a (-2) | .016 (-.001) | .023 (-.009) | .019 (-.004) | .029 (-.020) | .013 (-.008) | .020 (.000) | .027 (-.004) | .018 (.003) | .031 (-.003) |
| | OLS(-1) | .022 (-.005) | .026 (-.004) | .022 (.001) | .026 (-.003) | .022 (-.006) | .023 (-.003) | .039 (-.002) | .024 (-.002) | .033 (-.007) |
| | OLS(-2) | .024 (-.006) | .024 (-.002) | .022 (.001) | .020 (-.004) | .021 (-.007) | .022 (-.001) | .038 (-.004) | .022 (-.004) | .034 (-.007) |
| | F-검정 ³⁾ | 25.3 | 25.3 | 23.8 | 22.2 | 19.1 | 30.8 | 37.2 | 31.7 | 38.1 |
| | VAR(3) | .024 (-.006) | .021 (-.002) | .020 (-.000) | .019 (-.003) | .013 (-.007) | .019 (-.003) | .030 (-.003) | .021 (.004) | .031 (-.007) |
| | VAR(6) | .023 (.000) | .023 (-.002) | .021 (.003) | .022 (-.001) | .020 (-.003) | .022 (-.001) | .034 (-.006) | .020 (-.002) | .032 (-.002) |
| | VAR(9) | .021 (.001) | .023 (-.002) | .020 (.003) | .022 (-.002) | .017 (.000) | .022 (-.002) | .029 (-.006) | .019 (.000) | .036 (-.001) |
| AR(1) | | 0.020 (0.004) | | | | | | | | |

- 주 1) () 내는 표본내 중위수 오차를 나타냄.
- 2) GMM^a(-1) 및 OLS(-1)은 1개월 전, GMM^a(-2) 및 OLS(-2)은 2개월 전 예측을 나타내며 조 건부 기대치중 오차항과 관련된 부분을 제외한 근사치를 나타냄.
- 3) 표본의 중위수임. 5% 유의수준에서 F 검정은 3.84의 기각값(critical value)을 가짐.

그러나 본원통화를 사용할 경우는 근사 최적예측식의 성과가 분기VAR모형에 비

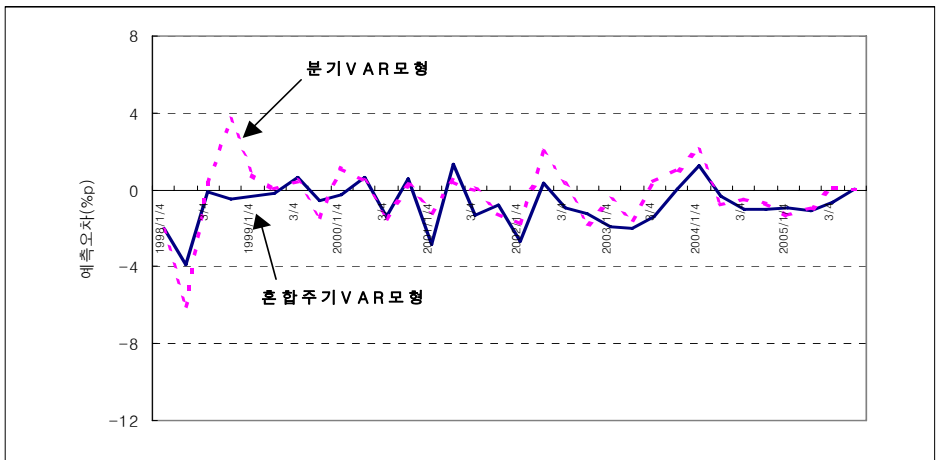
31) 다만 더 많은 변수를 예측에 사용한 GMM(-1) 또는 GMM^a(-1) 등을 사용한 1개월전 예측 성과가 GMM(-2) 또는 GMM^a(-2) 등을 사용한 2개월전 예측 성과보다 FMSE에서 반드시 우월한 결과를 나타내지는 않고 있는데, 이는 Clements and Hendry(1998) 등이 강조(“most forecasters seem to believe that parsimony is important” p. 280) 하고 있는 모형의 간결성 (parsimony)이 일부 반영된 것으로 판단된다.

해 미약한 것으로 나타났는데, 이는 본원통화의 변화가 일반적으로 단기보다는 6개월 이상의 시차를 가지고 실물경제에 영향을 미치기 때문인 것으로 보인다. 곧 1~2개월 전의 자료(본원통화) 추가가 예측에 기여하기 보다는 간접성을 훼손하는 잡음의 역할에 그쳤다는 것이다.³²⁾ 이와 함께 예측방정식에서 모든 금융변수가 (소비자물가지수와 같이) GDP성장률에 대하여 0의 계수를 갖는다는 귀무가설은 5% 유의수준에서 F 검정 결과 기각되었다.

결론적으로 우리나라의 금융변수를 이용한 단기 GDP성장률 예측은 M1또는 신용스프레드를 이용한 최적예측식의 근사치를 함께 고려하는 것이 하나의 방안인 것으로 보인다.

한편 아래의 <그림 1>과 <그림 2>에서는 금융변수로 M1을 사용하였을 때, 분기 및 혼합주기 모형들에서 가장 예측성고가 좋았던 VAR(9) 모형과 2개월전 근사 최적예측식 GMM^a(-2) 등 두 가지 예측방법의 분기 단위 예측치와 예측오차³³⁾를 그래프에 나타내고 있다.

<그림 1> 성장률 예측오차 추이 (M1의 경우)



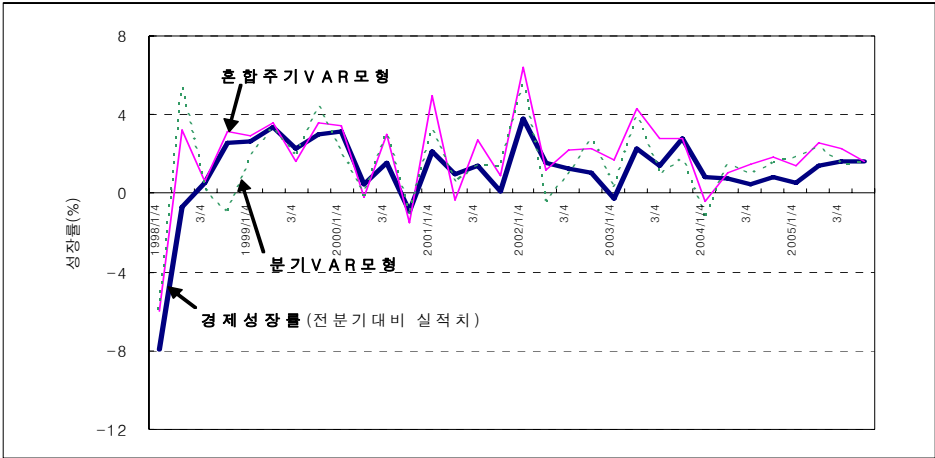
여기서 우리 경제가 불안정성을 보인 외환위기 직후의 기간(1998 ~ 1999년중)

32) 이와 같은 결과는 <표 3>에서도 간접적으로 확인할 수 있는데, 분기내 월별 M1 자료는 본원통화와는 반대로 F 통계량이 분기자료만 이용한 경우와 달리 유의성이 있는 것으로 나타났다.

33) 공표된 분기 GDP에 대한 예측을 의미한다.

동안 분기모형은 근사 최적예측식보다 상대적으로 큰 예측 오차를 나타내고 있는 것을 볼 수 있다. 이와 같은 결과는 II장 2절에서 언급한 바와 같이 ‘월별자료를 누락시킨 후 분기자료만으로 예측한 경우, 자료의 비정상성하에서 보다 더 큰 예측효율성 저하는 가져온다’ 는 이론적인 결과와 배치되지 않는다.

〈그림 2〉 성장률 실적치 및 예측치 추이 (M1의 경우)



IV. 맺음말

본고에서 우리는 좀 더 정확한 분기GDP성장률 예측을 위해서 분기별 GDP성장률외에 다른 월별 경제변수도 함께 이용할 수 있는 방법을 개발하였다. 이를 위하여 우선 GDP 성장률과 기타 관련된 경제변수 모두를 포함하는 월별 VAR(vector autoregressive) 모형을 상정한 후, 이로부터 오직 분기별로만 관측되는 GDP성장률을 위한 예측식을 유도하였다. 유도된 예측식은 유한한 시차를 가지는 이동평균 오차를 갖게 되며, 따라서 차분변수를 도구변수로 하는 GMM 방법으로 쉽게 추정할 수 있다. 제시된 방법이 현실적으로 잘 적용될 수 있는가 알아보기 위하여, 실제로 한국의 분기GDP성장률 예측을 시도해 보았다. 분기별 GDP성장률 외에 월별로 관측되는 소비자물가지수(CPI) 및 여타 금융변수 등을 고려한 바, 제 금융변수들 중 월별 통화량 M1과 회사채의 신용스프레드를 각각 추가 이용할 경우 기존의 분기VAR모형이나 단일변수 AR(autoregressive) 모형 등에 비해 예측평균자승오차

(forecasting MSE) 또는 중위수 오차(median bias)를 현저히 개선시킬 수 있는 것으로 나타났다.

이에 따라 GDP 성장률 예측시 본고와 같은 혼합주기자료를 이용한 VAR 모형 예측도 같이 고려할 경우, 최근의 금융 정보를 보다 일반적인 체계하에서 명시적으로 반영할 수 있을 것으로 보인다.

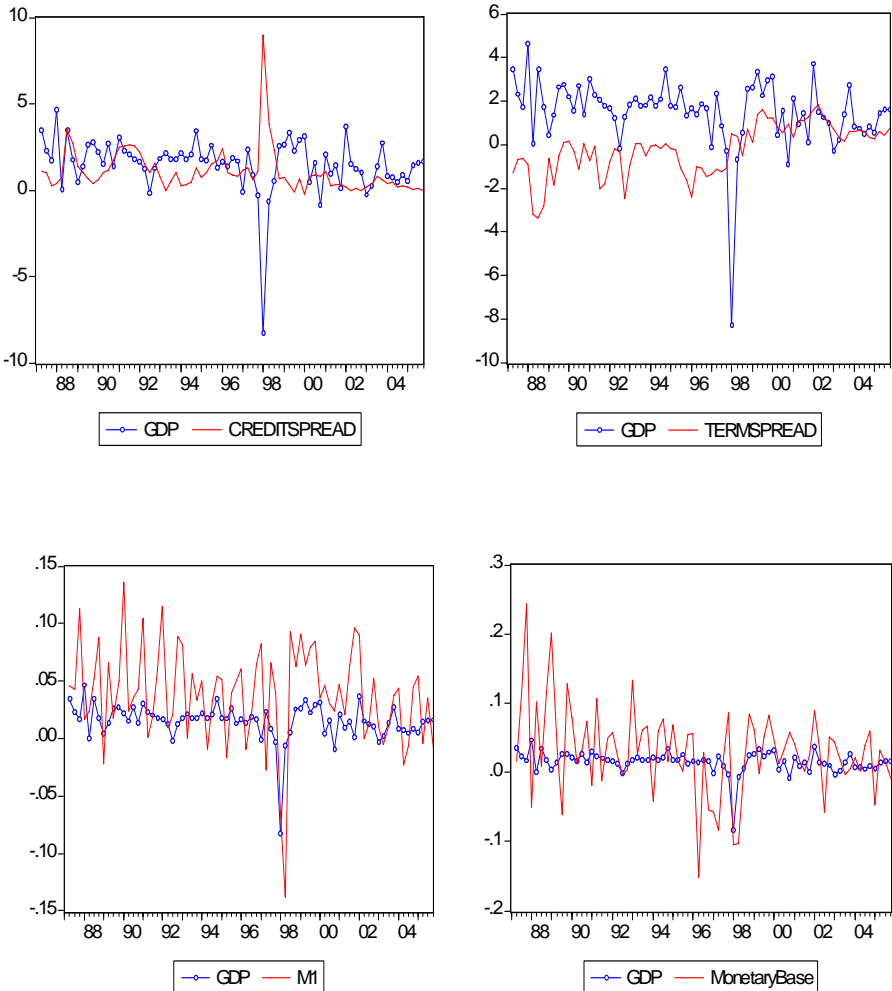
■ 참 고 문 헌

1. 박준용 · 장유순 · 한상범, 『경제시계열 분석』, 경문사, 2002.
2. Abeyasinghe, T., "Forecasting Singapore's Quarterly GDP with Monthly External Trade," *International Journal of Forecasting*, Vol. 14, 1998, pp.505-513.
3. Barro, R., "On the Determination of the Public Debt," *Journal of Political Economy*, Vol. 87, 1979, pp.940-71.
4. Cadzow, J.A., "Spectral Estimation: An Overdetermined Rational Model Equation Approach," *Proceedings of the IEEE*, Vol. 70, 1982, pp.907-939.
5. Chen, B. and P.A. Zadrozny, "An Extended Yule-Walker Method For Estimating A Vector Autoregressive Model With Mixed Frequency Data," *Advances In Econometrics: Messy Data-Missing Observations, Outliers, and Mixed-Frequency Data*, edited by T. Fomby, and R. Hill, 13, 1998, pp.47-73.
6. Clements, M. and D. Hendry, *Forecasting Economic Time Series*, Cambridge University Press, Cambridge, 1998.
7. Diebold, F. and J. Nason, "Nonparametric Exchange Rate Prediction?," *Journal of International Economics*, Vol. 28, 1990, pp.315-32.
8. Estrella, A. and F.S. Mishkin, "Predicting U.S. Recessions: Financial Variables as Leading Indicators," *Review of Economics and Statistics*, Vol. 80, 1998, pp.45-61.
9. Friedman, B.M. and K.N. Kuttner, "Money, Income, Prices and Interest Rates," *American Economic Review*, Vol. 82, 1992, pp.472-92.
10. _____, "Why Does the Paper-Bill Spread Predict Real Economic Activity?," *Business Cycles, Indicators, and Forecasting*, NBER; edited by Stock, J. and M.W. Watson, University of Chicago Press, 1993.
11. Hall, R., "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, Vol. 86, 1978, pp.971-987.

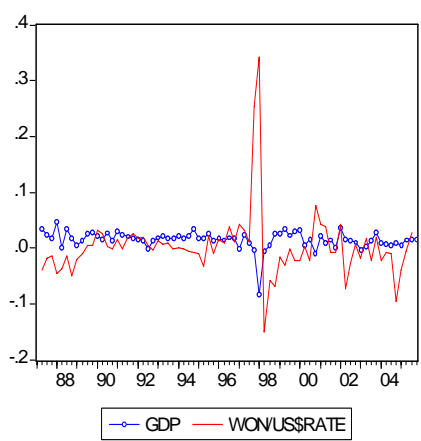
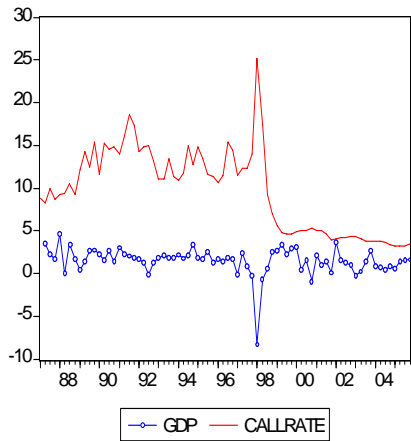
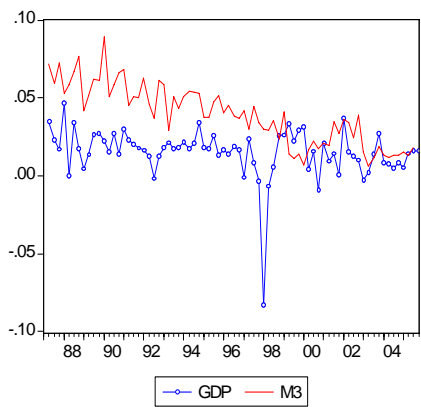
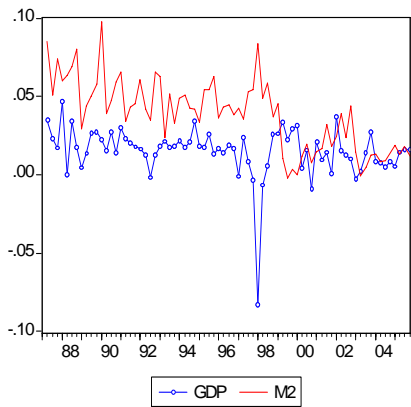
12. Hansen, L. R., "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, Vol. 50, 1982, pp.1029-54.
13. Ingentino, R. and B. Trehan, "Using Monthly Data to Predict Quarterly Output," *Federal Reserve Bank of San Francisco Economic Review*, Vol. 3, 1996, pp. 3-12.
14. Kashyap, A., C. Stein and D. Wilcox, "Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance," *American Economic Review*, Vol. 83, 1993, pp.78-98.
15. Kim, Yun-Yeong and Joon Y. Park, "VAR Model Estimation Under Mixed-Frequency Data," Manuscript, 2007.
16. Litterman, R., "Forecasting with Bayesian Vector Autoregressions-Five Years of Experience," *Journal of Business and Economic Statistics*, Vol. 4, 1986, pp.25-38.
17. Lupoletti W. and R. Webb, "Defining and Improving the Accuracy of Macroeconomic Forecasts: Contributions from a VAR Model," *The Journal of Business*, Vol. 59, 1986, pp. 263-285.
18. Mankiw, G., "The Optimal Collection of Seigniorage: Theory and Evidence," *Journal of Monetary Economics*, Vol. 20, 1987, pp. 327-41.
19. Miller, P. and D. Chin, "Using Monthly Data to Improve Quarterly Forecasts," *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 20, 1996, pp.16-33.
20. Nelson, C., "The Prediction Performance of the FRB-MIT-PENN Model of the U.S. Economy," *American Economic Review*, Vol. 62, 1972, pp.902-917.
21. Nunes, L.C., "Nowcasting Quarterly GDP Growth in a Monthly Coincident Indicator Model," *Journal of Forecasting*, Vol. 24, 2005, pp.575-592.
22. Sims, C., "Macroeconomics and Reality," *Econometrica*, Vol. 48, 1980, pp.1-48.
23. Stock, J. and M.W. Watson, "Interpreting the Evidence on Money Income Causality," *Journal of Econometrics*, Vol. 40, 1989, pp.161-82.
24. Thoma, M. and J.A. Gray, "Financial Market Variables Do Not Predict Real Activity," *Economic Inquiry*, Vol. 36, 1998, pp.522-39.
25. Webb, R.H., "Two Approaches to Macroeconomic Forecasting," *Federal Reserve Bank of Richmond Economic Quarterly*, Vol. 85, 1999, pp.23-40.
26. Weber, C.E., "Consumption Spending and the Paper-Bill Spread: Theory Evidence," *Economic Inquiry*, Vol. 36, 1998, pp.575-589.
27. Zadrozny, P.A., "Estimating A Multivariate ARMA Model with Mixed-frequency Data: An Application to Forecasting U.S. GNP at Monthly Intervals," *Research Paper*, Research Department, Federal Reserve Bank of Atlanta, 1990.
28. Zheng, I. and J. Rossiter, "Using Monthly Indicators to Predict Quarterly GDP," *Bank of Canada Working Paper*, August, 2006, pp.1-100.

〈참 고〉

분기 GDP 성장률과 금융변수(1개월 시차)와의 관계³⁴⁾



34) CREDITSPREAD: 신용스프레드, TERMSPREAD: 기간스프레드, MonetaryBase: 본원통화, CALLRATE: 콜금리, WON/US\$RATE: 원/달러환율



Forecasting GDP Growth Rates Using VAR with Mixed Frequency Data

Yun-Yeong Kim* · Joon Y. Park**

Abstract

In this paper, we develop a new forecasting method for the quarterly GDP growth rates, which utilize other economic variables that are observed at a monthly frequency. For this, we formulate a monthly VAR model including all the relevant variables, from which a forecasting equation for the quarterly GDP growth rates is derived. The derived equation has a finite order moving average error, and therefore, can be readily estimated by GMM using lagged observations as instruments. To investigate the practical relevancy, we used the proposed method to forecast the Korean GDP growth rates. We find in particular that M1 and paper-bill credit spread, among potentially relevant financial variables that are observed monthly, improve the forecast of the Korean GDP growth rates significantly in terms of both mean squared error and median bias.

Key Words: mixed-frequency data, VAR model, forecasting GDP growth rates

* Economist, Institute for Economic and Monetary Research, The Bank of Korea

** Professor, Department of Economics, Sungkyunkwan University and Texas A&M University