

VAR모형을 이용한 분기GDP 예측모형 연구

2006년 12월

연구자 : 통계개발원 경제통계실 정 선경

Tel. (XXXXXXXXXX

e-mail : XXXXXXXXXXXXX

主 要 內 容

- 월별 자료를 이용하여 당기 GDP를 예측하는 분기GDP모형과 상호 보완할 수 있도록 VAR모형을 개발하여 단일 모형운영에 따른 전망담당자의 부담을 경감시켜 주고자 한다.
- VAR모형은 회귀분석방법과 시계열분석 방법이 결합된 형태로서 최소자승법(OLS)에 의해 손쉽게 추정되고 모형의 예측력도 대체적으로 양호한 것으로 알려져 있다.
- GDP, 산업생산지수, 도소매업판매액지수 및 비농림취업자수 등의 계절조정계열을 구한 후 Log변환하여 추정에 사용하였으며, 1995년부터 2005년까지 분기자료를 대상으로 5분기 시차를 적용한 4변수 VAR(5) 모형을 설정하였다.
- 분기GDP모형과 예측력을 비교해 본 결과 VAR모형의 예측력이 다소 높은 것으로 추정되어 향후 예측작업에 활용할 수 있을 것으로 판단된다.

1. 서론

오늘날 세계화가 진전되고 불확실성이 확대됨에 따라 경제전망은 경제주체들이 합리적인 경제활동을 영위하는데 있어서 중요한 역할을 하고 있다. 특히 최근의 우리경제는 세계경기 하락, 환율 불안, 북한 핵 사태 악화, 건설경기 위축 등 국내외 경제여건 악화에 의해 전망시계가 더욱 불투명해지고 있어 한국은행 등을 비롯한 주요 전망기관의 향후 경제전망에 많은 관심이 모아지고 있다.

통계청의 경우에도 NSO-06모형과 분기GDP모형을 이용하여 GDP 등 주요 경제지표들을 예측하고 있으나 정책결정 및 수행과정에서의 피드백 목적보다는 통계청에서 작성하는 각종 통계지표가 경제흐름을 잘 반영하고 있는지 분석하기 위한 목적으로 주로 이용되고 있다.¹⁾

통계청은 1989년부터 계량경제모형 개발을 시작하여 1992년 「통계청 계량경제모형(NSO-92)」을 처음으로 개발하였다. 이후 기준년 및 주요 통계편제체제 변경 등을 반영한 3차례의 개정작업을 거쳐 현재의 NSO-06모형에 이르고 있다. NSO-06모형은 총수요, 국외거래, 물가 및 임금, 고용, 통화 및 금융 등 5개 부문 28개 행태식과 15개 정의식을 포함하는 총 43개의 연립방정식체제로 구성되어 있다.

한편 NSO-06모형 이외에 GDP가 공표되기 전 예측시점에서 이용 가능한 월별자료를 사용하여 당기 GDP를 효과적으로 예측하기 위한 분기GDP모형이 존재한다. 분기GDP모형은 공급측면의 증가율변수 모형과 지출측면의 수준변수 모형으로 나뉘어져 있으며 예측력에서는 공급측면 모형이 우수하여 공급측면 모형으로 전체GDP 규모를 추정한 후 이를 바탕으로 지출측면 모형을 운영하고 있다.

실제 예측에 있어서 공급측면 모형의 예측력이 상당히 우수하지만 전망담당자의 입장에서는 하나의 모형에서 나온 결과에 의존하여 GDP 예측결과를 확정하기에는 늘 부담이 따르게 마련이다. 이러한 부담을 경감하고 보다

1) 통계작성기관에서의 통계분석이나 경제예측을 실시하는 것에 대해서는 논란의 여지가 많으므로 본 연구에서는 논외로 한다.

효율적인 GDP예측이 가능하도록 기존의 분기GDP모형을 보완할 수 있는 소규모 모형으로, 단기예측 및 경제충격을 식별하기 위해 작성하는 VAR모형을 개발할 필요가 있는 것으로 판단된다.²⁾

GDP는 공급측면과 지출측면에서 각각 접근할 수 있지만 지출측면의 GDP에 대한 분석에 보다 많은 관심이 모아지고 있다. 본 연구는 산업생산지수와 같은 통계청에서 생산하는 지표를 바로 활용할 수 있는 이점이 십분 발휘될 수 있도록 공급측면에서 접근하도록 할 것이다.

본 연구의 구성은 다음과 같다. 제2장에서는 VAR모형에 대해 요약하고 제3장에서는 자료에 대한 여러 가지 통계적 검정을 실시한 후 VAR모형을 추정하기로 한다. 제4장에서는 VAR모형과 분기GDP모형의 예측력을 비교·평가하고 제5장에서는 각 충격에 대한 GDP의 반응을 살펴보고 GDP의 변동에 대한 각 충격의 중요도를 측정할 것이다. 마지막으로 제6장에서는 본 연구결과를 요약하고 시사점 및 결론을 제시하기로 한다.

II. 벡터자기회귀(VAR)모형의 이해³⁾

경제예측에는 거시계량모형에 의한 예측방법, 시계열모형에 의한 예측방법 또는 예측 전문가의 판단에 의한 방법 등 여러 가지가 있다.

거시계량모형은 전통적인 회귀모형에 의한 구조모형으로 경제이론에 근거하여 변수간의 인과관계를 설명하고 있다. 그러나 구조모형은 경제이론에 의해서 모형을 구축하고 있어 변수 선택 및 모형의 내·외생변수의 선정이 모형설계자의 주관에 의해서 결정되며 설명변수의 영향이 시간이 변하더라도 항상 일정하다는 가정을 하고 있어 구조적 변화가 급속히 진행되어 설명변수의 영향이 변한 경우 이를 적절하게 반영하지 못한다는 약점이 있다.

이에 Box and Jenkins 등이 개발한 ARIMA모형이 거시계량모형보다 정

2) 한국은행의 경우에도 전망시계를 달리하는 초단기예측모형, 단기예측모형 및 단순시계열모형 등을 개발하여 모형별 예측결과를 종합하는 단기예측시스템을 구축하여 경제예측에 이미 활용하고 있다. 김양우·이궁희·장동구(1997) 참조

3) 문권순(1999) 참조

확한 예측을 도출할 수 있다는 사실이 밝혀짐에 따라 시계열분석을 통한 예측이 각광을 받았다. 그러나 이러한 방법은 모형설정이 용이한 반면 변수들 사이의 상호작용을 무시하는 일변량분석이라는 문제점과 장기예측에 있어서는 그 정확도가 떨어진다는 한계를 지니고 있다.

Sims의 VAR모형은 회귀분석과 시계열분석이 결합된 형태로 전통적인 자기회귀이동평균모형을 다변량모형으로 확장시킨 모형으로 볼 수 있다. 최소자승법(OLS)에 의해 손쉽게 추정되고 모형의 예측력도 대체적으로 양호한 것으로 알려져 있다.

VAR모형의 장점은 (1)단순하다. 어느 변수가 내생변수이고 어느 변수가 외생변수인지를 결정할 필요가 없다. VAR 내의 모든 변수는 내생적이다. (2) 추정이 단순하다. 즉, 통상적인 OLS방법이 각 개별 회귀식에 대해 적용될 수 있다. (3)많은 경우에 있어서 이 방법으로 구한 예측결과가 더 복잡한 연립방정식 모형으로 구한 예측결과보다 더 낫다. (4)일반적으로 모형 내에 포함되는 변수가 많지 않아 실제 예측을 수행하는데 비용과 시간이 절약되고, 또 대규모 모형에서 제기되는 자료의 수집과 변형의 번거로운 문제도 없다.⁴⁾

VAR모형의 문제점은 다음과 같다. (1)연립방정식모형에 비해 VAR모형은 사전적인 정보를 덜 안고 있기 때문에 비이론적이다. (2)예측을 강조하기 때문에 정책효과 분석에 적합하지 않다. (3)VAR모형의 가장 큰 실제 운용상의 난관은 적절한 시차길이의 선택이다. (4)추정결과가 모형 내에 포함된 변수의 선택 및 시차의 길이에 따라 민감하게 변화하는 등의 문제점을 지니고 있다.

N개의 시계열로 구성된 X_t 가 p시차인 자기회귀모형으로 구성된 모형을 VAR(p)라 하며 다음과 같이 정의한다.

$$X_t = C + \sum_{i=1}^p A_i X_{t-i} + \epsilon_t$$

이때 C 는 (N×1) 상수벡터, A_i 는 (N×N)의 계수행렬, ϵ_t 는 (N×1)의 벡터백색 잡음과정으로 $E(\epsilon_t) = 0$ 이며, 다음의 공분산행렬을 갖는다.

$$\sum \epsilon = E(\epsilon_t \epsilon_s') = \begin{cases} \sigma^2, & t = s \\ 0, & t \neq s \end{cases}$$

4) 안충영 · 홍성표 · 박완규(2004) 참조

III. VAR모형의 추정

1. 변수의 선정

VAR모형은 다변량분석 방법으로 주요 관심변수인 GDP 이외에 GDP와 관련된 다른 경제변수를 고려하여야 한다. VAR모형을 이용한 GDP예측모형에 대한 선행연구를 살펴보면 김양우 등(1997)은 GDP, 산업생산지수, 도소매업활동지수, 비농림취업자수 및 GDP디플레이터 등을 분석대상변수로 채택하였다. 이충렬(1999)은 GDP와 경기동행지수, GDP와 경기선행지수를 각각 VAR모형으로 추정하였으며 성병희(2001)는 원재료가격, 총통화, 콜금리, 소비자물가, GDP, 실업률 등 6개 경제변수를 이용한 베이저안 VAR모형을 소개하고 있다. 이종원 등(2004)은 GNP, 총통화, 순재정적자, 환율, 물가, 임금, 이자율 등 7개 변수를 대상으로 VAR모형을 추정하였다.

본 연구에서는 김양우 등(1997)의 연구를 쫓아 GDP(GDP_x), 산업생산지수(IPI_x), 도소매업판매액지수(WRI_x) 및 비농림취업자수($EPNA_x$)를 대상으로 VAR모형을 구축하고자 한다.⁵⁾ 통계청에서 매월 발표하는 주요 통계자료를 이용하여 GDP를 예측할 경우 시의성 있는 예측이 가능하며 분기GDP모형에서도 이들 변수를 이용하여 예측작업을 하고 있어서 따로 변수를 수집하는 번거로움을 줄일 수 있는 이점이 있다.

각 변수들은 공표기관이 제공하는 계절조정계열을 그대로 사용하였으며 공표되는 계절조정계열이 없는 경우에는 X-12-ARIMA방법을 통해 계절성을 제거한 후 log변환을 취한 다음 분석에 이용하였다.

분석에 이용된 자료는 1995년 1/4분기부터 2006년 3/4분기까지의 분기자료이다. 1995년 1/4분기부터 2005년 4/4분기까지의 자료는 모형을 식별하고 추정하는데 사용하였으며, 나머지 2006년 1/4분기부터 2006년 3/4분기까지의 자료를 이용하여 모형의 예측력을 검증하기로 하였다.

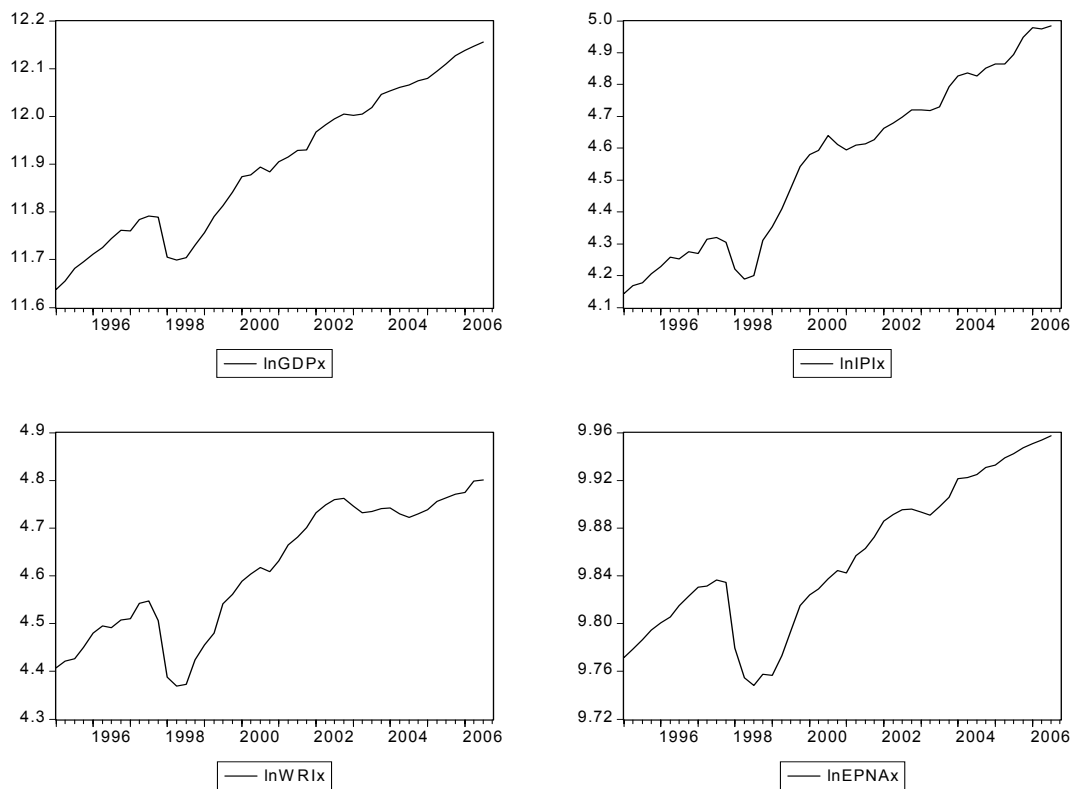
본 연구에서 모형의 추정과 검정 작업에는 Eviews 4.0 프로그램을 사용하였다. 자료에 대한 기초통계량을 <표 1>에 정리하였다.

5) 이외에도 서비스업활동지수와 건설기성액 등을 고려하였으나 이용 가능한 시계열이 짧아서 시계열의 안정성 미확보 등 추정에 어려움이 있어 최종적으로는 분석대상에서 제외하였다.

<표 1> 기초통계량

	$\ln GDP_x$	$\ln IPI_x$	$\ln WRI_x$	$\ln SPI_x$
평 균	11.88	4.53	4.60	9.85
중 위 수	11.88	4.59	4.61	9.84
최 대 값	12.13	4.95	4.77	9.95
최 소 값	11.64	4.14	4.37	9.75
표준편차	0.15	0.25	0.13	0.06
왜 도	0.08	-0.05	-0.20	0.07
첨 도	1.66	1.58	1.57	1.83
Jarque-Bera (p-value)	3.34 (0.19)	3.70 (0.16)	4.06 (0.13)	2.56 (0.28)
관측치수	44	44	44	44

<그림 1>의 시계열 추이를 살펴보면 직관적으로도 각 시계열의 평균치가 증가하는 추세를 보여 비정상시계열임을 짐작할 수 있다. 시계열의 안정성에 관해서는 다음 절에서 분석하기로 한다.



<그림 1> 시계열 추이

2. 단위근 검정

기본적으로 시계열 자료를 분석할 때는 시계열이 안정적이라는 가정을 내재하고 있다. 시계열의 안정성이란 어떤 시계열 자료가 단기 충격에 의해 추세치에서 이탈하더라도 결국에는 그 추세치로 돌아가려는 성질을 말한다. 즉, 어떤 시계열이 안정적이라면 평균, 분산이 해당 기간동안 일정하고 두 시점간의 공분산이 두 시점간의 거리 혹은 시차에만 의존한다는 의미이며 시계열이 이러한 조건을 만족하지 못 할 때 불안정 시계열이라고 한다.

불안정 시계열의 대표적인 예가 단위근이 있는 시계열이다. 어떤 시계열이 단위근을 가진다는 의미는 어떤 시점에서 시계열에 충격이 가해졌을 때 그 충격의 효과가 시간이 경과함에 따라 소멸되지 않고 지속적으로 영향을 준다는 것이다. 단위근을 갖는 시계열을 회귀분석에 그대로 이용하면 상관관계가 없는 변수 사이에도 강한 상관관계가 있는 것으로 나타나는 가성회귀(Spurious regression)의 문제가 발생한다.

분석에 사용된 변수들이 안정적 시계열인지 확인하기 위해 ADF검정을 이용하여 단위근 검정을 실시하였다. ADF검정은 다음과 같이 상수항만 존재하는 모형(c), 추세와 상수항이 있는 모형(t), 추세항과 상수항이 존재하지 않는 모형(n)을 가정하여 $\rho = 1$ 이라는 귀무가설을 검정하는 것으로 귀무가설을 기각하지 못하면 시계열이 단위근을 가진다는 의미이다. 표본기간은 1995년 1/4분기부터 2005년 4/4분기까지를 대상으로 하였다.

$$c: y_t = \beta_0 + \rho y_{t-1} + \epsilon_t$$

$$t: y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \epsilon_t$$

$$n: y_t = \rho y_{t-1} + \epsilon_t$$

<표 2>에서 수준변수에 대한 단위근 검정결과를 살펴보면 각 변수들은 1% 유의수준에서 단위근을 갖는다는 귀무가설을 기각하지 못 하였으므로 본 연구에 사용되어진 시계열들은 모두 불안정한 시계열로 판명되었다.

한편 1차 차분변수에 대한 단위근 검정 결과에서는 모든 변수들이 1% 유의수준에서 단위근이 존재한다는 귀무가설을 기각하여 안정적인 시계열인

것으로 나타났다.

<표 2> 수준변수에 대한 ADF검정 결과

	Lag	수준변수			1차 차분		
		c	t	n	c	t	n
<i>lnGDPx</i>	1	-0.3660	-2.5583	2.4682 *	-3.5159 *	-3.4802	-2.7523 *
	2	-0.2130	-2.9954	2.0401 *	-3.8947 **	-3.8549 *	-2.7950 **
	3	-0.1161	-2.3929	2.4974 *	-2.9943 *	-2.9760	-2.0690 *
	4	-0.0835	-2.8004	2.0685 *	-3.2533 *	-3.2484	-2.0953 *
	5	0.1115	-2.4340	2.3651 *	-3.1632 *	-3.1949	-1.9060
<i>lnIPIx</i>	1	-0.2507	-3.0220	2.1241	-3.5521 *	-3.5109	-2.7037 **
	2	-0.3144	-3.0610	2.1351	-3.8404 **	-3.8038 *	-2.7707 **
	3	-0.0364	-2.6126	2.4506	-2.7977	-2.7788	-1.9098
	4	-0.0964	-3.2882	1.9616	-3.3323 *	-3.3145	-2.1357 *
	5	0.1115	-2.7379	2.4276	-2.8630	-2.8417	-1.6241
<i>lnWRIx</i>	1	-1.1927	-2.7088	1.0271	-3.5749 *	-3.5324 *	-3.3439 **
	2	-1.1365	-2.4698	1.1757	-3.8355 **	-3.7782 *	-3.5576 **
	3	-0.8030	-2.0420	1.3146	-2.8905	-2.8377	-2.6892 **
	4	-0.7413	-2.5196	1.0384	-2.8057	-2.7581	-2.5330 *
	5	-0.7125	-2.3481	1.1556	-2.7144	-2.6689	-2.3590 *
<i>lnEPNax</i>	1	-0.7955	-2.4183	1.1551	-3.3800 *	-3.3818	-3.1540 **
	2	-0.6497	-2.3521	1.1716	-3.0763 *	-3.1021	-2.8391 **
	3	-0.5299	-2.3686	1.1521	-2.7597	-2.8023	-2.5061 *
	4	-0.4860	-2.4369	1.1314	-3.2423 *	-3.3147	-2.8543 **
	5	-0.1083	-1.9232	1.4471	-3.0796 *	-3.2519	-2.6698 **

주) *(**)은 5%(1%) 유의수준에서 각각 유의함을 나타낸다.

3. 공적분 검정

단위근 검정결과에 의하면 모든 변수들이 불안정 시계열인 것으로 판정되었다. 이런 경우 변수들에 대한 전통적인 회귀분석은 무의미하거나 가성회귀가 될 수 있다. 그러나 개별 시계열이 단위근이 있는 불안정한 시계열이지만 이들 시계열 간에 장기적으로 안정적인 균형관계를 갖도록 하는 선형결합이 존재한다면 이들 시계열은 공적분 관계에 있다고 한다.

변수들 간의 공적분 관계가 성립하는지 Johansen검정에 의한 공적분 검정을 실시하였다. 시차를 5로 했을 때 공적분검정을 실시한 결과는 <표 3>과 같다.⁶⁾ 공적분검정 결과에 따르면 각 변수들 간에는 장기균형관계가 존재하고 2개의 공적분 벡터가 존재함을 알 수 있다.

6) VAR모형의 시차를 따른 것이다.

<표 3> Johansen 공적분검정 결과

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value	
			5 Percent	1 Percent
None **	0.82112	117.05630	62.99	70.05
At most 1 **	0.54392	51.65668	42.44	48.45
At most 2	0.31714	21.82317	25.32	30.45
At most 3	0.17538	7.327492	12.25	16.26

주) *(**)는 5%(1%) 유의수준에서 각각 유의함을 나타낸다.

4. Granger 인과검정

VAR모형은 변수 간의 순서와 시차의 길이에 의해 결과가 달라질 수 있다. 본 연구에서는 변수의 순서를 Granger의 인과검정을 통해서 정하기로 한다.

여기서 X가 Y를 Granger 인과한다는 의미는 Y의 과거정보만을 가지고 Y를 예측할 때 보다 X와 Y의 과거정보를 이용하여 Y에 대한 예측을 더 잘 할 수 있다는 뜻한다. 각 변수들 간의 Granger 인과검정 결과는 <표 4>에 정리하였다.

Granger 인과검정 결과를 살펴보면 1시차에서 산업생산지수가 GDP를 인과하고 GDP는 산업생산지수를 4시차에서 인과하는 것으로 나타났다. 도소매업 판매액지수의 변동이 GDP 변동의 원인이 되지만 그 반대의 경우는 인정되지 않는 것으로 나타났다. 비농림취업자수와 GDP, 도소매업판매액지수와 산업생산지수, 비농림취업자수와 산업생산지수 간에는 양방향 모두 인과관계가 성립한다. 1시차에서 비농림취업자수는 도소매업판매액지수를 인과하고 모든 시차에서 도소매업판매액지수는 비농림취업자수를 인과하는 것으로 나타났다.

따라서 이상을 결과를 종합하여 도소매업판매액지수, 산업생산지수, 비농림취업자수 및 GDP의 순서로 VAR모형을 설정하기로 하였다.⁷⁾ 이는 소비가 생산을 촉진시키고 생산이 다시 일자리를 창출하는 경제순환현상을 잘 설명하는 것으로 판단된다.

7) VAR모형에서는 변수의 나열순서에 따라 충격반응함수의 결과가 달라지기 때문에 외생성이 강한 변수부터 먼저 나열하는 것이 일반적이다.

<표 4> Granger 인과검정 결과

귀무가설	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5
$\ln IPI_x \nrightarrow \ln GDP_x$	5.49 (0.02)*	2.11 (0.14)	1.21 (0.32)	1.49 (0.23)	1.23 (0.32)
$\ln GDP_x \nrightarrow \ln IPI_x$	0.88 (0.35)	0.40 (0.67)	0.34 (0.80)	2.34 (0.08)*	2.01 (0.11)
$\ln WRI_x \nrightarrow \ln GDP_x$	0.04 (0.84)	6.09 (0.01)*	4.02 (0.02)*	2.78 (0.04)*	2.23 (0.08)*
$\ln GDP_x \nrightarrow \ln WRI_x$	0.08 (0.78)	1.93 (0.16)	0.96 (0.42)	0.77 (0.55)	0.75 (0.59)
$\ln EPNA_x \nrightarrow \ln GDP_x$	3.94 (0.05)*	2.62 (0.09)*	1.88 (0.15)	1.84 (0.15)	2.93 (0.03)*
$\ln GDP_x \nrightarrow \ln EPNA_x$	4.72 (0.04)*	4.23 (0.02)*	2.80 (0.05)*	3.10 (0.03)*	2.94 (0.03)*
$\ln WRI_x \nrightarrow \ln IPI_x$	2.23 (0.14)	4.08 (0.02)*	2.52 (0.07)*	4.55 (0.01)*	3.61 (0.01)*
$\ln IPI_x \nrightarrow \ln WRI_x$	2.26 (0.14)	3.71 (0.03)*	2.21 (0.10)*	1.41 (0.25)	1.27 (0.30)
$\ln EPNA_x \nrightarrow \ln IPI_x$	7.36 (0.01)*	2.12 (0.13)	2.22 (0.10)*	2.04 (0.11)	1.69 (0.17)
$\ln IPI_x \nrightarrow \ln EPNA_x$	6.80 (0.01)*	2.86 (0.07)*	1.94 (0.14)	1.70 (0.17)	1.32 (0.28)
$\ln EPNA_x \nrightarrow \ln WRI_x$	6.23 (0.02)*	1.05 (0.36)	0.51 (0.68)	1.08 (0.38)	1.19 (0.34)
$\ln WRI_x \nrightarrow \ln EPNA_x$	9.14 (0.00)*	7.22 (0.00)*	4.50 (0.01)*	4.92 (0.00)*	4.15 (0.01)*

주) 1. 수치는 F-statistics와 p-value를 나타낸다.
 2. *는 10% 유의수준에서 유의함을 나타낸다.

5. 시차의 길이

한편, VAR모형의 시차는 LR, AIC 등과 같은 정보기준 통계량을 사용하여 적정시차를 선택하기로 한다.

<표 5>의 정보기준에 의하면 SIC는 2분기 시차를, LR는 5분기 시차를, AIC와 HQ는 7분기 시차를 각각 최적시차로 제시하고 있다. 따라서 본 연구에서는 분석변수가 분기별 자료인 것을 고려하여 LR기준에 의해 5분기 시차를 적정시차로 결정하였다.⁸⁾

<표 5> 시차길이 결정을 위한 정보기준

Lag	LogL	LR	AIC	SIC	HQ
0	272.4484	NA	-14.51073	-14.33657	-14.44933
1	447.4631	302.7282	-23.10612	-22.23535 *	-22.79913
2	468.2403	31.44654	-23.36434	-21.79696	-22.81177
3	485.1801	21.97598	-23.41514	-21.15115	-22.61698
4	495.5349	11.19437	-23.11000	-20.14939	-22.06624
5	527.2328	27.41438 *	-23.95853	-20.30131	-22.66919
6	554.9527	17.98048	-24.59204	-20.23821	-23.05711
7	597.9947	18.61277	-26.05377 *	-21.00332	-24.27325 *

주) 1) LR: sequential modified LR test statistic, AIC: Akaike information criterion
 SC: Schwarz information criterion, HQ: Hannan-Quinn information criterion
 2) *는 5% 유의수준에서 각 정보기준에 의해 선택된 시차를 나타낸다.

8) 너무 적은 시차를 택하면 모형의 동태적 관계를 충분히 반영하지 못 하고 너무 많은 시차를 택하면 과다추정으로 자료의 손실이라는 문제가 발생한다.

6. VAR모형의 추정

앞의 과정에서 변수의 순서와 시차길이가 결정됨에 따라 1995년 1/4분기부터 2005년 4/4분기까지를 표본기간으로 하여 4변수 VAR(5) 모형을 설정하였으며 모형의 추정결과는 <표 6>에 제시하였다.

VAR모형은 경제이론을 바탕으로 하지 않기 때문에 계수값에 대한 언급은 생략하기로 한다. 다만 GDP에 대한 추정결과만 아래에 요약하였다.

$$\begin{aligned} \ln GDPx(t) = & 0.6265 + 0.8593 \sum_i^5 \ln GDPx(t-i) + 0.0764 \sum_i^5 \ln WRIx(t-i) \\ & + 0.0175 \sum_i^5 \ln IPIx(t-i) + 0.0640 \sum_i^5 \ln EPNAx(t-i) \end{aligned}$$

각 변수의 회귀계수 값은 표기의 편의를 위해 각 시차별 회귀계수 값을 합한 것으로 모두 양의 값으로 추정되어 GDP와 서로 정(+)의 관계에 있음을 나타내 준다.

IV. 모형의 예측력 평가

1. 평가방법

모형이 변수의 예측에 이용되는 예측모형의 경우에 피설명변수의 예측능력으로 모형을 평가하는 것이 바람직할 것이다. 이 연구에서는 여러 가지 예측능력 평가방법 중에서 post-sample 예측능력 평가⁹⁾를 모형의 평가방법으로 채택하였다.

예측능력 평가대상 기간은 2006년 1/4분기부터 2006년 3/4분기까지 3분기로 설정하였으며 모형에 의해 나온 예측치는 자연대수 값이므로 다시 지수함수를 사용하여 수준값으로 변환한 후 예측오차를 측정하였다. <표 7>에는 예측오차의 크기를 비교하기 위해 평방근평균제곱오차(Root Mean Squared

9) post-sample 예측능력 평가방법은 t 기까지의 자료를 가지고 추정된 모형으로 $t+s$ 기까지의 피설명변수를 예측한 후 그 예측치와 실적치의 오차를 평가하는 것이다.

<표 6> VAR모형 추정 결과

	<i>lnGDPx</i>	<i>lnWRIx</i>	<i>lnIPlx</i>	<i>lnEPNAx</i>
<i>lnGDPx(-1)</i>	0.676405 (1.10454)	0.607054 (0.61827)	1.224923 (1.23961)	0.449483 (1.15709)
<i>lnGDPx(-2)</i>	-0.04346 (-0.07740)	0.284773 (0.31630)	0.766707 (0.84618)	0.010603 (0.02977)
<i>lnGDPx(-3)</i>	0.016621 (0.03433)	-0.207706 (-0.26756)	-0.130767 (-0.16738)	-0.140668 (-0.45801)
<i>lnGDPx(-4)</i>	0.586651 (1.25283)	0.404109 (0.53826)	0.137687 (0.18223)	0.146869 (0.49445)
<i>lnGDPx(-5)</i>	-0.376936 (-0.74086)	-0.834327 (-1.02277)	-0.319318 (-0.38895)	-0.525828 (-1.62925)
<i>lnWRIx(-1)</i>	0.535692 (1.71481)	1.405846 (2.80682)	0.486763 (0.96565)	0.234305 (1.18239)
<i>lnWRIx(-2)</i>	-0.443319 (-1.16163)	-0.658968 (-1.07694)	-0.584149 (-0.94859)	-0.358288 (-1.48001)
<i>lnWRIx(-3)</i>	-0.393485 (-1.00158)	-0.375714 (-0.59647)	-0.854267 (-1.34758)	-0.120403 (-0.48314)
<i>lnWRIx(-4)</i>	0.698112 (1.72023)	0.633811 (0.97409)	1.280017 (1.95471)	0.31389 (1.21932)
<i>lnWRIx(-5)</i>	-0.37954 (-1.15648)	-0.175062 (-0.33270)	-0.543176 (-1.02571)	-0.092044 (-0.44214)
<i>lnIPlx(-1)</i>	0.049028 (0.18927)	-0.321865 (-0.77498)	0.466085 (1.11509)	-0.055312 (-0.33662)
<i>lnIPlx(-2)</i>	0.070154 (0.28234)	-0.048986 (-0.12296)	-0.18191 (-0.45372)	-0.02051 (-0.13013)
<i>lnIPlx(-3)</i>	-0.028057 (-0.11554)	0.238389 (0.61230)	0.142519 (0.36373)	0.067292 (0.43686)
<i>lnIPlx(-4)</i>	-0.014817 (-0.05502)	-0.017971 (-0.04162)	-0.035082 (-0.08074)	0.018322 (0.10726)
<i>lnIPlx(-5)</i>	0.00014 (0.00061)	0.24595 (0.67140)	0.08136 (0.22068)	0.068096 (0.46985)
<i>lnEPNAx(-1)</i>	-0.063555 (-0.11085)	-0.5109 (-0.55577)	-0.792053 (-0.85613)	0.541426 (1.48868)
<i>lnEPNAx(-2)</i>	0.194643 (0.27695)	0.348891 (0.30962)	-0.131932 (-0.11634)	0.258958 (0.58087)
<i>lnEPNAx(-3)</i>	0.292612 (0.41328)	0.011341 (0.00999)	-0.0583 (-0.05103)	0.024349 (0.05421)
<i>lnEPNAx(-4)</i>	-1.853894 (-2.58137)	-1.879246 (-1.63202)	-1.493863 (-1.28908)	-0.572736 (-1.25718)
<i>lnEPNAx(-5)</i>	1.494233 (2.24191)	1.306353 (1.22246)	0.960827 (0.89340)	0.654949 (1.54912)
C	0.626489 (0.30337)	4.463388 (1.34801)	-1.644288 (-0.49344)	1.369732 (1.04560)
R-squared	0.993025	0.980157	0.993864	0.98517
Adj. R-squared	0.985274	0.958109	0.987047	0.968692
Sum sq. resids	0.004767	0.012254	0.012411	0.001918
S.E. equation	0.016273	0.026091	0.026259	0.010323
F-statistic	128.124	44.45605	145.7863	59.78789
Log likelihood	120.3497	101.9384	101.6894	138.1014
Akaike AIC	-5.094855	-4.150687	-4.137917	-6.005198
Schwarz SC	-4.199091	-3.254923	-3.242153	-5.109434
Mean dependent	11.90483	4.620651	4.569135	9.854913
S.D. dependent	0.134101	0.127479	0.230722	0.05834
Determinant Residual Covariance		1.63000E-16		
Log Likelihood (d.f. adjusted)		487.53090		
Akaike Information Criteria		-20.69389		
Schwarz Criteria		-17.11084		

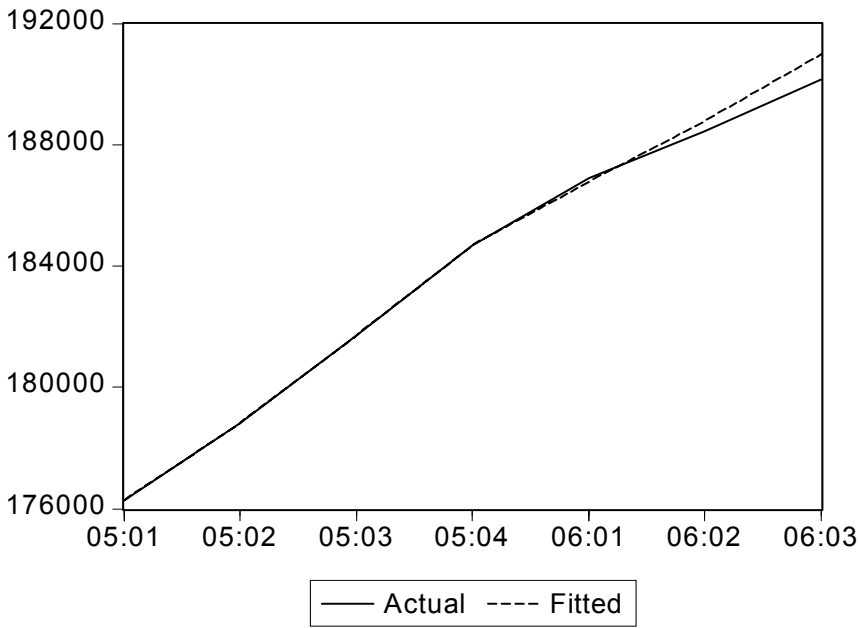
주) () 안의 t-statistics을 나타낸다.

Error%, RMSE%)¹⁰⁾를 계산하여 제시하였다. 예측력 평가결과에 의하면 모든 변수에 대한 예측오차의 크기가 5% 내외로 추정되어 예측력은 우수한 것으로 나타났다.

<표 7> VAR모형의 예측력 평가결과

	GDP_x	IPI_x	WRI_x	$EPNA_x$
RMSE%	0.2703	1.3027	2.9281	0.667

<그림 2>에 GDP에 대한 실제값과 예측값을 그림으로 나타내었다. 그림을 살펴보면 시간이 경과할수록 예측정도는 떨어지는 것을 알 수 있다.



<그림 2> GDP에 대한 실제값과 예측값 비교

2. 기존 모형과의 예측력 비교

분기GDP모형은 공급측면의 GDP를 광공업, 전기가스수도업, 건설업, 도소매음식숙박업, 서비스업 및 기타 등의 구성항목으로 구분하여 이들에 대해

10) $RMSE\% = \sqrt{\frac{1}{N} \sum_{s=1}^n \left(\frac{FX_{t+s} - X_{t+s}}{X_{t+s}} \right)^2}$, FX_{t+s} 는 예측치, X_{t+s} 는 실적치, N 는 관측치

각각 회귀방정식을 설정하는 6개의 회귀방정식으로 구성되어 있다.¹¹⁾

피설명변수와 설명변수 모두 전년동기대비 증가율을 사용하여 추정하는 증가율모형으로 설명변수로는 산업생산지수, 건축허가면적, 건설용중간재내수출하지수, 도소매업판매액지수, 취업자수, 경기동행지수 등이 사용되었다.

예측력 비교를 위해 VAR모형과 마찬가지로 1995년 1/4분기부터 2005년 4/4분기까지를 대상으로 모형을 추정한 후 2006년 1/4분기부터 2006년 3/4분기까지의 GDP를 예측하였다.

예측력 평가결과를 살펴보면 VAR모형의 예측력이 분기GDP모형의 예측력 보다 다소 우수한 것으로 나타났다. <표 8>은 VAR모형과 기존 분기GDP모형의 예측력을 비교한 것이다.

<표 8> VAR모형과 분기GDP모형의 예측력 비교결과

	VAR	분기GDP모형
RMSE%	0.2703	0.3799

V. 충격반응함수 및 분산분해

1. 충격반응함수

충격반응함수는 어떤 변수에 충격이 발생할 경우 모형 안의 다른 변수들이 시간이 지남에 따라 어떻게 동태적으로 반응하는지 경로를 나타낸 것이다. 따라서 VAR모형 내 충격반응함수를 통해 GDP에 대한 모형 내 변수의 영향을 살펴보고자 한다.

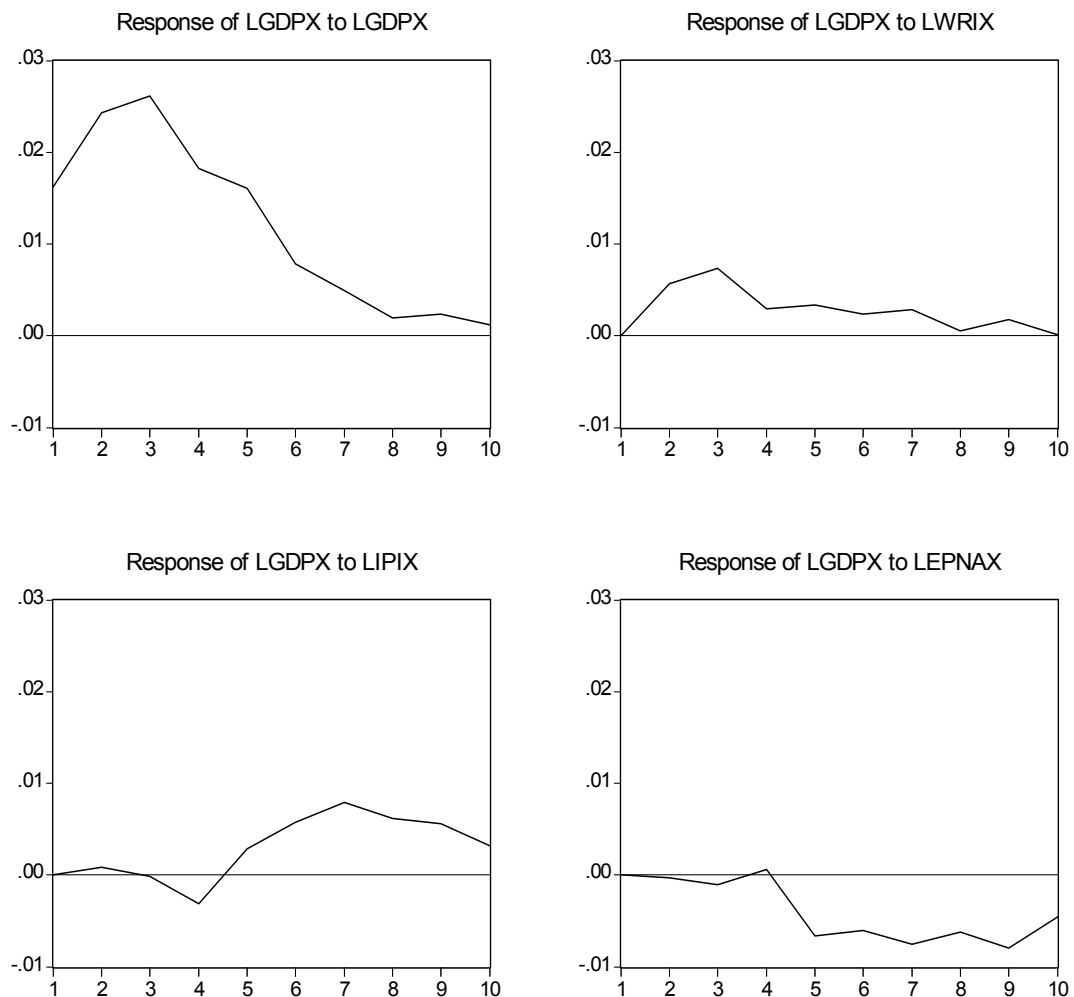
<그림 3>는 4개의 면으로 구성되어 있는데 이는 각 충격변수에 1 표준편차¹²⁾의 충격이 각각 가해졌을 때 시간에 따른 GDP의 동태적 반응을 나타낸

11) GDP를 추정할 때는 먼저 모형에 의해 각 구성항목별 GDP를 추정한 후 이를 합산하는 항등식 방법을 사용하고 있다. 자세한 내용은 민경삼 등(2002) 참조

12) 1 표준편차는 예측오차의 공분산행렬의 대각행렬을 평방근 한 값이다. GDP, 산업생산지수, 도소매업판매액지수 및 비농림취업자수의 표준편차는 각각 0.0163, 0.0261, 0.0263, 0.0103이다.

것이다. 즉, 도소매업판매액지수가 1 표준편차(2.61%) 만큼 상승하였을 때 3분기까지 GDP를 증가시키다가 그 이후는 증가폭이 점차 줄어드는 것으로 나타났다. 산업생산지수가 1 표준편차(2.63%) 만큼 증가하면 4분기까지는 별 영향을 미치지 못 하다가 5분기 이후부터는 GDP를 증가시키는 것으로 나타났다. 한편 비농림취업자수가 1 표준편차(1.03%) 만큼 상승하면 GDP는 4분기까지 영향을 미치지 못 하다가 5분기부터는 부(-)의 효과를 주는 것으로 나타났다.

Response to Cholesky One S.D. Innovations



<그림 3> GDP의 충격반응함수

2. 분산분해

분산분해는 한 내생변수의 예측기간을 늘려 나가면서 얻어지는 예측오차의 분산에서 모형 내 개별 설명변수의 변동에 기인하는 부문이 차지하는 비중을 백분율로 나타낸 것이다. 즉, 내생변수 변동에 대한 모형 내 개별 설명변수들의 상대적 중요성을 예측오차의 분산분해를 통해서 측정할 수 있다.

우리가 추정한 VAR(5)모형으로부터 예측오차의 분산분해 결과는 <표 9>에 제시되어 있다. 이 표의 숫자는 반응변수에 대한 예측오차의 분산 중에서 각 충격변수들에 의해서 설명되어지는 부분을 나타낸다.

<표 9>에 의하면 VAR(5) 모형으로 예측할 때 발생하는 예측오차의 총분산을 100%이라 할 때 예측오차의 분산 중 79.3%가 GDP 자체에 내재된 변화에 의해서 발생하고, 산업생산지수는 4.4%, 도소매업판매액지수는 6.9%, 비농림취업자수는 9.4%를 차지하는 것으로 나타났다. 이는 GDP 자체의 변동이 예측오차의 분산에 가장 크게 기여하고 있음을 알 수 있다.

예측오차의 분산에서 GDP의 비중이 서서히 줄어드는 반면 산업생산지수, 도소매업판매액지수 및 비농림취업자수는 1분기에는 아무런 영향이 없다가 점차 영향이 증가하지만 중요도는 GDP에 비해 낮은 편이다. 각 충격별 중요도는 GDP, 비농림취업자수, 도소매업판매액지수 및 산업생산지수 순서이다.

<표 9> GDP의 예측오차의 분산분해

분기	$\ln GDPX$	$\ln WRIX$	$\ln IPIx$	$\ln EPNAx$
1	100.0000	0.0000	0.0000	0.0000
2	96.2769	3.6325	0.0781	0.0124
3	94.5687	5.3163	0.0441	0.0710
4	94.5824	4.8030	0.5366	0.0781
5	92.5652	4.6220	0.8128	2.0000
6	89.8995	4.5955	2.1183	3.3867
7	85.5350	4.6409	4.4167	5.4075
8	83.0753	4.5112	5.7056	6.7080
9	80.1740	4.4588	6.6102	8.7570
10	79.3190	4.4090	6.8869	9.3851

VI. 결론

VAR모형에는 GDP이외에 산업생산지수, 도소매업판매액지수 및 비농림취업자수를 변수로 사용하였다. VAR모형의 단점의 하나로 지적되는 변수선정의 자의성을 보완하기 위해 주성분분석¹³⁾과 같은 방법을 이용하여 변수를 선정하기도 하지만 본 연구에서는 통계청에서 생산하는 월별지표를 이용할 경우 보다 시의성 있는 예측이 가능하다는 점을 들어 이들 세 변수를 이용하여 모형을 설정하였다.¹⁴⁾

분석에 앞서 시계열의 안정성을 검정을 위한 단위근 검정을 실시하였으며 단위근 검정 결과에 의하면 $I(1)$ 인 불안정시계열인 것으로 판정되었다. 따라서 이들 불안정시계열 간의 선형결합이 안정적인지 확인하기 위해 Johansen 공적분 검정을 실시하였으며 공적분 관계가 성립하는 것을 확인하였다.

변수 간의 순서를 정하기 위해서는 Granger 인과검정을 실시하였다. 인과검정 결과에 따르면 도소매업판매액지수가 산업생산지수를, 산업생산지수가 다시 비농림취업자수를 인과 하는 것으로 나타났다. 시차의 길이는 LR기준에 의해 5분기 시차를 선택하였다. 이상의 결과를 종합하여 4변수 VAR(5) 모형을 추정하였다.

VAR모형의 예측력을 평가하기 위해 2006년 1/4분기부터 2006년 3/4분기까지를 대상으로 평방근평균제곱오차(RMSE%)를 계산한 결과에 의하면 기존 분기GDP모형에 비해 VAR모형의 예측오차가 다소 낮게 나타났다. 따라서 GDP예측에 있어서 VAR모형이 분기GDP모형을 보완하는 역할을 할 수 있음을 시사하고 있다.

본 연구는 새로운 이론의 적용이나 독창적인 시도는 포함되어 있지 않다. 다만 그동안 통계청에서는 다루지 않았던 VAR모형을 이용하여 분기GDP 예측모형을 구축을 시도하는데 의의를 두어야 할 것이다. 앞으로 이를 기초로 하여 베이지안 VAR, 구조VAR 등 보다 진일보된 시계열모형을 구축하는 것은 향후 과제로 남긴다.

13) 최공필(1989) 참조

14) 변수 선정에 있어서 통계청 자료에만 국한하기 보다는 이자율, 물가, 통화 등과 같이 경제이론을 기초로 하여 분석이 가능한 변수에 대해서도 고려할 필요가 있을 것이다.

참고문헌

- [1] 김 양우 · 이 공희 · 장 동구(1997), “韓國의 短期經濟豫測시스템”, 『경제연구』 제3권 제3호, 한국은행금융경제연구소
- [2] 문 권순(1997), “벡터자기회귀(VAR)모형의 이해”, 『통계분석연구』 제2권 제1호, 통계청
- [3] 민 경삼 · 박 진호 · 박 순옥(2002), “월별 지표를 이용한 분기예측모형”, 『통계분석연구』 제7권 제2호, 통계청
- [4] 성 병희(2001), “Bayesian VAR 모형을 이용한 경제전망”, 『경제분석』 제7권 제2호, 한국은행
- [5] 안 충영 · 홍 성표 · 박 완규 공역(2004), 『제3판 기초 계량경제학』
- [6] 이 종원 · 이 상돈(2004), “結合方式을 利用한 豫測模型의 效率性 分析”, 『응용경제』 제1권 제1호, 한국응용경제학회
- [7] 이 충렬(1999), “超短期 經濟豫測 模型에 대한 연구”, 『産銀調査月報』 제527호, 한국산업은행
- [8] 최 공필(1989), “벡터自己回歸(VAR)方法에 의한 經濟豫測模型”, 『조사월보』 제56호, 대우증권