

그래프이론(DAG)을 이용한 국민소득과 에너지 간의 동시적 인과관계 분석*

박해선** · 오완근***

본 연구는 방향지시 비순환성 그래프(DAG)이론을 이용하여 국민소득과 에너지 간의 동시적 인과관계를 분석하였다. 분석은 1981:1~2004:4 분기 데이터와 GDP, 에너지, 자본, 노동의 다변량모형을 사용하였다. 오차수정모형과 DAG를 접목하여 실증분석한 결과, 국민소득은 에너지를 인과하는데 반하여 그 역은 성립하지 않음을 발견하였다. 또한 예측오차 분산분해도 동일한 결과를 보여 주었다. 이러한 실증분석 결과는 경제성장을 제약함 없이도 어느 정도 에너지 소비를 줄일 수 있음을 시사해 주는 것으로 그랜저 시차적 인과관계를 이용하여 장기에 동일한 일방향 인과관계를 발견한 Oh and Lee(2004b)와 일치하는 것이기도 하다.

핵심주제어: GDP, 에너지, DAG, 오차수정모형, GES

경제학문헌목록 주제분류: Q43

I. 서 론

국민소득과 에너지와의 관계에 관한 연구는 지난 수십 년간 크게 장단기 탄력성 분석과 인과관계 분석을 중심으로 이루어져 왔다. 특히, 지구온난화가 진행되고 기후변화협약이 전 세계 차원에서 논의되면서 최근 들어 경제성장과 에너지 소비 양자 간의 관계에 대한 연구는 더욱 관심을 끌고 있다.

그 동안 국민소득과 에너지 간의 관계는 여러 대상에 대해 다양한 방법론을

* 저자들은 본 논문에 대하여 유익한 논평을 주신 이성표 박사님과 익명의 두 분 심사위원님들께 깊은 감사를 드리며, 논문에 남아 있을지 모르는 오류는 전적으로 저자들의 책임임을 밝힌다. 이 논문은 2006학년도 한국외국어대학교 학술연구비의 지원에 의하여 이루어진 것임.

** 한국수출입은행 해외경제연구소 선임조사역(제1저자), 전화: (02) 3779-6679, E-mail: haesunp@gmail.com

*** 한국외국어대학교 경제학과 부교수(교신저자), 전화: (02) 2173-2333, E-mail: wanoh@hufs.ac.kr

논문투고일: 2006. 10. 16 수정일: 2006. 12. 19 게재확정일: 2006. 12. 20

사용하여 이루어져 왔지만 일치된 견해가 존재하기 보다는 여전히 다양한 결론들이 제시되고 있다. 이의 원인은 여러 가지가 있을 수 있으나 대표적인 것으로 데이터의 빈도(연도별 대 분기별), 에너지 집계방법(단순합계 대 디비지아), 시계열 방법론(VAR모형 대 오차수정모형), 모형에 포함된 변수 개수(이변량 대 다변량) 등을 들 수 있다.

연도별 데이터 사용 및 시계열 방법론상에서 VAR모형을 사용한 한계를 지니고 있는 연구로는 Kraft and Kraft(1978)를 들 수 있다. 이들은 1947~1974년의 연도별 미국 자료를 가지고 Sims기법을 사용한 결과 GNP가 총에너지 소비량에 영향을 주는 것으로 파악했지만, 이 연구는 같은 기법을 쓰더라도 분석기간만 달리 해도 결과가 달라진다는 점에서 비판받았다(Akarca and Long, 1980). 특히, Granger(1988)는 변수 간 공적분(cointegration)관계에 있을 경우 VAR를 이용한 그랜저 인과관계(standard Granger causality)검정 대신 오차수정모형(error correction model)을 써야 인과관계의 원천을 모두 파악할 수 있음을 보였다. 실제로 Glasure and Lee(1997)는 한국과 싱가포르를 대상으로 한 그랜저 인과검정에서 전통적인 그랜저 인과검정으로 찾아내지 못한 실질GDP와 에너지 간의 인과관계를 오차수정모형을 써서 찾아낸 바 있다.

다음으로 GDP와 에너지에 국한된 이변량(bivariate)모형의 한계를 지니고 있는 연구로는 Yu and Jin(1992), Masih and Masih(1996), Glasure and Lee(1997), Cheng and Lai(1997), Yang(2000), Oh and Lee(2004a) 등을 들 수 있다. 이들 연구는 오차수정모형에 기초하여 그랜저 인과관계 분석을 수행하였으나 두 개의 변수만을 사용하였다. 그렇지만 이들 이변량모형을 사용한 연구결과는 Lükepohl(1982)이 지적한 바와 같이 이변량모형에서 발견한 그랜저 인과관계는 다른 중요한 변수가 빠진 데 기인할 수 있다는 문제점을 지니고 있다.

이에 따라 상기의 문제점을 극복하기 위하여 최근에는 다변량 오차수정모형을 이용한 연구들이 진행되고 있다. 다변량 오차수정모형도 두 가지 유형으로 구분된다. Masih and Masih(1997, 1998)와 Asafu-Adjaye(2000)의 경우는 에너지 소비와 소득의 이변량에 소비자물가 변수를 추가하였고, Oh and Lee(2004b)는 소득과 에너지 소비는 동일하되 실질에너지가격 변수를 추가한 모형을 사용하였다. 반면, Stern(1993, 2000)¹⁾과 Oh and Lee(2004b)는 GDP, 자본, 노동, 에너지 데이터를 사용한 생산 측면의 접근방법을 시도하고 있다.

1) 이들 논문은 비록 다변량모형을 채택하였으나 오차수정모형에 근거한 인과관계가 아닌 전통적인 그랜저 인과관계 검정기법을 사용했다는 점에서 한계를 갖는다.

이들 인과관계를 분석한 기존 연구들의 공통점은 모두 그랜저 인과관계 검정법을 사용한다는 점이다. 비록 그랜저 인과관계가 현재 다양한 응용분야에서 광범위하게 적용되고는 있지만 인과방향을 결정하는 동 방법론은 그 자체의 단점이 있는데, 특히 영향을 미치는 다양한 채널의 존재가능성 및 일시적인 인과순위 결여가 있을 때 단점이 더 드러난다.²⁾ 즉, 그랜저 인과관계가 두 변수들 간의 직접적인 인과경로(direct causal path)에 대한 정보를 제공한다 하더라도 확률적 교란(stochastic disturbances)의 상관관계로 인해 여전히 존재할 수 있는 간접적인 인과경로는 설명하지 못한다.³⁾

또한 Hamilton(1994, p. 307)에서 지적하고 있는 바와 같이 그랜저 인과관계의 결과는 해석상에 다소 주의를 요한다. 예를 들어, 에너지가 GDP를 인과한다는 것은 에너지가 GDP를 늘이거나 줄인다는 것을 의미하기보다는 에너지가 GDP를 어느 방향으로 움직일지에 대한 최상의 정보를 반영한다는 것이다. 즉, 에너지변수가 GDP를 예측할 수 있는 변수라는 것을 의미한다. 이와 같이 그랜저 인과관계는 시간상의 인과성(time sequence causality)으로서 어떤 원인(cause)이 효과(effect)에 대해 시간적으로 선행(precede)함을 의미한다.⁴⁾

반면에 Pearl(2000)과 Hoover(2005) 등은 최근에 그랜저 인과관계 개념을 보완하거나 확장하는 인과관계 개념을 논의하고 있고, Pearl(2000)과 Spirtes *et al.* (2000)은 구체적으로 방향지시 비순환성 그래프(directed acyclic graphs, 이하 DAG) 분석법을 제시하고 있다. 본 논문에서 소개하고 적용하는 DAG 분석에 의존한 인과관계는 그랜저 인과관계와 달리 시간의 흐름에 따르지 않는 비대칭적 인과관계(non-time sequence asymmetry in causal relations)인 동시적 인과관계(contemporaneous causality)를 의미한다. DAG를 이용한 연구가 최근 활발히 진행되고 있는데, 예를 들면 Haigh and Bessler(2003)는 VECM-DAG모형을 이용하여 미국 내 주요 곡물시장 간 가격의 발견과정을, Bessler and Yang(2003)은 국가 간 주가지수의 변동과 상호 영향을 분석하였으며, Park, Mjelde, and Bessler(2006)는 미국 전기 현물시장 간의 가격동학에 대해 연구하였다.

이와 같이 그래프이론의 적용이 일종의 병렬적 형태를 가지는 가격변수의 인과관계 검증에 보다 많이 이루어졌다고 할 수 있지만 분기 또는 연간 표본빈도를 가진 서로 다른 형태의 변수에 대한 인과관계의 연구에 활용된 것도 있다.

2) Granger(1980); Hausman(2003) 참조.

3) Awokuse(2006), p. 594.

4) 박호경·윤원철(2006), p. 481.

예를 들면, Bessler and Lee(2002)는 미국의 연간 데이터를 활용하여 통화량과 물가의 인과관계를 분석하였고, Awokuse(2006)은 수출, 생산성, 실질교역조건, 자본, 산업생산지수의 분기별 데이터를 활용하여 수출과 GDP 간의 인과관계를 분석하였다. 또한 Bessler and Loper(2001)는 그래프이론을 통해 무역개방도, 자원수출, 평균수명 등의 연간 변수를 활용하여 개도국의 경제성장의 요인을 분석한 사례이며, Wang and Bessler(2003)도 분기별 가격, 소비 데이터를 활용하여 그래프이론을 적용한 사례이다.

본 연구는 다음과 같이 구성된다. 제I절 서론에 이어 제II절에서는 분석모형인 오차수정모형과 DAG를 소개하고, 제III절에서는 실증분석으로서 데이터를 설명하고, 단위근검정 및 공적분검정 결과와 오차수정모형 결과를 이용한 DAG 결과를 제시한다. 제IV장은 요약 및 결론이다.

II. 모 형

1. 오차수정모형 및 동시적 인과순위

GDP와 에너지와의 관계를 보는 시계열 다변량모형으로는 벡터자기회귀(vector autoregressive: VAR)모형이나 오차수정모형(vector error correction model: VECM)이 가장 많이 쓰인다.

수준에서의 VAR모형 (1)은 식 (2)와 같은 오차수정모형으로 표현할 수 있다.

$$Y_t = \mu + \Pi_1 Y_{t-1} + \dots + \Pi_k Y_{t-k} + \varepsilon_t, \quad t=1, 2, \dots, T, \quad (1)$$

$$\Delta Y_t = \mu + \Gamma_1 \Delta Y_t + \dots + \Gamma_{k-1} \Delta Y_{t-k+1} - \Pi Y_{t-k} + \varepsilon_t. \quad (2)$$

단, $Y = [GDP, E, K, L]$, $\Gamma_i = -1 + \Pi_1 + \dots + \Pi_i$,

$$i=1, \dots, k-1, \quad \Pi = 1 - \Pi_1 - \dots, \quad -\Pi_k$$

여기서, GDP: 국내총생산

E: 에너지

K: 자본

L: 노동

위 식에서 행렬 Π 의 cointegration rank를 추정함을 통하여 공적분관계의 존재 여부를 알아볼 수 있다. 세 가지 경우가 가능한데 ① $\text{rank}(\Pi)=0$ 으로 변수들

간에 장기적 관계가 없는 경우, ② $\text{rank}(\Pi) = p$, 즉 Y_t 의 모든 변수가 수준에서 안정적인 경우, ③ $0 < \text{rank}(\Pi) = r < p$: 이는 $\Pi = \alpha\beta'$ 을 만족하는 $p \times r$ 행렬 α 와 β 가 존재한다는 것을 의미한다. 여기서 β 는 공적분 벡터행렬로 해석되며, 설사 Y_t 자체는 불안정적이라도 $\beta'Y_t$ 는 안정적이라는 성질을 갖는다. α 는 오차수정파라미터들의 행렬이다.

그런데 VAR나 오차수정모형은 시계열 분석에 매우 유용하고 많이 쓰이는 도구이지만 추정할 파라미터의 개수가 많고 또한 이들을 직관적으로 해석하기가 쉽지 않아 모형을 통해 내생변수들에 대해 정보를 제공하는 데 한계가 있다. 따라서 대부분의 연구에서 외생적인 확률충격(innovation)이 내생변수에 미치는 영향 등을 살펴보기 위해 예측오차분산분해(forecast error variance decomposition) 등을 활용하게 된다.

그러나 이러한 방법을 활용하는데 있어 경제통계자료에서 흔히 발견되는 오차항 간의 이분산문제, 즉 오차항 간에 서로 동시 상관하지 않도록 오차항의 동시적 구조(contemporaneous structure)가 독립적(직교)이어야 한다는 점을 해결하기 위해 이른바 출레스키 분해(Choleski decomposition)가 이용된다. 일반적으로 오차항의 이분산문제의 해결은 다음과 같은 방법으로 이루어진다. 직교잔차항 벡터를 $e_t = A^{-1}\epsilon_t$ 로 정의할 경우, ϵ_t 는 직교화된 잔차항 벡터를 의미하고 A행렬은 비직교 잔차항(e_t)이 직교화된 잔차항과 어떤 관계를 갖고 있는가를 보여 주고 있다고 볼 수 있다. 행렬 A를 VAR모형에 곱함으로써 직교화된 오차항 벡터를 구하는 것이다. 즉, 식으로 표현하면 $Ae_t = \epsilon_t$ 과 같다. 출레스키 분해의 경우 A행렬은 역행렬이 존재하는 하방삼각행렬로서 대각행렬원소가 1로 이루어진다. 출레스키 분해방법이 활용될 경우, A행렬의 상위에 위치한 원소의 변수가 아래에 위치한 원소의 변수를 인과하는 구조를 갖게 된다. 따라서 이 방법은 변수의 나열순서에 따라 충격반응 및 예측오차분산분해의 결과가 달라지는 문제점을 안고 있는 한계가 있는 것으로 알려지고 있다.

이러한 변수의 임의나열문제를 해결하기 위해 Sims(1986)와 Bernanke(1986)는 VAR모형의 확률오차를 구조적으로 식별하기 위한 구조적 VAR모형(structural VAR model)을 제시한 바 있다. 구조적 VAR모형은 통계적으로 검증하고자 하는 경제이론이 잘 갖추어진 분야에서 매우 유용한 방법으로 활용될 수 있다.⁵⁾ 하지만 상이한 경제이론이 존재하거나 연구자의 주관적인 경험과 판단이 일치

5) 김영덕(2005)은 구조 VAR모형을 이용하여 우리 나라에서 유가의 변동이 국내시장과 수출 시장을 통하여 어떻게 전달되는지를 설명하고 있다.

하지 않을 경우, 적용되는 인과순위가 연구자에 따라 달라질 수 있으며 이에 따라 상이한 충격반응 분석결과를 낳는 한계를 여전히 갖고 있다.

Swanson and Granger(1997)는 VAR모형의 식별과정에서 변수의 임의적인 나열의 문제점과 구조적 VAR모형이 갖고 있는 한계점을 극복하기 위한 방법으로 그래프이론이 유용하게 활용될 수 있음을 보여 주었다. Swanson and Granger(1997)는 VAR모형 내의 동시적 인과순위에 관한 유용한 정보가 VAR모형 잔차항의 공분산행렬에 담겨져 있다고 보고, DAG(directed acyclic graph)를 활용하여 Bernanke(1986)가 제시한 구조모형의 인과순위를 제공할 수 있다는 것을 보여 주었다. 또한 Hoover(2005)는 VAR모형의 동시적 인과순위를 파악하는 데 있어 그래프를 활용한 방법이 몇몇 문제점에도 불구하고 유용한 접근법이 될 수 있다고 주장하고 있다.

또한 그래프이론을 활용한 인과순위 결정방법은 특히 사전적으로 뚜렷한 인과관계를 정하기 어려운 경우, 인과관계 설정은 연구자의 주관적인 판단에 의존할 수밖에 없다는 문제점을 통계자료에 근거하여 동시적 인과순위(causal ordering)를 해결할 수 있기 때문에 분석의 객관성을 제고할 수 있다. 즉, 그랜저 인과관계 분석이 기간(분기) 간의 시차를 둔 반면, DAG를 활용한 인과관계 분석은 기간(분기) 내의 동시적 인과관계를 보여 준다는 점에서 의미가 있다. 또 그래프이론을 활용할 경우 VECM의 충격반응함수와 예측오차분산분해에서 변수의 오차항 간의 동시적 인과관계를 적용하여 변수의 임의적인 나열보다는 자료로부터 구할 수 있는 동시적 인과관계를 적용함으로써 분석의 객관성을 제고할 수 있는 장점이 있어 최근 시계열 분석에 도입되기 시작하여 연구가 활발히 진행되고 있다.

2. 그래프이론

(1) 방향지시 비순환성 그래프

그래프이론(directed acyclic graphs)⁶⁾의 기본적인 개념을 소개하는 차원에서 본항에서 간략히 설명하고자 한다. 다시 말해, 방향지시 그래프는 변수들 간의 인과흐름(causal flow)을 나타내 주는 그림이라고 할 수 있다. 화살표는 그러한 흐름을 표시하는 것으로 사용된다. 방향지시 화살표는 연결선 또는 모서리(edge)

6) 그래프이론에 대한 보다 자세한 내용은 Pearl(2000) 및 Park, Mjelde, and Bessler(2006)와 박호정·윤원철(2003, 2006)을 참조하기 바란다.

라고 부르며 한 연결선을 두고 연결된 변수는 인접(adjacent)하다고 말한다. 예를 들어, 간단한 방향지시 그래프(directed graph)인 $X \rightarrow Y$ 는 “변수 X 는 변수 Y 를 인과한다(cause)”라는 것을 의미한다. 만약 두 변수가 $W-X$ 와 같이 연결되어 있을 경우에는 W 와 X 는 정보흐름(information flow)에 의해 연결되어 있다고 할 수 있다. 그러나 방향지시 화살표(arrow)가 아닌 선분으로 연결된 위의 경우, W 변수가 X 변수를 인과하는지 그 반대인지가 구별되지 않음을 의미한다. DAG에서는 $W \rightarrow X \rightarrow Y \rightarrow W$ 와 같이 한 변수(W)에서 생성된 정보가 다른 변수(X, Y)를 거쳐 결국 변수 W 로 돌아오는 순환성 체계는 포함되지 않는다.

만약 두 변수(예를 들어, A 시장과 B 시장의 확률충격)가 상관관계에 있다면 A 가 B 를 인과하는지 B 가 A 를 인과하는지 또는 제3의 변수 C 에 의해 두 변수(A, B)가 인과되는지를 구별할 수 없다. 이러한 인과흐름을 파악하기 위한 시도는 세 개 이상의 변수 간에 인과관계(causal relationships)가 조건부 독립조건(conditional independence conditions)의 활용에 의해 밝혀질 수 있다는 것으로부터 출발한다.

예를 들어, 그래프 $A \leftarrow B \rightarrow C$ 를 고려해 보자. 이 그래프는 B 가 A 와 C 에 대한 공통원인(common cause)이 되는 경우를 표시한다. 이 경우 두 변수 A 와 C 간의 무조건부 연관성(unconditional association)은 0이 아니다. 그러나 B 라는 공통원인에 대한 정보 또는 지식이 주어지면 A 와 C 의 조건부 관계는 0이 된다. Pearl(2000)은 이러한 구조를 특히 “인과분기(causal fork)”라고 언급하고 있으며, 이 경우 공통원인이 A 와 C 변수 간의 연관성을 “차단(screen-off)”하고 있다고 말한다. “차단”은 변수 간의 인과흐름을 결정짓는데 중요한 역할을 하는 개념으로 이용되고 있다. 다른 유사한 차단조건이 존재하는 경우로 “인과사슬(causal chain)”관계, 즉 $D \rightarrow E \rightarrow F$ 의 예를 고려해 보기로 하자. 이 경우 D 와 F 변수 간의 무조건부적 연관성은 0이 아니지만 만약 중간에 위치한 E 의 정보가 주어진다면 D 와 F 간의 조건부적 연관성은 0이 된다. 왜냐하면, 인과사슬의 중간변수에 대한 정보가 다른 두 변수 간의 연관성을 “차단”하기 때문이다.

다음으로 변수 G, H 와 I 간의 인과관계가 $G \rightarrow H \leftarrow I$ 로 설정되어 있는 경우를 고려해 볼 수 있다. Pearl(2000)은 이 경우를 특히 “인과역분기(causal inverted fork)”이라고 언급하고 있다. 이 경우 H 는 충돌부(collider)로 불리며 G 와 I 의 공통효과(common effect)라고 할 수 있다. 이 경우는 앞서의 두 예와 달리 무조건부적 연관성은 0이나 만약 H 의 정보 또는 지식이 주어질 경우 G 와 I 의 연

관성은 0이 되지 않는다. 즉, 공통결과에 대한 정보는 다른 두 변수(G, I) 간의 연관성을 “차단”하지 못한다.

요약하면 인과분기와 인과사슬은 중간 변수가 양쪽 끝 두 변수 간의 연관성을 차단한다는 측면에서 동일한 독립적 구조를 지니고 있다고 할 수 있지만, 인과역분기의 경우에는 중간변수가 다른 두 변수 간의 연관성을 차단하지 않는다. 이러한 인과역분기의 존재는 연구자로 하여금 이론적으로 관련성이 있는 변수들 간의 독립성 또는 조건부적 독립성을 활용하여 인과구조를 파악하는 것을 가능하게 해 준다는 것이다.

이러한 공통원인, 인과사슬, 공통결과와 관련된 “차단”현상은 오래 전부터 문헌에서 인식되어져 왔다(Orcutt, 1952). 그러나 차단개념이 세 개 이상의 변수들 간의 인과흐름을 파악하는데 본격적으로 활용되기 시작한 것은 비교적 최근 들어 Pearl(2000)과 Spirtes, Glymour, and Scheines(2000)에 의해 이루어졌다. Pearl과 몇몇 연구자들은 차단개념을 인과흐름과 확률표현(probability representation) 간의 관계를 설정하는 d-분리(d-separation) 개념을 활용하여 일반화하였다. Pearl(2000)은 두 변수 A 와 C 간의 모든 경로를 B 가 차단하면 “ B 는 A 와 C 를 d-분리한다”고 정의한다.

(2) 탐욕동급검색 알고리즘

DAG를 찾기 위한 탐색 알고리즘은 Pearl(2000)과 Spirtes, Glymour, and Scheines(2000)에서 몇 가지 방법을 소개하고 있는데, 그 중 가장 널리 활용되는 알고리즘으로 PC 알고리즘과 탐욕동급검색(greedy equivalence search, 이하 GES) 알고리즘을 들 수 있다. Park, Mjelde, and Bessler(2006)와 박호정·윤원철(2003, 2006)은 PC 알고리즘을 활용한 DAG를 연구에서 활용하였다. PC 알고리즘은 세 가지 가정, 즉 인과충분성(causal sufficiency) 조건, 마르코프(Markov) 조건 및 진정성(faithfulness) 조건⁷⁾이 충족된다는 가정하에 앞서 언급한 d-분리의 개념을 활용하여 변수 간의 상관계수와 편상관계수(partial correlation)를 이용하여 변수들 간의 인과관계를 나타내는 DAG를 찾게 된다.⁸⁾

그러나 실제의 많은 연구에 있어 고려할 수 있는 변수가 충분치 않은 경우가

7) 박호정·윤원철(2003)에서는 PC 알고리즘이 잘 작동하기 위한 전제조건으로 인과충분성과 마르코프 조건만 언급되어 있으나 진정성(faithfulness) 조건도 충족되어야 한다(Bessler and Lee, 2002).

8) PC 알고리즘에 대한 보다 자세한 논의는 Pearl(2000); Spirtes, Glymour, and Scheines(2000); Park, Mjelde, and Bessler(2006); 박호정·윤원철(2003, 2006)을 참조하기 바란다.

많은데 이러한 경우에 있어, “분석에 포함된 두 개 이상의 변수를 인과하는 변수는 반드시 분석에 포함되어야 한다”는 충분성 조건은 현실적으로 PC 알고리즘을 적용하는데 제약이 될 수 있다. 또한 PC 알고리즘은 표준적인 Neyman-Pearson 가설검정에 기초하고 있기 때문에 유의수준(significance level)이 적절히 선택되어야 한다. 유의수준의 선택에 따라 DAG의 결과도 달라지기 때문이다. 반면에 GES 알고리즘은 PC 알고리즘과 같이 엄격한 가정을 필요로 하지 않을 뿐만 아니라 베이저안 평가점수에 의해 DAG를 탐색하게 되므로 적절한 유의수준의 선택이라는 문제에서 자유로울 수 있는 장점을 지니고 있다. 본 연구에 선 Meek(1997)에 의해 제시되고 Chickering(2002)에 의해 보다 구체화된 베이저안 검색 알고리즘인 GES 알고리즘에 대해 간략히 소개하는 차원에서 설명하기로 한다.

GES 알고리즘은 간단히 베이저안 평가기준(Bayesian scoring criterion)을 활용하여 가능한 DAG들을 평가하는 방법이라고 할 수 있다. Schwarz에 의한 베이저안 평가기준의 근사치는 다음과 같이 표시된다.

$$S(G^K, D) = \log p(D | \hat{\phi}, G^K) + \frac{1}{2} \log T. \quad (3)$$

식 (3)에서 p 는 D 로 표시되는 통계자료를 생성하는 확률분포를 의미하고, $\hat{\phi}$ 는 알려지지 않은 모수의 최우도(maximum-likelihood) 추정치를 의미하며, d 는 그래프 G^K 의 자유모수(free parameters)의 개수를, 그리고 T 는 관측치의 개수를 의미한다. 함수 S 는 $\log p(\hat{D} | \hat{\phi}, G^K)$ 항으로 표시되는 적합도(fit)와 $\{(d/2)\log T\}$ 로 표시되는 자유모수의 최소화 간의 trade-off 관계를 보여 주고 있다.

GES 알고리즘을 설명하기 위해 먼저 동급(equivalence class)개념을 이해할 필요가 있다. 만약 두 개의 DAG가 서로 동일한 골격(skeleton)과 동일한 v -구조(structure)를 갖고 있으면 이들 두 DAG는 동일한 동급에 속한다고 할 수 있다. 여기서 골격이란 방향성을 무시한 변수 간의 연결성만을 의미하며, v -구조는 두 개의 변수가 한 다른 변수를 인과하는 구조, 즉 앞서 언급한 인과역분기의 구조를 의미한다. 예를 들어, 다음의 DAG들을 생각해 보자.

$$\begin{aligned} (D.1) \quad & A \leftarrow B \rightarrow C, & (D.2) \quad & A \leftarrow B \leftarrow C, & (D.3) \quad & A \rightarrow B \rightarrow C, \\ (D.4) \quad & A \rightarrow B \leftarrow C, & (D.5) \quad & A \leftarrow C \rightarrow B \end{aligned}$$

(D.1), (D.2) 및 (D.3)은 동급에 속한다고 할 수 있다. 골격이 서로 동일하며 v -

구조에 있어 인과역분기가 전혀 없다는 점에서 동일한 v -구조를 갖고 있기 때문이다. 이에 반해 (D.4)는 이들과 동일한 골격은 갖고 있으나 하나의 인과역분기를 갖고 있는 다른 v -구조를 보이고 있기 때문에 (D.1), (D.2) 및 (D.3)과 함께 동급에 속하지 않는다. 한편, (D.5)는 (D.1), (D.2) 및 (D.3)과 같이 인과역분기가 전혀 없는 v -구조를 갖고 있으나 골격에서 다른 차이를 보이고 있으므로 (D.1), (D.2) 및 (D.3)과 같은 동급에 속하지 못한다. 즉, 다른 동급에 속한다고 할 수 있다.

GES 알고리즘은 어느 한 동급에 속한 DAG들과 다른 동급에 속한 DAG들을 단계적으로(stepwise) 비교 검색하게 된다. 그러나 GES 알고리즘은 같은 동급에 속해 있는 DAG들에 대해서는 평가의 우열을 구분하지 않는다.

GES 알고리즘은 두 단계로 나뉘는데, 첫 번째 단계에서는 먼저 모든 변수들 간의 모서리가 없는(모든 변수들이 상호 독립적인) 상태에서 출발하여 모서리를 추가하고 방향지시 화살표의 방향을 변경해 가면서 베이저안 사후적 평가점수(Bayesian posterior score)가 이러한 변화에 따라 개선되어 가는지를 살피게 된다. 이러한 과정에서 알고리즘은 변수 간 모서리의 추가 또는 화살표 방향의 변화에 따라 베이저안 평가점수(Bayesian score)를 가장 많이 증가시키는 변화를 선택하게 되는데 첫 번째 단계는 더 이상 모서리의 추가나 화살표 방향의 변화가 베이저안 사후적 평가점수를 개선시키지 못할 때 종료된다. 이로부터 출발하여 두 번째 단계에서는 첫 번째 단계에서 구해진 DAG 후보로부터 모서리를 하나씩 제거하고 화살표의 방향을 역으로 변화시키는 방법을 순차적으로 적용하여 베이저안 사후적 평가점수를 가장 많이 증가시키는 방향으로의 변화를 선택하는 식으로 알고리즘이 작동된다. 두 번째 단계에서 첫 번째 단계와 마찬가지로 알고리즘은 더 이상 베이저안 사후적 평가점수를 개선시킬 수 없을 때 중단되고, 이 때의 DAG가 최종적으로 탐욕검색 알고리즘으로 탐색된 최적 DAG라고 할 수 있다. 이상의 GES 알고리즘은 PC 알고리즘과 같이 Spirtes를 비롯한 연구진들이 개발한 소프트웨어 TETRAD IV에 의해 간단히 수행된다.

Ⅲ. 실증분석

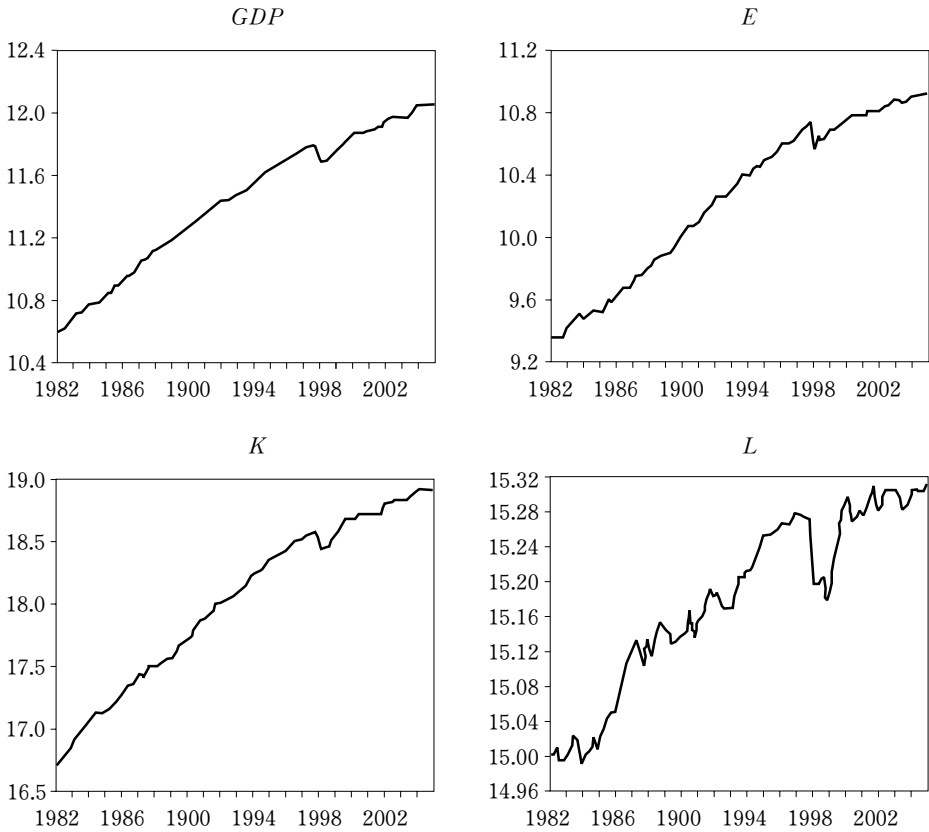
1. 모형 및 데이터

모형은 Stern(1993, 2000)과 Oh and Lee(2004b)에서와 같이 실질GDP(GDP), 에너지(E), 자본(K), 노동(L)의 4개 변수 다변량모형을 채택하였다. 먼저 실질GDP는 한국은행이 작성한 계절조정 실질GDP(2000년 불변, 10억 원)이다. 에너지와 경제성장 간의 초기연구에는 실질GNP가 더러 사용되기도 하였으나, 에너지 소비가 국내에서의 생산활동에 쓰이기 때문에 실질GDP를 사용하는 것이 논리적으로 더 타당하다고 판단된다.

에너지 소비는 에너지경제연구원에서 발표한 1차 에너지 소비량(석유 환산톤)이다. 1차 에너지에는 석유, 석탄, 액화천연가스(LNG), 수력, 원자력, 신탄 등이 포함된다. 자본투입량은 Pyo and Ha(2006)의 실질자본 스톡에다 통계청에서 구한 계절조정 제조업 가동률 지수를 곱하였다. 노동은 산업별 취업자수에 근로시간을 곱한 총노동시간을 썼다.

분석대상기간은 1981년 1분기에서 2004년 4분기까지의 분기별 데이터인데 이는 1981년 이전은 에너지 소비량 분기 데이터를 구하기 어려워 분석대상기간에 포함시키지 못하였으며, 분기별 실질자본 데이터는 현재 2004년 4/4분기까지만 존재하기 때문이다. 분기별 데이터를 사용한 이유는 분기 데이터를 사용할 경우 무엇보다도 연간 데이터를 사용할 경우에 발생할 수 있는 변수 간 단기변동이 상쇄되는 단점을 보완할 수 있는 장점이 있기 때문이다. 또한 연간 데이터를 사용할 경우 표본수가 많지 않아 검정력이 떨어지는 문제도 완화된다. Yu and Hwang(1984)에서 같은 방법론을 적용했음에도 불구하고 연간 데이터를 썼을 때 발견되지 않은 에너지와 GDP 간의 인과관계가 분기 데이터를 쓴 경우에 나타났다는 점에서 연간 데이터보다는 분기 데이터가 더 나음을 시사한다고 볼 수 있다(Stern, 1993). 또한 분기자료를 사용하면 분기단위의 동시적 인과관계를 분석할 수 있는 장점이 있다.

그리고 계절조정이 된 실질GDP와 자본을 제외한 모든 시계열 데이터값은 $X-12$ 로 계절조정을 하고 모든 값은 자연대수를 취하였다. <그림 1>은 이를 나타낸 것이다. 네 변수 모두 외환위기의 영향을 보여 주고 있어 추정식에서 1997년 4분기부터 1998년 4분기까지는 1, 나머지는 0인 외환위기 더미변수를 포함하였다.



〈그림 1〉 연도별 실질GDP, 에너지, 자본, 노동의 로그값

2. 단위근검정

먼저 시계열 변수들의 안정성 여부를 알아보기 위해 ADF 단위근검정을 실시하였다. 단위근검정 결과는 <표 1>에 나타나 있다. 검정시 확정적 시간 추세항(deterministic time trend)과 상수항을 포함하는 경우와 상수항만 포함하는 경우 모두를 검정하였다. 최적시차수를 정하는 데에는 Schwert(1987)를 따랐다. 그 결과 최적시차는 모두 3으로 나타났다.

검정결과 모두 수준에서는 단위근이 존재하지만 1차 차분한 경우는 단위근이 없는 것으로 나타났다. 즉, 모든 변수가 수준에서는 불안정적이지만 1차 차분에서는 안정적인 $I(1)$ 으로 나타났다.⁹⁾

9) 일반적으로 불안정한 계열을 수준변수로 추정할 경우 가성회귀의 문제가 나타난다는 것은

〈표 1〉 ADF 단위근검정

변 수	시 차 수	상수항만 포함시		상수항과 시간추세항 포함시	
		검정통계량	임계치(5%)	검정통계량	임계치(5%)
<i>GDP</i>	3	-2.5056	-2.8935	-1.2123	-3.4599
<i>E</i>	3	-2.3210	-2.8935	0.1830	-3.4599
<i>K</i>	3	-2.5823	-2.8935	-1.3188	-3.4599
<i>L</i>	3	-0.9361	-2.8935	-2.2738	-3.4599
ΔGDP	3	-3.9683	-2.8935		
ΔE	3	-5.4194	-2.8935	-	-
ΔK	3	-4.4472	-2.8935		
ΔL	3	-5.5674	-2.8935		

주: *Y*=실질GDP, *E*=1차 에너지, *K*=자본, *L*=노동.
 Δ 는 1차 차분이며, 임계치는 McKinnon(1991)임.

3. 공적분검정 및 오차수정모형 추정

대상 변수들이 모두 단위근을 갖는 불안정한 시계열로 드러남에 따라 공적분 관계를 찾아본다. 공적분검정 방법에는 여러 가지가 있다. Gonzalo(1994)는 공적분을 검정하는 주요 다섯 가지 방법을 비교·검토한 후 이들 가운데서 Johansen and Juselius(1990)가 제시한 최우추정방법(full information maximum likelihood estimation: FIML)이 여러 가지 면에서 가장 우월한 것임을 보이고 있다. 이에 기초하여 본 연구에서는 최우추정방법을 사용한다.

나아가 Johansen(1992)은 공적분계수(階數)와 선형추세를 동시에 결정하는 일반적인 가설검정법을 제안하고 있다. Johansen(1992)은 Trace검정을 이용하여 $r=0, 1, \dots, n, p-1$ 에 대해서 순차적으로

$$\begin{aligned}
 H_r^* & \text{ 수락 if } T_r^* < C_r^*(\alpha), \\
 H_r & \text{ 수락 if } T_r^* \geq C_r^*(\alpha) \text{ and } T_r < C_r(\alpha), \\
 H_{r+1} & \text{ 수락 if } T_r^* \geq C_r^*(\alpha) \text{ and } T_r \geq C_r(\alpha).
 \end{aligned}$$

여기서, H_r : 최대 r 개의 공적분벡터와 선형추세모형

H_r^* : 최대 r 개의 공적분벡터와 선형추세가 없는 모형

대립가설 $H_p = H_p^*$: p 개의 공적분벡터모형

잘 알려진 사실이다. Granger and Newbold(1974); Phillips(1986) 등 참조.

T : 검정통계량
 C : 임계치
 α : 유의수준

물론 만일 $r=0, \dots, p-1$ 에 대해서 모든 가설이 기각되면 우리는 모든 수 준변수가 안정적이라고 결론 내릴 수 있다.

공적분검정 전에 최적시차수의 결정이 필요하다. Geweke and Meese(1981)는 Schwarz Information Criterion(SIC)이 다른 시차수 결정방법에 비해 우월하다는 것을 보여 주고 있다.¹⁰⁾ 그들의 연구결과에 기초하여 SIC를 사용하였으며, 최대차수는 5로 하였다. <표 2>에서 보듯이 시차수를 달리할 경우의 SC값들을 비교한 결과 최적시차수는 2로 나타났다.

공적분검정 결과는 <표 3>에 나타나 있다. 여기서 T^* 와 C^* 는 각각 귀무가설 H_0^* : 최대 r 개의 공적분벡터와 선형추세 없는 모형에 대한 검정통계량과 임계치이다. 마찬가지로 T 와 C 는 각각 귀무가설 H_r : 최대 r 개의 공적분벡터와 선형추세모형에 대한 검정통계량과 임계치이다.

<표 3>의 Trace검정은 유의수준 5%에서 공적분벡터수가 최대 1이라는 귀무

<표 2> 최적시차 결정을 위한 SC값

SC값			
시차수 2	시차수 3	시차수 4	시차수 5
-20.55	-20.15	-19.75	-19.20

<표 3> 공적분검정 결과

공적분 계수	T^*	$C^*(5\%)$	T	$C(5\%)$
$r=0$	114.61	54.07	67.21	47.85
$r \geq 1$	52.60	35.19	36.29	29.79
$r \geq 2$	27.47	20.26	12.78	15.49
$r \geq 3$	4.90	9.16	4.68	3.84

주: 임계치 출처는 Osterwald-Lenum(1992).

10) 또한 Lütkepohl and Breitung(1997)은 AIC는 점근적(asymptotic)으로 시차를 과대 추정하는 경향이 있는 반면, SIC는 필요한도 내에서 최소한의 시차로 파라미터를 추정토록 한다고 주장하였다. 즉, AIC는 최적시차를 과대하게, SIC는 최적시차를 가장 최소로 결정한다는 것이다.

〈표 4〉 오차항의 분산-공분산의 상관계수행렬

	<i>GDP</i>	<i>E</i>	<i>K</i>	<i>L</i>
<i>GDP</i>	1			
<i>E</i>	0.6436	1		
<i>K</i>	0.7124	0.4965	1	
<i>L</i>	0.5642	0.4822	0.3710	1

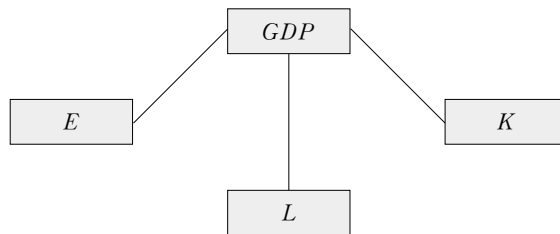
가설들(H_1^* 와 H_1)을 기각한다. 그리고 귀무가설 H_2^* : 최대 2개의 공적분벡터와 선형추세 없는 모형도 기각한다. 그러나 H_2 : 최대 개의 공적분벡터와 선형추세모형은 기각하지 못한다. 따라서 실질GDP, 에너지, 자본, 노동의 네 개의 변수 간 공적분벡터수는 2개이며, 이들 변수 간에는 장기적으로 안정적인 관계가 있음을 알 수 있다.

한편, 이로부터 구해지는 오차수정모형은 변수들 간의 동태적 관계를 나타내는 것이지만 추정된 모든 계수에 대한 해석이 쉽지 않기 때문에 변수들 간의 예측오차분산분해를 살펴볼 필요가 있다. 이에 따라 사전적으로 인과순위를 정하지 않도록 DAG를 사용하여 동시적 인과관계 구조를 파악한다. 이를 위해 오차수정모형으로부터 구한 오차항의 분산-공분산행렬로부터 구한 상관계수행렬이 〈표 4〉에 제시되어 있다.

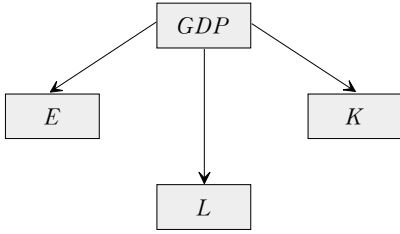
상관계수행렬에서 노동과 자본의 상관관계가 가장 약한 반면, GDP와 자본 간의 상관관계가 가장 큰 것으로 나타났다. 또한 GDP와 에너지 변수 간의 상관관계도 큰 것을 보여 주고 있다.

4. 방향지시 비순환성 그래프

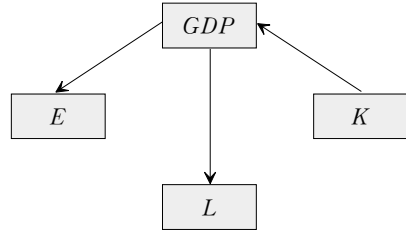
GES 알고리즘을 활용한 동시적 인과관계를 보여 주는 DAG는 〈그림 2〉와



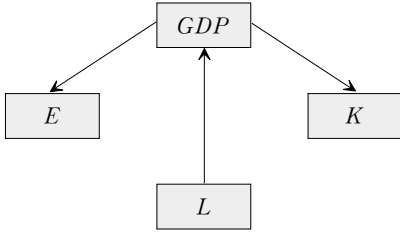
〈그림 2〉 GES 알고리즘을 활용한 DAG 결과



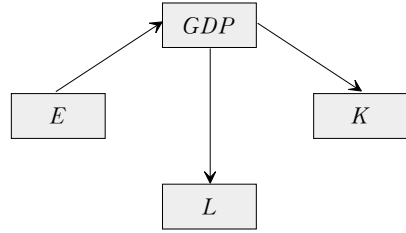
〈그림 3〉 DAG 1



〈그림 4〉 DAG 2



〈그림 5〉 DAG 3



〈그림 6〉 DAG 4

같이 나타났다.

GES 알고리즘은 DAG의 방향성을 제시하지 못하고 있다. 이는 동급에 속한 〈그림 3〉, 〈그림 4〉, 〈그림 5〉, 〈그림 6〉의 DAG 간의 우열을 가리지 못하고 있다는 것을 의미한다. 즉, DAG 1, 2, 3 및 4는 동일한 골격과 v -구조를 가진 동일한 동급에 속해 있으므로 GES 알고리즘은 〈그림 2〉에서처럼 방향성을 제시하지 못하고 있는 것이다. 4개의 DAG 중 최적의 DAG를 찾기 위해 Wang and Bessler(2005)와 Park, Mjelde, and Bessler(2006)에서 사용된 Seemingly unrelated regression과 변형된 Schwartz information criteria를 활용한 방법을 적용하기로 한다. 즉, 추정된 오차수정모형은 4개의 변수를 각각 종속변수로 하는 4개의 방정식으로 구성되는데 이들 4개의 식으로부터 각 변수의 확률충격(innovation) 시계열을 구할 수 있다. 이렇게 구해진 각 변수의 확률충격을 활용하여 각각의 DAG에 해당하는 4개의 SUR 체계를 추정하고 이들 4개의 SUR 추정에서 얻게 되는 분산-공분산행렬을 이용하여 SIC를 구하여 이를 비교하여 최적의 DAG를 탐색한다. 예를 들어, DAG 1의 경우 GDP가 에너지와 노동, 자본을 인과하는 관계를 보여 주고 있는데 이 경우 SUR은 다음 4개의 식으로 구성된다. GDP식으로부터의 확률충격을 종속변수로 하고 상수항을 독립변수로 하는 한 개의 식과, 에너지식으로부터의 확률충격을 종속변수로 하고 GDP 확률충격과 상수항을 독립변수로 한 식, 그리고 노동식으로부터의 확률충격을 종속변수로 하고

GDP 확률충격과 상수항을 독립변수로 한 식 및 자본식으로부터의 확률충격을 종속변수로 하고 GDP 확률충격과 상수항을 독립변수로 한 식으로 구성된다. 종속변수를 인과하는 독립변수로 구성된 식으로 구성된 4개의 SUR체계가 SIC에 의거 비교된다.

이러한 방법을 적용한 결과, <그림 5>의 “DAG 3”가 가장 적합한 DAG인 것으로 나타났다. DAG 3은 동시적인(contemporaneous) 관점에서 GDP가 자본과 에너지를, 노동이 GDP를 인과하는 흐름을 보여 주고 있다. GDP에서 에너지로의 인과관계는 있지만 그 역은 성립하지 않는 일방향 인과관계가 존재한다는 연구결과는 생산요소로서의 에너지가 차지하는 비중이 매우 낮은 것과 일맥상통하는 것으로 판단된다. 이러한 결과는 시차적 인과관계를 의미하는 그랜저 인과관계 검정법을 이용하여 양자 간에 동일한 일방향 인과가 있음을 발견한 Oh and Lee(2004b)의 결과와 일치하는 것으로 본 연구의 실증분석 결과는 경제 성장을 제약함이 없이도 어느 정도 에너지 소비를 줄일 수 있음을 시사해 주는 것이다.¹¹⁾

한편, GDP와 자본 및 노동과의 관계는 경제가 노동투입을 늘려 산출을 늘리지만, 자본의 경우는 산출을 늘리기 위해 먼저 가동률을 최대한 높이고 난 후에 추가적인 산출이 필요하면 투자를 하여 자본이 축적되는 현상과 연관된 것으로 판단된다.

5. 예측오차분산분해

앞의 DAG에서 구해진 인과관계를 사용하여 변수 간의 단기구조를 분석하기 위해 예측오차분산분해를 구하였다. 예측오차분산분해는 동태적으로 어떤 변수가 외생적인지 아니면 내생적인지를 판단할 수 있는 정보를 제공해 준다. 즉, 어떤 변수의 예측오차분산의 대부분이 그 변수 자체의 확률오차에 의해 설명될 경우 그 변수는 통계적으로 외생적인 것으로 간주된다. 또한 예측오차분산분해는 시간의 경과에 따라 어떤 변수에 대한 타 변수의 설명력의 변화과정에 대한 정보를 제공해 준다. <표 5>는 DAG의 동시적 인과관계를 이용하여 구해진 예측오차분산분해를 보여 주고 있다.

<표 5>에서 GDP의 예측오차는 1의 시차에서 노동(L)에 의해 31.83%, 자체

11) 추가적 분석을 통해 이를 보다 명확히 논리적으로 설명할 필요가 있다는 심사자의 지적이 있었다.

〈표 5〉 예측오차분산분해

시 차		<i>GDP</i>	<i>K</i>	<i>L</i>	<i>E</i>
<i>GDP</i>	1	68.17	0.00	31.83	0.00
	5	59.70	11.62	28.57	0.11
	10	54.23	14.10	30.38	1.29
	15	52.27	13.29	32.08	2.37
<i>E</i>	1	28.25	0.00	13.19	58.56
	5	24.79	8.05	11.95	55.21
	10	36.62	12.87	21.26	29.25
	15	43.42	12.37	27.86	16.35
<i>K</i>	1	34.60	49.25	16.16	0.00
	5	40.82	27.95	26.97	4.26
	10	42.52	19.03	31.88	6.57
	15	44.25	15.38	33.77	6.60
<i>L</i>	1	0.00	0.00	100.00	0.00
	5	2.44	4.73	67.54	25.29
	10	29.99	21.60	34.99	13.41
	15	46.42	18.16	30.77	4.66

GDP에 의해 68.17% 설명되고 자본과 에너지에 의한 설명력은 거의 없다. 그러나 시차가 늘어날수록 자본(*K*)과 에너지의 설명력이 다소 증가하나 여전히 자체 GDP(52.27%)와 노동(32.08%)의 설명력이 높은 수준에 머물러 있음을 알 수 있다.

에너지(*E*)는 GDP에 비해 훨씬 내생적 변수의 경향이 강한 것으로 나타나고 있다. 즉, 에너지의 자기변수 설명력은 시차 1의 경우 58.56%로 비교적 내생적 성향이 큰 것으로 판단된다. 시차가 진행될수록 에너지에 대한 자체 설명력은 크게 감소하고, GDP의 설명력이 43.42%(시차 15)로 가장 큰 비중을 차지하며, 노동의 설명력 또한 27.86%(시차 15)로 커짐을 알 수 있다. 이러한 결과는 DAG에서 GDP가 에너지를 인과했던 결과와 일치하는 것이다.

자본(*K*)도 GDP에 비해 내생적 변수의 경향이 강한 것으로 나타나고 있다. 즉, 자본의 예측오차는 시차 1에 있어 자체 변수(49.25%) 이외에 GDP(34.60%), 노동(16.16%)에 의해 각각 설명된다. 시차가 증장기로 진행될수록 자체 변수의 설명력보다 GDP와 노동에 의한 설명력이 커지는 특징을 보여 주고 있다.

노동은 다른 변수와 달리 시차 1에 있어 자기변수의 설명력이 100%를 보이며 가장 외생성이 높은 변수인 것으로 나타났다. 이러한 결과는 노동이 DAG에

서 유일하게 외생적 변수로 나타난 결과와 일치하는 것으로 볼 수 있다. 시차 5의 단기의 경우 자체 변수 설명력은 67.54%로 낮아진 반면, 노동에 대한 에너지의 설명력이 급격히 증가한 후 시간이 흐름에 따라 설명력이 감소하는 특징을 보여 준다. 노동은 중장기에 있어 대체적으로 다른 변수들에 대한 설명력이 높아지는 흥미로운 모습을 보여 주고 있다.

VI. 요약 및 결론

그 동안 국민소득과 에너지 간의 관계를 살펴본 많은 연구들은 연도별 데이터를 사용하기도 하고, 국민소득이나 에너지 변수들이 수준에서는 불안정적인 변수임에도 불구하고 VAR모형을 이용하였으며 국민소득과 에너지 이변량모형을 사용하는 등 여러 가지 한계를 지니고 있었다. 이에 대한 대안으로 최근 들어서는 다변량 오차수정모형을 이용하여 그랜저 인과관계를 살펴보는 것이 가장 많이 사용되고 있다.

인과관계를 분석한 기존 연구들의 공통점은 모두 그랜저 인과관계 검정법을 사용하고 있다는 점이다. 그렇지만 그랜저 인과관계는 두 변수들 간의 직접적인 인과경로에 대한 정보를 제공한다 하더라도 확률적 교란의 상관관계로 인해 여전히 존재할 수 있는 간접적인 인과경로는 설명하지 못하는 문제점을 지니고 있다. 즉, 변수들에 영향을 미치는 채널이 다양하게 존재하면 올바른 결론을 도출하지 못할 가능성이 존재한다.

한편, 시간상의 인과성(time sequence causality)을 의미하는 그랜저 인과관계 개념과는 다른 인과개념으로 Pearl(1995, 2000), Hoover(2005), Spirtes *et al.* (2000) 등은 시간의 흐름에 따르지 않는 동시적 인과관계(contemporaneous causality) 개념을 제시하고 구체적으로 DAG 방법론을 제시하고 있다.

본 연구에서는 분기별 데이터와 GDP, 자본, 노동, 에너지의 다변량모형을 이용하여 국민소득과 에너지 간의 인과관계를 살펴봄에 있어 오차수정모형과 DAG 기법을 접목하였다. 먼저 단위근검정 결과 모든 변수가 수준에서는 불안정적이지만 1차 차분에서는 안정적으로 나타났고, Johansen and Juselius(1990)의 공적분검정을 수행한 결과 이들 변수들 간에 두 개의 안정적인 선형관계가 존재함을 밝혔다.

그리고 GES 알고리즘을 적용한 DAG의 결과는 변수들 간에 연결성은 있었

지만 방향성은 제시하지 못하였다. 이에 따라 방향성을 찾기 위해 Wang and Bessler(2005)와 Park, Mjelde, and Bessler(2006)에서 사용된 Seemingly unrelated regression을 적용하여 방향성을 구하였다. 그 결과 동시적인 관점에서 GDP에서 에너지로의 인과관계는 있지만 그 역은 성립하지 않는 일방향 인과관계가 존재한다는 것을 발견하였다. 또한 DAG에서 구해진 인과관계를 사용하여 변수 간의 단기구조를 분석하기 위해 예측오차분산분해를 구하였다. 그 결과 에너지의 경우 GDP에 비해 훨씬 내생적 변수의 경향이 강한 것으로 나타났는데 이는 DAG에서 GDP가 에너지를 인과하는 결과와 일치하는 것이다.

이러한 일방향 실증분석 결과는 경제성장을 제약함이 없이도 어느 정도 에너지 소비를 줄일 수 있음을 시사해 주는 것이다. 국민소득에서 에너지로의 일방향 인과관계는 생산요소로서의 에너지가 전체 생산비에서 차지하는 비중이 매우 낮다는 점과 일맥상통하는 것으로 판단된다. 또한 이러한 결과는 시차적 인과관계를 의미하는 그랜저 인과관계 검정법을 이용하여 장기에 동일한 일방향 인과가 있음을 발견한 Oh and Lee(2004b)의 결과와 일치하는 것이기도 하다.

본 연구에서 사용한 DAG도 다른 시계열 분석법과 마찬가지로 1997년에 발생하였던 외환위기와 같은 구조변화에 영향을 받을 수 있다. 본 연구에서는 동 구조변화를 반영하기 위하여 외환위기 더미를 사용하기는 하였지만 보다 현실성 있는 과제는 과연 이들 변수들 간의 관계가 외환위기를 전후로 변화하였는지, 변화하였다면 어떻게 변화하였는지를 살펴보는 것일 수 있다. 따라서 향후 연구과제는 외환위기 전후로 구분하여 외환위기가 이들 변수 간의 관계에 미친 영향 분석하는 것이다. 이는 데이터가 충분히 축적된 후의 과제로 남겨 둔다.

참 고 문 헌

- 김영덕, 「유가변동의 산업별 파급효과」, 『자원·환경경제연구』 제14권 제2호, 2005. 6.
- 노동부 DB, <http://laborstat.molab.go.kr/>
- 박호정·윤원철, 「VAR와 그래프 이론을 이용한 시계열의 인과성 분석」, 『자원·환경경제연구』 제12권 제4호, 2003, 687~708.
- _____, 「오차수정모형과 그래프 이론을 이용한 국제유가의 동시 및 단기가격 발견과정에 관한 연구」, 『자원·환경경제연구』 제15권 제3호, 2006, 479~

504.

에너지경제연구원 DB, <http://www.keei.re.kr/index.html>

통계청, 통계정보시스템, <http://www.nso.go.kr>

한국은행, 경제통계시스템, <http://ecos.bok.or.kr>

Akarca, A. T. and T. V. Long, "On the relationship between energy and GNP: A reexamination," *Journal of Energy and Development*, 5, 1980, 326~331.

Asafu-Adjaye, J., "The relationship between energy consumption, energy prices and economic growth: time series evidence from Asian developing countries," *Energy Economics*, 22, 2000, 615~625.

Awokuse, T. O., "Export-led growth and the Japanese economy: evidence from VAR and directed acyclic graphs," *Applied Economics*, Vol. 38, 2006, 593~602.

Bernanke, B., "Alternative Explanations of the Money-Income Correlation," *Carnegie-Rochester Conference Series on Public Policy*, Vol 25, 1986, 49~99.

Bessler, D. A. and N. Loper, "Economic Development: Evidence From Directed Acyclic Graphs," *The Manchester School*, Vol. 69, 2001, 457~476.

Bessler, D. A. and Seongpyo Lee, "Money and Prices: U.S. Data 1869-1914 (A Study with Directed Graphs)," *Empirical Economics*, 27, 2002, 427~446.

Bessler, D. A. and J. Yang, "The Structure of Interdependence in International Stock Markets," *Journal of International Money and Finance*, 22, 2003, 261~287.

Bessler, D. A. and Z. Wang, "Price and Quantity Endogeneity in Demand Analysis: Evidence from Directed Acyclic Graphs," *Agricultural Economics*, 34, 2006, 87~95.

Cheng, Benjamin S. and T. W. Lai., "An investigation of co-integration and causality between energy consumption and economic activity in Taiwan," *Energy Economics*, 19, 1997, 435~444.

Chickering, D. M., "Optimal Structure Identification with Greedy Search," *Journal of Machine Learning Research*, 3, 2002, 507~554.

Geweke, J. and R. Meese, "Estimating Regression Models of Finite but Unknown Order," *International Economic Review*, 22, 1981, 55~69.

Glasure, Y. U. and A. Lee, "Cointegration, error-correction, and the relationship between GDP and energy: The case of South Korea and Singapore," *Resource and Energy Economics*, 20, 1998, 17~25.

- Gonzalo, J., "Five Alternative Methods of Estimating Long-Run Equilibrium Relationships," *Journal of Econometrics*, 60, 1994, 203~233.
- Granger, C. W. J., "Testing causality: a personal viewpoint," *Journal of Economic Dynamics and Control*, Vol. 23, 1980, 29~52.
- _____, "Some Recent Developments in a Concept of Causality," *Journal of Econometrics*, 39, 1988, 199~211.
- Granger, C. W. J. and Paul Newbold, "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2, 1974, 111~120.
- Haigh, M. and D. Bessler, "Causality and Price Discovery: An Application of Directed Acyclic Graphs," *Journal of Business*, 2003.
- Hamilton, J. D., *Time Series Analysis*, Princeton: Princeton University Press, 1994.
- Hausman, J. A., "Triangular structural model specification and estimation with application to causality," *Journal of Econometrics*, Vol. 112, 2003, 107~113.
- Hoover, R., "Automatic Inference of the Contemporaneous Causal Order of a System of Equations," *Econometric Theory*, 21, 2005, 69~77.
- Johansen, S., "Determination of Cointegration Rank in the Presence of a Linear Trend," *Oxford Bulletin of Economics and Statistics*, 54, 1992, 383~397.
- Johansen, S. and Katarina Juselius, "Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money," *Oxford Bulletin of Economics and Statistics*, 52, 1990, 169~210.
- Kraft, J. and A. Kraft, "On the relationship between energy and GNP," *Journal of Energy and Development*, 3, 1978, 401~403.
- Lükepohl, H., "Non-causality due to omitted variables," *Journal of Econometrics*, Vol. 19, 1982, 367~368.
- Lükepohl, H. and H. Breitung, "Impulse Response analysis of Vector Autoregressive Process," in C. Heij, J. Schumacher, B. Hanzon, and C. Praagman ed., *System Dynamics in Economic and Financial Models*, John Wiley & Sons, 1997.
- McKinnon, J., "Critical Values for Cointegration Tests," in R. F. Engle and C. W. J. Granger, eds., *Long-Run Economic Relationships: Readings in Cointegration*, New York: Oxford University Press, 1991, 267~276.
- Masih, A. M. M. and R. Masih, "Energy consumption, real income and temporal

- causality: results from a multi-country study based on cointegration and error-correction modelling techniques,” *Energy Economics*, 18, 1996, 165~183.
- _____, “On the Temporal Causal Relationship between Energy Consumption, Real Income, and Prices: Some New Evidence from Asian-Energy Dependent NICs based on a Multivariate Cointegration/Vector Error-Correction Approach,” *Journal of Policy Modeling*, 19(4), 1997, 417~440.
- _____, “A Multivariate Cointegrated Modelling Approach in Testing Temporal Causality between Energy Consumption, Real Income and Prices with an Application to Two Asian LDCs,” *Applied Economics*, 30(10), 1998, 1287~1298.
- Meek, C., “Graphical Models: Selecting Causal and Statistical Models,” Unpublished Ph. D. Dissertation, Carnegie Mellon University, 1997.
- Oh, Wankeun and Kihoon Lee, “Causal Relationship between Energy Consumption and GDP revisited: The case of Korea 1970-1999,” *Energy Economics*, Vol. 26, No. 1, 2004a, 51~59.
- _____, “Energy Consumption and Economic Growth in Korea: testing the causality relation,” *Journal of Policy Modeling*, Vol. 26, No. 9, 2004b, 973~981.
- Orcutt, G., “Toward a Partial Redirection of Econometrics,” *Review of Economics and Statistics*, 34, 1952, 195~213.
- Osterwald-Lenum, M., “Practioners Corner A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics,” *Oxford Bulletin of Economics and Statistics*, 54, 1992, 461~472.
- Park, Haesun, J. W. Mjelde, and D. A. Bessler, “Price dynamics among U.S. electricity spot markets,” *Energy Economics*, Vol. 28, 2006, 81~101.
- Pearl, J., *Causality: Models, Reasoning, and Inference*, Cambridge Press, 2000.
- Phillips, P. C. B., “Understanding Spurious Regressions in Econometrics,” *Journal of Econometrics*, 33, 1986, 311~340.
- Pyo, Hak Kil and Bongchan Ha, “The Estimation of Potential GDP in Korea: A Review,” presented on Feb. 14 at a seminar organized by Asia Pacific Division of International Monetary Fund, Washington D.C., 2006.
- Schwarz, G., “Estimating the Dimension of a Model,” *Annals of Statistics*, 6, 1978,

461~464.

- Schwert, G. W., "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data," *Journal of Monetary Economics*, 20, 1987, 73~103.
- Sims, C., "Are Forecasting Models Usable for Policy Analysis?," *Federal Reserve Bank of Minneapolis Quarterly review*, 10, 1986, 2~15.
- Spirtes, P., C. Glymour, and R. Schines, *Causation, Prediction and Search*, Cambridge, MA: MIT Press, 2000.
- Stern, D. I., "Energy and Economic Growth in the USA," *Energy Economics*, 15, 1993, 137~150.
- _____, "A Multivariate Cointegration Analysis of the Role of Energy in the US Macroeconomy," *Energy Economics*, 22(2), April 2000, 267~283.
- Swanson, N. R. and C. W. J. Granger, "Impulse Response Functions Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions," *Journal of the American Statistical Associations*, 92, 1997, 357~367.
- Wang, Z. and D. A. Bessler, "Forecast Evaluations in Meat Demand Analysis," *Agribusiness*, Vol. 19, 2003, 505~524.
- Yang, H. Y., "A note on the causal relationship between energy and GDP in Taiwan," *Energy Economics*, 22, 2000, 309~317.
- Yu, E. S. H. and B. K. Hwang, "The Relationship Between Energy and GNP: Further Results," *Energy Economics*, 6, 1984, 186~190.
- Yu, E. S. H. and J. C. Jin, "Cointegration Tests of Energy Consumption, Income, and Employment," *Resources and Energy*, 14(3), 1992, 259~266.

[Abstract]

A Study on Contemporaneous Causal Relationship between GDP and Energy Using Graph Theory

Haesun Park · Wankeun Oh

Applying a multivariate error correction model of GDP, energy, capital, and labor, we have studied contemporaneous causality between income and energy using directed acyclic graph. Empirical results from DAG and forecast error variance decomposition for Korea over the period 1981:1–2004:4 suggest a unidirectional causal relationship running from GDP to energy in contemporaneous time. It implies an energy conservation policy may be feasible without compromising economic growth in the long run. The result is also consistent with Oh and Lee (2004b) which adopts Granger time sequence causality.

Keywords: GDP, energy, DAG, VECM, GES

JEL Classification: Q43