

한국의 미세먼지 발생요인 분석: 공간계량모형의 적용

강희찬*

요약 : 본 논문에서는 한국의 미세먼지 발생원인을 분석하는 과정에서 기존 논문에서는 고려하지 않았던 지역 간 공간상관성(Spatial correlation)을 고려한 패널계량분석을 진행하였다. 기존 환경쿠즈네츠곡선(EKC, Environmental Kuznets Curve)에 대한 연구들에서, 인접한 국가 및 지역 간에 오염물질의 상호영향이 존재할 가능성이 있음에도 각 유닛이 독립이라고 가정한다. 본 논문에서는 한국의 미세먼지농도에 대한 지역 패널데이터를 이용하여 기존 EKC가 지역의 상호상관성을 고려하는 때도 성립할 수 있으며, 이러한 영향을 고려하지 않았을 때 미세먼지농도의 원인에 대해 과소 혹은 과대 추정될 수 있음을 규명하였다.

주제어 : 미세먼지, 공간계량모형, 공간상관성, 환경쿠즈네츠곡선가설

JEL 분류 : Q1, R1

접수일(2019년 5월 10일), 수정일(2019년 8월 6일), 게재확정일(2019년 8월 21일)

* 인천대학교 경제학과, 부교수, 교신저자(e-mail: henrykang@inu.ac.kr)

An Analysis of the Causes of Fine Dust in Korea Considering Spatial Correlation

Heechan Kang*

ABSTRACT : In this paper, we conducted panel data analysis considering spatial correlation between regions, which were not considered in previous papers in analyzing the causes of fine dust concentration in Korea. Many existing researchers implicitly assume the independence of the effects of incomes and other explanatory variables of adjoining countries(or regions). Using panel data on fine dust concentration, this paper has established that existing EKC can be established even when considering the spatial correlation of the region, and when these effects are not taken into account, it can be underestimated or overestimated on the effects and causes of fine dust concentration.

Keywords : Spatial correlation, PM10 concentration, Environmental Kuznets Curve, Spatial Durbin Model

Received: May 10, 2019. Revised: August 6, 2019. Accepted: August 21, 2019.

* Associate professor, Department of Economics of Incheon National University, Corresponding author (e-mail: henrykang@inu.ac.kr)

I. 서론

최근 한국에 미세먼지는 가장 골치 아픈 문제가 되고 있다. 특히 미세먼지 문제가 대부분 도시지역을 중심으로 나타나는 것으로 여겨졌지만, 최근에는 농촌지역, 산악지역, 해안지역을 불문하고 전국적인 심각한 문제로 나타나고 있다. 발생원인에 대한 다양한 논쟁이 있지만 비단 중국에서 넘어온 미세먼지뿐만 아니라 국내 발생 미세먼지에 대한 발생원과 확산 경로에 대해서도 명쾌한 결론이 없다. 중국발 미세먼지가 주요 발생원으로 주목되고 있지만 모든 귀책사유를 중국에만 전가하는 것도 문제가 된다. 만일 한국에서 발생하는 미세먼지의 발생원인과 경로를 제대로 파악하여 철저히 대비하지 못하면서 중국에 감축을 제안할 수 없을 뿐만 아니라 국내적 대처도 소홀히 하는 도덕적 해이가 발생할 수 있다. 이에 따라 본 논문에서는 한국의 미세먼지에 한정하여 분석하고자 한다.

현재 높은 미세먼지농도로 인한 국민의 건강을 보호하기 위해 다양한 대책이 나올 필요가 있지만 동시에 미세먼지의 국내 발생원인을 조기에 발견하고, 이를 줄이기 위한 엄격한 대책이 상호 보완적으로 진행되어야 할 것이다. 미세먼지 발생원인과 파급경로는 국내·외 경제활동과 기상변화 등 매우 복잡한 양상을 보인다. 국내적으로는 국내 화력발전소, 사업장, 자동차, 선박, 석탄 야적장, 가정 등 다양하며, 중국에서 발생하는 미세먼지도 주요 국외 발생원이 될 수 있다. 또한, 파급경로는 계절, 바람의 방향, 습도, 날씨, 강수량 등 다양한 요인에 영향을 받는다. 발생한 미세먼지는 축적되어 이동과 정체를 거듭하며 농도를 높이고, 국민의 건강을 위협하고 있다.

이처럼 다양한 요인에 의해 변하는 미세먼지농도는 국가 전체적 요인들(예를 들어, 자연 환경적 요인들)에 의해 영향을 받을 뿐만 아니라 미시적으로는 한 지역의 미세먼지농도는 해당 지역 내의 미세먼지 발생원에 의해 영향을 받을 수 있고, 간접적으로 인접 지역의 미세먼지 발생원으로부터 영향을 받을 수도 있다. 그러나 기존의 논문들에서는 지역 내 미세먼지농도에 미치는 요인들을 분석하면서 지역 간 발생하는 영향을 배제한 채 지역 내 미세먼지 발생원의 영향에만 집중하였다. (윤인주·한상연, 2010; 이광훈, 2010; 배정환·김유선, 2016) 그러나 경계를 맞대고 있는 지역 간 미세먼지농도 영향을 무시할 경우 발생원인에 대한 과소 혹은 과대 추정의 가능성을 담고 있었다. 이는 결과적으로는 미세먼지 저감을 위한 정부 차원의 정책결정도 왜곡시킬 수 있다.

이러한 ‘지역 간 상호영향의 패턴’은 크게 두 가지로 구분할 수 있다. 첫 번째는 미세먼지 농도 자체에 관한 것으로, 한 지역의 미세먼지농도 증가는 주변 지역의 미세먼지농도에 영향을 미칠 수 있다. 이러한 미세먼지농도에 미치는 영향은 계절, 날씨, 기상, 바람, 기류 등 다양한 요인에 의해 영향을 받을 수 있기에 하나의 방향성을 두지는 않는다. 일반적으로는 한 지역의 미세먼지농도 증가는, 누출 효과(spillover effect)로 인해 주변 지역의 미세먼지농도를 높일 가능성이 크다. 따라서 근접한 지역 간에는 강한 공간적 상관관계(spatial correlation)가 존재한다.¹⁾ 두 번째는 한 지역의 미세먼지 발생원의 영향은 해당 지역의 미세먼지 농도를 높일 뿐만 아니라 인접 지역의 미세먼지 농도에도 영향을 미칠 수 있다. 예를 들어 한 지역의 생산 활동 증가는 해당지역의 미세먼지 농도를 높일 수 있고, 인접한 다른 지역의 미세먼지 농도도 높일 가능성이 있다. 물론 인접도가 높다고 해서 미세먼지 농도에 미치는 영향이 항상 크다고 할 수는 없다. 그러나 다른 조건이 동일할 때 인접도가 높은 지역 간에 한 지역의 미세먼지 발생원이 인접 지역의 미세먼지 농도에 미치는 영향은 그렇지 않은 지역에 비해서 클 수밖에 없다. 한 국가 내 지역들(혹은 국경을 맞댄 국가들)에서 미세먼지 농도변화를 유발하는 주요 요인을 분석할 때는 특별히 두 번째 영향 패턴에 관심을 두게 된다. 즉 특정 지역에서 미세먼지농도가 변하는 것은 해당 지역의 미세먼지 발생량의 증가로 인한 것일 수도 있지만, 동시에 인접 지역의 소득이나 에너지 사용과 같은 독립변수에 의한 영향도 고려해야 한다.

따라서 본 논문에서는 이처럼 한 국가 내 지역의 미세먼지 농도변화 유발요인을 분석하면서, 기존의 연구와 같이 지역을 독립된 단위로 접근하지 않고 지역 간 변수들의 상호영향을 고려한 공간계량방법론을 채택하여 분석하기로 한다. 그리고 이러한 분석방법을 채용했을 경우, 기존 연구들에서 분석한 미세먼지 농도변화 유발요인의 추정량이 혹시 과대 혹은 과소 추정되었는지를 비교함으로써, 더 정밀한 분석이 가능한지 검증해보고자 한다. 만일 지역 간에 강한 공간적 상관관계(spatial correlation)가 존재할 경우, 기존의 계량분석 방법들을 적용할 때는 추정의 편의(bias)가 발생한다고 알려져 있다 (Anselin, 1988).

1) 그러나 이러한 변화를 분석하고자 할 때는 실시간 혹은 매일 단위일 경우에만 포착할 수 있고, 이를 연간 데이터로 분석하고자 할 때는 한 지역에 다른 지역에서 유입된 미세먼지는 그 지역에서 다른 지역으로 유출되는 미세먼지 먼지의 양에 의해 상쇄됨으로써 미세먼지의 순농도에 미치는 영향은 크지 않을 가능성도 있다.

본 연구에서는 국내 16개 시·도²⁾의 2003~2014년 PM₁₀ 데이터를 이용하여, 미세먼지의 농도와 미세먼지 발생원(특히 사회경제적 변수) 간의 관계를 분석하였다. 이를 위해 공간계량모형을 통해 한 지역 그리고 인접 지역의 경제 수준, 산업 활동, 교통의 변화가 상호간의 미세먼지농도에 어떤 영향을 미치는지 분석한 것이다. 본 논문에서는 적절한 공간계량모형을 적용하여, 잠재적인 공간적 상관관계(spatial correlation)를 통제하도록 한다.

II. 선행연구 검토

오염물질과 경제성장 간의 논문들에서 흔히 활용하는 것이 환경쿠즈네츠 곡선(EKC, Environmental Kuznets Curve) 가설이다. 즉 경제발전 초기에는 오염물질 배출량이 증가하지만 일정 소득수준을 넘으면 오염이 다시 감소한다는 가설에 관한 다양한 실증연구가 진행되었다. Grossman and Kruger(1995)가 다양한 오염물질에 대해 역 U자형 영향을 실증적으로 규명한 이래로 수많은 논문에서 이러한 관계를 유발하는 다양한 변수들을 도입하여 분석하였다(Seldon and Song, 1994; Cole et al., 1997; Suri and Chapman, 1998; Stern, 2004)

최근에는 오염물질과 경제성장에 관한 연구에서 인접 국가 혹은 국가 내 지역 간 공간 상호영향을 고려한 논문들이 발표되고 있다. (Anselin, 2001; Maddison, 2006). 이들 논문에서는 지리적으로 인접한 지역에서 얻은 데이터는 더 먼 지역에 비해 서로 관련되어 있을 수 있다는 점에 주목하고, 이러한 공간상관성을 EKC 실증분석에 고려해야 한다고 주장한다. 이러한 공간상관성은 인접 국가 간 연구에 있어서 월경성(Transboundary) 대기오염물질에 관한 연구(Stern, 2000)와 오염피난처가설(pollution haven hypothesis)에 따른 먼 국가보다는 인접 국가로 오염배출사업장 이동 연구(Rothman, 1988), 저감기술의 인접 국가 간 파급에 관한 연구(Keller, 2004) 등에서 간략하게 연구되었다. 한편 한 국가 내 지역의 공간상관성에 관한 연구도 다양하게 진행되고 있다. Rupasingha et al. (2004)은 미국 내 지역의 소득수준과 독성오염물질에 관한 EKC분석에서 횡단면 데이터를 통해 공간영향, 인종다변도 등을 고려하였고, Poon et al.(2006)은 중국의 이산화황

2) 세종시는 데이터 일관성의 이유로 제외함.

과 그을음(soot)에 관한 EKC 분석에서 공간상관성을 고려한 횡단면 데이터를 이용해 소득수준에 대한 이산화황의 역 U자 관계를 발견하였다. Hao and Liu (2016)는 중국 내 PM_{2.5} 농도에 영향을 미치는 사회경제학적 요인들을 분석하면서 공간상관성을 고려하였다. 최근에는 중국 내 오염물질과 소득수준 간에 EKC 실증연구에서 공간 패널데이터를 활용한 연구가 진행되고 있다(Kang et al., 2016; Xiangyu et al., 2018).

이러한 공간상관성을 고려한 공간계량방법론은 Anselin(1988)이 학문체계를 제시한 이후 많은 분석에서 응용되고 있다. Anselin(1988)은 공간상관관계를 고려하지 않는 전통적인 회귀분석을 적용할 경우 편의(bias)와 불일치(inconsistent) 추정량을 갖게 된다는 사실을 증명하고 공간시차모형(Spatial lag model, SAM 혹은 SAR)과 공간오차모형(Spatial error model, SEM)을 제시하였다. SAM(SAR) 모형에 따르면, 종속변수의 공간시차 자기상관을 고려할 때 한 국가(지역)의 경제성장이 주변 국가(지역)의 성장에 미치는 영향을 분석한다. SEM은 오차의 공간 자기상관을 고려하여 공간가중치 행렬을 통해 오차항이 모형에 반영되어 지역 경제성장에 영향을 준다는 점을 고려한다. 즉, 오차항의 공간 자기상관을 통해 외부충격이 전 지역 경제성장에 영향을 미친다는 것이다. 공간간의 상호작용은 지역 간 거리가 가까울수록 효과가 크고, 멀어질수록 감소하는 효과를 반영하고 있다. 이뿐 아니라 Kelejian and Prucha(1988)는 공간시차 종속변수와 공간 자기상관 오차를 모두 포함하는 SAC모형을 제안하였다. LeSage and Pace(2009)는 한 국가(지역) 소득이 이웃한 지역의 소득수준뿐만 아니라 기타 독립변수들에 의해서도 영향을 받을 수 있다는 가정에 따라 공간시차 종속변수에 공간시차 독립변수들을 포함하는 공간터빈모형(Spatial Durbin model, SDM)을 활용하였다. 이 연구에 따르면, 한 국가(지역)의 독립변수의 변화에 따른 종속변수의 영향을 직접효과(Direct effect)라고 하였고 한 국가(지역)의 독립변수의 변화에 따른 다른 지역의 종속변수의 영향을 간접효과(Indirect effect, Spillover effect)라고 하였다.

기존 공간상관성 연구가 횡단면이나 시계열 자료 분석에 따라 대부분 이뤄졌으나 최근에는 공간 패널(spatial panel)에 기반을 둔 모형 설정과 추정이 증가하고 있다. Elhorst (2014)는 다양한 패널 공간계량 경제모형을 소개하였다.

공간상관성을 고려한 EKC 실증연구가 주로 국가 간 연구에서 이뤄졌고, 국가 내 지역 연구에서도 중국, 미국 등 영토가 넓은 국가에 관한 연구였던 반면, 본 연구에서는 미세

먼지농도에 관한 대한민국 내 지역 간 공간상관성을 고려한 패널데이터를 활용한 연구라는 측면에서 본 연구의 의의가 있다. 따라서 본 연구에서는 한 지역의 미세먼지농도 변화가 인접한 지역의 미세먼지농도 변화에 영향을 미칠 수 있고, 지역이 인접해 있다면 미세먼지농도를 변화시킬 수 있는 한 지역의 설명변수 변화가 인접 지역의 미세먼지농도에도 영향을 미칠 가능성이 있다고 보고, 국내 16개 시·도에 대한 공간패널을 구축하고, 다양한 패널 공간계량 모형을 통해 분석하고자 한다.

III. 데이터와 추정모형

1. 데이터

미세먼지농도는 에어코리아의 대기환경 연보에서 추출하였다.³⁾ 기간은 2003~2014년까지 16개 시도에 대한 패널데이터를 구성하였다. 선행연구⁴⁾를 참조하여 미세먼지농도에 영향을 미칠 수 있는 설명변수로는 1인당 GRDP, 석탄화력 발전량, 경유자동차 등록대수, 황사일수, 비발전용 유연탄 소비량을 선택하였다. 다음 <표 1>에서는 변수를 설명하고 있다. 이외에도 다양한 변수들이 미세먼지 농도변화에 영향을 미칠 수 있으나 다중공선성 문제를 줄이기 위해 제외하였다.⁵⁾ 석탄화력발전량은 지역의 에너지산업 연소의 영향을 파악하기 위한 변수⁶⁾이며, 경유차 등록대수는 도로이동오염원이나 비산먼지의 영향을 파악하기 위한 변수⁷⁾이다. 황사일수는 기상 등 물리적 영향으로 미세먼지농

3) 미세먼지배출량 대신 농도데이터를 사용한 것은 본 논문이 분석하고자 하는 것이 지역 간 미세먼지 농도에 미치는 영향을 파악하고자 하는 것이기 때문이다. 논리적으로 미세먼지 배출량은 인접한 지역 간에도 영향을 받을 수 있지만 그 영향은 미미하고, 이에 반해 미세먼지 농도는 미세먼지가 월경가능성이 있으므로 인접 지역 간에 영향을 받을 수 있기 때문이다.

4) Grossman and Krueger(1995); Wang et al.(2019); 배정환·김유선(2016)

5) 미세먼지농도에 영향을 미칠 수 있는 변수로 온도, 강수량, 제조업연소 배출량, 건설기계 및 선박 등 비도로이동 오염 배출량 등이 있으나, 이들 변수와 본 논문에서 선택한 설명변수 간에 높은 상관관계로 인해 제외하였음. 특히 온도나 강수량은 일 혹은 월별 데이터에서는 의미 있는 변수일 수 있으나, 연간 데이터에서는 큰 차이를 보이지 않음.

6) 본 논문에서 사용한 석탄화력발전량(GWh) 변수는 석탄발전분야의 미세먼지농도 기여를 파악하기 위한 것이며, 이외에 고려할 수 있는 석탄화력발전 설비 용량(GW)은 변수의 설명력이 낮고, 그 외 전체 산업의 석탄사용량은 소득변수와 높은 상관관계로 인해 다중공선성 문제를 제거하기 위해 포함시키지 않음.

7) 본 논문에서 차량용 경유 사용량, 운행거리 대신 경유차 등록대수를 설명변수로 사용한 것은 다른 변수(경유 사용량, 운행거리)는 소득변수와 높은 상관관계로 다중공선성의 문제가 있기 때문이다.

〈표 1〉 변수 설명

변수명	변수설명	단위	출처
$\ln PM$	미세먼지 농도 로그값	$\mu\text{g}/\text{m}^3$	에어코리아 대기환경연보
$\ln Y$	일인당 GRDP 로그값	10억 원/명	한국통계연감 지역계정
$\ln PS$	석탄화력발전량 로그값	GWh	한국전력통계
$\ln RD$	경유차 등록대수 로그값	천 대	국토교통부 월간자동차등록현황
$\ln Yellow$	황사일수 로그값	일	기상청
$\ln NGC$	非발전용 유연탄 소비량 로그값	1,000톤	대한석탄협회

도에 미치는 영향을 제어하기 위한 변수이며, 非발전용 유연탄 소비량은 제조업 연소나 제조업공정에서 배출되는 오염물질의 영향을 파악하기 위한 변수이다. 본 논문에서는 실시간 데이터인 미세먼지에 비해 다른 설명변수들은 연간 데이터로 구성되어 시간 통일을 위해 월별 미세먼지 농도 데이터를 12개월 평균값으로 사용하였다. 이로 인해 계절적 영향은 파악할 수 없다는 한계가 있다. <표 2>는 설명변수들 상관계수를 보여주고 있다. 모든 상관계수가 낮거나 크게 높지 않게 나타난다. Myers(1990)가 제안한 VIF(variance inflation factor) 검정치도 1.02~1.11 범위로 평균은 1.07로 나타나 10보다 작게 나타났다. 모든 자료는 변동성과 분산의 영향을 줄이기 위해 자연대수를 취한 값을 사용한다.

각 변수의 기술통계량은 <표 3>과 같다.

미세먼지농도의 공간상관성을 살펴보기 위해 2003~2014년간 국내 16개 시도의 미세먼지농도의 상관계수를 인접 지역과 비인접 지역으로 구분하여 측정하고 그 결과를 <표 3>에 정리하였다. 여기서 인접 지역이란 시·도의 경계를 맞대고 있는 지역을 의미하고 비인접 지역은 그렇지 않은 지역이다. 예를 들어 인천의 경우 서울시와 경기도가 인접 지

〈표 2〉 상관계수 행렬과 VIF 검정

	VIF	$\ln Y$	$\ln PS$	$\ln RD$	$\ln Yellow$	$\ln NGC$
$\ln Y$	1.11	1.00				
$\ln PS$	1.09	0.21	1.00			
$\ln RD$	1.05	-0.12	-0.05	1.00		
$\ln Yellow$	1.05	0.15	0.08	-0.05	1.00	
$\ln NGC$	1.02	0.22	0.12	-0.11	0.18	1.00

〈표 3〉 기술통계량

변수명	관측수	평균	표준편차	최솟값	최댓값
<i>PM</i>	192	50	7	33	69
<i>Y</i>	192	23,689	9,879	10,898	62,938
<i>PS</i>	192	10,707	24,801	0	103,843
<i>RD</i>	192	395	333	83	1,868
<i>Yellow</i>	192	7	4	1	27
<i>NGC</i>	192	2,058	4,021	0	16,227

역이고, 그 외 13개 시도는 비인접지가 된다. 이때 2003~14년 동안 인천과 인접 지역의 미세먼지농도의 상관관계(평균값)가 인천과 비인접 지역의 미세먼지농도의 상관관계(평균값)보다 통계적으로 크다면, 인천과 인접 지역의 미세먼지에 관한 공간상관성이 존재할 수도 있다고 조심스럽게 판단할 수 있다.

<표 3>에서 보는 바와 같이, 서울과 인천에 대해서는 인접 지역이 비인접지에 비해 미세먼지의 상관관계가 통계적으로 더 높게 나타나고 있다. 그 외 지역은 통계적으로는 인접 지역과 비인접지 간에 차이가 나지 않는다.⁸⁾ 결과적으로 국내 16개 시·도에서 평균적으로 미세먼지농도 간에 공간상관성이 존재한다는 가설은 단순 상관관계 분석을 통해

〈표 4〉 미세먼지 농도의 상관관계

지역	인접 지역	비인접 지역	t-통계량	지역	인접 지역	비인접 지역	t-통계량
서울*	0.94	0.53	4.05	강원	0.64	0.66	-0.04
부산	0.71	0.69	0.10	충북	0.21	0.08	0.83
대구	0.80	0.57	1.48	충남	0.77	0.66	1.01
인천*	0.94	0.65	3.88	전북	0.36	0.35	0.05
광주	0.75	0.58	0.68	전남	0.57	0.67	-0.51
대전	0.64	0.74	-0.33	경북	0.55	0.71	-1.03
울산	0.51	0.41	0.51	경남	0.61	0.55	0.43
경기	0.78	0.64	0.68	제주	1.00	0.37	14.22

주: *는 5% 유의수준에서 두 통계치가 같다는 귀무가설이 기각된 것을 나타냄.

자료: AirKorea

8) 대전, 강원, 전남, 경북 지역에서는 통계적으로는 차이가 없더라도 각 지역의 인접 지역에 비해 비인접 지역에 대해 미세먼지 농도의 상관관계가 더 높은 것으로 나타난다.

서는 명쾌하게 밝혀질 수 없다. 하지만, 서울과 인천에 대해서는 확실히 미세먼지농도에 있어서 통계적으로 공간상관성이 나타나고 있다고 예상할 수 있다.

이제 지역 간에 미세먼지농도에 있어 공간상관성을 보다 체계적으로 분석하기 위해, 다음 분석을 살펴보자. 추가적인 연구를 진행하기 위해서는 공간계량경제모형(Spatial econometrics)을 이용해야 한다. 본 논문에서는 16개 시·도의 2003~2016년 자료이므로 패널데이터를 구축하되 공간상관성을 고려한 공간패널모형을 구축한다.

〈그림 1〉 공간가중치 행렬

<1> 인접 여부를 이용한 공간가중치 행렬

[서울 부산 대구 인천 광주 대전 울산 경기 강원 충북 충남 전북 전남 경북 경남 제주]

W=	서울															
	부산															
	대구															
	인천															
	광주															
	대전															
	울산															
	경기	.000	.000	.000	.500	.000	.000	.000	.500	.000	.000	.000	.000	.000	.000	.000
	강원	.000	.000	.000	.000	.000	.000	.500	.000	.000	.000	.000	.000	.000	.500	.000
	충북	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.500	.500
	충남	.500	.000	.000	.000	.000	.000	.000	.500	.000	.000	.000	.000	.000	.000	.000
	전북															
	전남															
	경북															
	경남															
	제주															

<2> 거리(좌표)를 이용한 공간가중치 행렬

[서울 부산 대구 인천 광주 대전 울산 경기 강원 충북 충남 전북 전남 경북 경남 제주]

W=	서울															
	부산															
	대구															
	인천															
	광주															
	대전															
	울산	.000	.001	.002	.173	.002	.006	.001	.009	.781	.007	.008	.003	.001	.003	.001
	경기	.008	.000	.118	.007	.020	.021	.461	.009	.008	.022	.011	.020	.014	.044	.231
	강원	.014	.123	.000	.013	.028	.058	.181	.016	.015	.071	.022	.038	.017	.212	.184
	충북	.586	.004	.006	.000	.006	.020	.004	.020	.271	.019	.034	.010	.004		
	충남															
	전북															
	전남															
	경북															
	경남															
	제주															

본 연구의 분석단위인 행정구역 시·도는 미세먼지농도 혹은 실제 경제 지역과 정확하게 일치하지 않는다. 따라서 관측되지 않는 요소가 인접 지역 간 공간적으로 상호 연관성 (Spatial dependence)을 가지고 있을 개연성을 배제할 수 없다.

공간계량경제모형(Spatial econometrics)은 공간자료의 문제를 모형에서 명시적으로 처리할 수 있다는 장점이 있다. 이는 계량모형의 시간 대신 공간을 주 대상으로 분석하는 것이다. 다시 말해 시계열 계량모형에서는 시간의 자기상관관계를 분석하는 반면, 공간 계량모형에서는 공간적 자기상관관계를 분석하는 것이다. 공간계량모형에서 우선해야 하는 것은 공간가중치행렬(spatial weighted matrix)을 구성하는 것이다. 이 공간가중치 행렬에서는 공간적으로 더 인접한 지역에 높은 가중치를 부여함으로써 지역 간 상호작용 혹은 공간상호연관성(spatial dependence)을 검정하거나 추정할 수 있다. 공간가중치 행렬을 구성하는 방식은 다양한데, 그중에서 이용빈도가 가장 높은 것은 인접 (Contiguity)에 가중치를 부여하는 방식⁹⁾과 공간좌표(Coordinates)를 이용하는 방법¹⁰⁾이 있다. <그림 1>에서는 본 논문에서 사용한 두 가지 형태의 공간가중치 행렬들을 나타 내고 있다. W 는 16×16 행렬이며, W 의 대각행렬의 모든 원소는 모두 0이다.

2. 모형 설명

본 연구에서 사용하는 계량모형은 기본적으로 환경쿠즈네츠 가설을 기반으로 한다. 이 가설에 따르면, 종속변수인 오염물질이 설명변수, 특히 소득수준에 따라 비선형의 특징을 보이는데, 특히 소득이 낮은 구간에서는 소득과 함께 오염물질이 증가하고, 소득이 높은 구간에서는 소득과 함께 감소하는 특징을 보인다고 가정한다. 이러한 환경쿠즈네츠 가설검정을 위한 일반적인 패널계량모형은 다음의 식 (1)과 같이 나타낼 수 있다.

$$\ln PM_{i,t} = \alpha_i + \beta_1 \ln Y_{i,t} + \beta_2 (\ln Y_{i,t})^2 + \beta_3 (\ln Y_{i,t})^3 + x_{i,t} \eta + \mu_{i,t} \quad (1)^{11)}$$

-
- 9) 인접에 가중치를 부여하는 방식인 *Threshold distance* 방법은 지역이 서로 인접(contiguity)하는 경우에는 1, 그렇지 않으면 0으로 가중치를 부여한다. 예를 들어, 서울과 서로 인접한 인천과 경기도에 대해서는 1을 부여하고, 그렇지 않으면 0을 가중치로 부여한다. 일반적으로 가중치는 행(row)을 기준으로 합이 1이 되도록 다시 표준화한다.
- 10) 공간좌표를 이용하는 방법은 지역의 중심을 기준으로 다른 지역의 중심까지 좌표거리를 계산하고, 이를 횡단 표준화함으로써 행의 합이 "1"이 되도록 하는 방법이다.

여기서 $PM_{i,t}$ 는 t 시점에서 각 i 지역별 미세먼지농도, Y 는 일인당 GRDP, x 는 통제 변수 벡터, α_i 는 알려지지 않은 i 지역의 시간과 무관한 특수성¹²⁾, μ_{it} 는 *iid*로 표현된다. 이 패널데이터는 전통적인 분석방법인 고정효과모형(FE, Fixed effect model)이나 확률 효과모형(RE, Random effect model)의 분석기법을 사용한다. 그러나 식 (1)과 같은 일반적 패널데이터의 모형에서 지역 간 공간상호연관성이 존재함에도 이를 고려하지 않으면 추정량에 편의(bias)나 불일치(consistency)가 발생한다고 한다. (Anselin, 1988) 따라서 식 (1)에 있는 패널 일반 계량모형에 공간상관성을 고려하기 위해서는 앞서 설명한 몇 가지 방식을 고려한다. 첫째는 패널 공간시차모형(SLM 혹은 SAR)으로 다음의 식 (2)과 같이 나타낼 수 있다.

$$\ln PM_{i,t} = \lambda \sum_{j=1}^N W_{ij} \ln PM_{jt} + \alpha_i + \beta_1 \ln Y_{i,t} + \beta_2 (\ln Y_{i,t})^2 + \beta_3 (\ln Y_{i,t})^3 + x_{i,t} \quad (2)$$

식 (1)과 차이는 $\lambda \sum_{j=1}^N W_{ij} \ln PM_{jt}$ 가 식 (1)에 포함되는 것으로, 여기서 λ 은 “공간자기 회귀계수(Spatial autoregressive coefficient)”라고 한다. 식 (2)에서 보는 것처럼, 종속변수($\ln PM_{i,t}$)가 내생적으로 공간상관이 존재한다고 가정한다. W_{ij} 는 앞서 설정한 공간가중치행렬이다.

둘째, 공간오차모형(SEM, Spatial error model)은 다음의 식 (3)과 같이 쓸 수 있다.

$$\ln PM_{i,t} = \alpha_i + \beta_1 \ln Y_{i,t} + \beta_2 (\ln Y_{i,t})^2 + \beta_3 (\ln Y_{i,t})^3 + x_{i,t} \eta + \epsilon_{i,t} \quad (3)$$

$$\epsilon_{i,t} = \rho \sum_{j=1}^N W_{ij} \epsilon_{j,t} + \mu_{i,t}$$

11) 본 논문에서는 N자형 환경쿠즈네츠 곡선을 가정한다. 즉, 소득이 증가함에 따라 오염물질이 증가하지만, 일정 수준 이후에는 오염물질이 감소하고, 이후 다시 오염물질이 증가할 수 있음을 가정한다.
 12) 본 논문에서는 시간 특수성을 고려하지 않는다. 따라서 시간의 변화에도 불구하고 변하지 않는 시·도 변수의 특성만 고려한다.

식 (3)이 식 (1)과 다른 점은 오차항($\epsilon_{i,t}$)이 *iid*가 아닌, 다른 지역의 오차항과 공간상 관되어 있다는 점이다. 이는 종속변수의 결정요인 중에 모형에서 누락된 부분이 공간적으로 자기상관되어 있는 경우이다. 여기서 ρ 는 “공간자기상관계수(Spatial autocorrelation coefficient)”라고 한다.

마지막으로 미세먼지농도(종속변수)가 공간적 자기상관이 존재하는 동시에 인접 지역의 독립변수(들)에도 영향을 받는다고 가정할 때는 식 (4)와 같이 SDM(Spatial Durbin Model)¹³⁾을 사용한다.

$$\ln PM_{i,t} = \lambda \sum_{j=1}^N W_{ij} \ln PM_{jt} + \alpha_i + X_{i,t} \beta + \theta \sum_{j=1}^N W_{ij} X_{j,t} + \mu_{ti} \quad (4)$$

여기서 $X_{i,t}$ 는 모든 설명변수를 포함하는 벡터이다.

물론 이외에도 다양한 공간계량모형들이 존재한다. 첫째, 설명변수에만 외생적 공간상관성이 존재하는 SLX모형이 있다. 그러나 모형의 단순함에도 불구하고 종속변수(미세먼지 농도)의 공간상관성을 고려하지 못해 본 논문에서는 제외한다. 둘째, 종속변수와 오차항 모두에 내생적인 공간상관성이 존재하는 것을 전제로 하는 SAC모형이 있다. 그러나 본 논문에서는 미세먼지농도에 영향을 미치는 요인 분석을 목적으로 하므로, 이 모형도 제외한다. 셋째, 오차항의 자기상관되어 있고, 종속변수가 인접 지역의 독립변수에 영향을 받는 것을 가정한 SDEM(Spatial Durbin Error model)이나, 이를 더욱 일반화시켜서 종속변수의 내생적 자기상관, 오차항의 내생적 자기상관, 독립변수의 외생적 자기상관이 있다는 가정에 의한 가장 일반적 형태인 GNS(General nesting spatial model)가 있으나, 본 연구에서는 이들 모형은 제외한다.¹⁴⁾ 본 논문에서는 특별히 SDM모형에 관심을 둔다. 앞서 설명한 바와 같이, 미세먼지농도는 인근 지역의 미세먼지농도에 영향을 받는 동시에, 인근 지역의 다른 독립변수(예를 들어, 소득, 황사, 석탄화력발전량, 에너지 소비 등)에 영향을 받을 수 있기 때문이다.

13) LeSage and Pace(2009)과 Anselin (1988)에 따르면, SDM은 시계열 계량모형의 Durbin 모형을 본떠서 명명된 것임.

14) Elhorst (2014)에 따르면, SDM과 SDEM은 구분이 어려울 정도로 유사하고, GNS는 식별문제가 발생한다고 함.

다양한 패널 공간계량모형은 우도함수법(Maximum likelihood)이나 도구변수법, GMM법, Bayesian법 등 다양한 방식을 통해 추정할 수 있다. 본 논문에서는 우도함수법을 활용하여 추정한다.

3. 직접효과와 간접효과

공간계량모형에서 직접효과라는 것은 한 지역의 설명변수 변화가 해당 지역의 종속 변수에 미치는 영향을 나타낸다. 반면, 간접효과는 누출 효과(Spillover effect)라고도 하며, 한 지역의 설명변수 변화가 다른 지역, 특히 인근 지역의 종속변수에 미치는 영향을 나타낸다. 일반적으로 직접효과와 간접효과는 어떤 공간계량모형을 선택하느냐에 따라 그 정보를 얻을 수도 있고 없을 수도 있으며, 그 크기도 다르다.

우선 일반 패널계량모형(예를 들어 FE, RE)이나 SEM의 경우에는 직접효과와 간접효과를 구분할 수 없다, 특히 패널 SEM의 경우 오차항의 공간적 상관관계를 고려한다고 해도 간접효과에 관한 어떠한 정보도 제공해 줄 수 없다. 이는 식 (3)을 다시 정리해서 확인할 수 있다. 식 (5)에서 X_t 는 모든 설명변수를 나타내며, R_t 는 절편과 오차항 등 식의 나머지 부분이다.

$$\ln PM_t = X_t \beta + R_t \quad (5)$$

그러나 식 (5)에 기댓값을 취하면 오차항 부분이 사라지기 때문에 k^{th} 번째 설명변수에 대해 편미분을 구하면 직접효과인 β_k 만 나타나고, 간접효과는 파악할 수 없게 된다. 따라서 간접효과를 파악하기 위해서는 SEM은 적절치 않다. <표 4>는 본 논문에서 활용한 세 가지 공간계량모형의 직접효과와 간접효과를 정리해 놓은 것이다.

SEM과 달리 SAR의 경우에는 간접효과의 크기와 중요도를 파악할 수 있다. 앞서 식 (2)를 다시 정리하면 식 (6)과 같다.

$$\ln PM_t = (I - \lambda W)^{-1} (X_t \beta) + R_t \quad (6)$$

<표 5> 다양한 공간계량모형 구성에 따른 직접효과와 간접효과

	직접효과	간접효과
SEM	β_k	0
SAR	$(I - \lambda W)^{-1} \beta_k$ 의 대각행렬	$(I - \lambda W)^{-1} \beta_k$ 의 비대각행렬
SDM	$(I - \lambda W)^{-1} [\beta_k + W\theta_k]$ 의 대각행렬	$(I - \lambda W)^{-1} [\beta_k + W\theta_k]$ 의 비대각행렬

식 (6)에 기댓값을 취한 후, k^{th} 번째 설명변수에 대해 편미분을 구하면 다음 식 (7)과 같다.

$$\left[\frac{\partial E(\ln PM_t)}{\partial x_{1k}} \cdot \frac{\partial E(\ln P_t)}{\partial x_{Nk}} \right] = (I - \lambda W)^{-1} \beta_k \quad (7)$$

직접효과와 간접효과를 이해하기 쉽게 보려면 식 (7)의 공간 승수 행렬에 대한 무한급수 전개를 해보면 식 (8)과 같다.

$$(1 - \lambda W)^{-1} = I + \lambda W + \lambda^2 W^2 + \lambda^3 W^3 + \dots \quad (8)$$

여기서 첫 번째 단위행렬(I)의 비대각행렬은 0이므로 이를 β_k 에 곱한 것은 직접효과를 나타낸다. 두 번째 행렬인 λW 의 대각행렬은 <그림 1>의 공간가중치행렬의 특성에 따라 모든 원소의 값이 0이 되므로 직접효과는 없고, 간접효과만 존재한다. 세 번째 항 이상의 나머지 행렬에서는 직접효과와 간접효과 모두를 포함하고 있다. 그러나 SAR의 경우 Pinkse and Slade(2010)가 지적한 대로 모든 설명변수에 대해 동일한 간접효과와 직접효과가 나타나게 되어 실제 적용과 한계효과를 해석하기에 적합하지 않다.

마지막으로 종속변수의 내생적 공간상관성과 설명변수의 외생적 공간상관성을 모두 고려한 모형으로 공간터빈모형(SDM)의 직접효과와 간접효과를 살펴보자. 식 (4)를 행렬식으로 나타내고, 식 (8)과 같이 정리할 수 있다. 여기서 R_k 은 절편과 오차항 등 식의 나머지 부분¹⁵⁾을 나타낸다.

15) 이 부분은 기댓값을 취하면 상수가 됨.

$$\ln P_t = (I - \lambda W)^{-1} (X_t \beta + W X_t \theta) + R_t \quad (8)$$

식 (8)에 기댓값을 취한 후, k^{th} 번째 설명변수에 대해 편미분을 구하면 다음 식 (9)와 같다.

$$\left[\frac{\partial E(\ln P_t)}{\partial x_{1k}} \cdot \frac{\partial E(\ln P_t)}{\partial x_{Nk}} \right] = \left[\frac{\frac{\partial E(\ln P_{1t})}{\partial x_{1k}} \cdot \frac{\partial E(\ln P_{1t})}{\partial x_{Nk}}}{\frac{\partial E(\ln P_{Nt})}{\partial x_{1k}} \cdot \frac{\partial E(\ln P_{Nt})}{\partial x_{Nk}}} \right] \quad (9)$$

$$= (I - \lambda W)^{-1} \begin{bmatrix} \beta_k & w_{12}\theta_k \cdot w_{1N}\theta_k \\ w_{21}\theta_k & \beta_k \cdot w_{2N}\theta_k \\ \cdot & \cdot & \cdot \\ w_{N1}\theta_k & w_{N2}\theta_k \cdot \beta_k \end{bmatrix}$$

이 편미분 값(행렬)의 의미는 만일 한 지역의 특정 설명변수(x_{nk})가 변하면, 그 해당 지역의 종속변수(미세면지농도, $\ln P_n$)가 변할(직접효과)뿐만 아니라, 다른 지역의 미세면지농도($\ln P_{\neq n}$)도 변한다는 것(간접효과)이다. 여기서 대각행렬에 있는 값 $((I - \lambda W)^{-1} \beta_k)$ 은 직접효과를 나타내고, 비대각행렬에 있는 값들은 간접효과를 나타낸다. 결과적으로 만일 $\lambda = 0$ 이고 $\theta_k = 0$ 이면 모든 비대각행렬 값은 0이 되어서 간접효과는 사라진다. 여기서 주목할 점은 직접효과와 간접효과의 값은 표본에 있는 지역별로 다르게 나타나게 된다. $(I - \lambda W)^{-1} \beta_k$ 에서 $(I - \lambda W)^{-1}$ 의 값은 모든 지역에 대해 다른 값이 나오므로, 직접효과는 지역마다 다르다. 간접효과의 경우에도 $(I - \lambda W)^{-1}$ 의 값이 지역마다 다른데, 이는 $\lambda \neq 0$ 이고 $\theta_k \neq 0$ 이기 때문이다.

4. 패널자료 검증

1) 공간상관성 검사

앞서 <표 1>에서 살펴본 바와 같이, 단순상관관계 분석을 통해서도 지역 간 미세면지농도의 상호관계가 존재하는지 명확히 설명해 낼 수 없었다. 그러나 공간계량분석모형

에서 간접효과의 존재여부를 파악하면 지역 간 미세먼지농도의 상호영향을 확실히 파악할 수 있다. 앞서 3.3에서 본 바와 같이, 만일 간접효과가 존재한다면, 한 지역의 미세먼지농도 변화는 인접 지역의 미세먼지농도를 변화시킬 수 있고, 또한 한 지역의 설명변수 변화도 인접 지역의 미세먼지농도를 변화시킬 수 있다. 따라서, 공간상관성이 존재하는지를 검증한다는 것은 간접효과가 존재하는지를 검증하는 것과 동일하다. 일반적으로 공간상호영향을 검정하는 방법은 Moran I를 통해 검정한다.¹⁶⁾ 그러나 패널데이터에는 이 검증방법은 적절치 않으며, 패널데이터의 횡단면 상관관계 검증방법을 활용하여 공간패널의 상호영향을 검증할 수 있다. 특히 $N > T$ 인 경우에는 Pesaran's CD 검증을 활용한다. 여기서 $\hat{\rho}_{ij}$ 는 상관관계 표본 추정치이다. 귀무가설은 공간상호영향이 없는 것이다. 이 검정통계량은 점근적으로 χ^2 분포하며, 자유도는 $N(N-1)/2$ 가 된다.

$$Pesaran's\ CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij} \right) \quad (10)$$

검정결과 검정통계량 값이 4.337로 계산되어, 귀무가설은 기각되었다. 따라서 패널데이터는 간접효과가 존재하며, 지역 간 미세먼지농도의 상호연관성이 존재한다.

2) 이분산 검정과 자기상관성 검증

본 논문의 패널데이터는 그룹 간 이분산이 존재하는지를 판단해야 한다. Green(2000)에 따르면, Modified Wald 검정통계량은 동분산이라는 귀무가설 하에 $\chi^2_{N_g}$ 분포한다고 한다. 여기서 N_g 는 횡단면 단위 개수이다. 검정결과 미세먼지 농도모형에서 그룹 간 이분산이 존재하는 것(Modified Wald test 검정통계량=157.95, p-value=0.00)으로 나타났다.

본 논문의 패널데이터는 시계열 개체 간 상관성 문제가 존재하는지 파악할 필요가 있다. 만일 선형 패널데이터에 자기상관이 존재하는 경우에는 표준오차에 편의가 발생하게 되어 추정량의 효율성이 낮아진다(Drukker, 2003). 패널데이터의 자기상관성을 검정

16) 인접한 지역의 변수들을 비교하여 통계량을 산출하는 것으로 그 값이 [-1, 1] 사이에 놓이게 된다. 만일 Moran I 값이 0에 가까우면 공간상호영향이 없는 것으로, 1이면 양의 공간 자기상관, -1에 가까우면 음의 공간 자기상관이 있다고 해석한다.

하기 위해 Wooldridge F-test를 이용한다. 귀무가설은 자기상관성이 없는 것이다. 검정 결과 미세먼지 농도모형에서 1차 자기상관성이 존재하는 것(Wooldridge F-test=58.975, p-value=0.00)으로 나타났다. 패널데이터 모형에서 그룹 간 이분산성만 존재하는 경우 패널고정효과 또는 확률효과모형을 적용하여 제거할 수 있으나 시계열 자기상관성까지 동시에 존재하면 패널일반화최소자승법(GLS)을 적용해야 한다.

3) 모형적합도 검증

우선 미세먼지농도 패널에 대한 Breusch-Pagan LM검증을 통해서, 단순 OLS 모형으로 분석할 수 있는지 파악한다. 분석결과 1% 유의수준에서 귀무가설이 기각되어 단순 OLS 모형은 적합하지 않다. 두 번째 하우스만 검정을 통해 고정효과모형이 더 적합하다고 판단하였다.¹⁷⁾

공간계량모형 중 어떤 모형이 주어진 데이터를 가장 설명하는지를 판단하기 위해서는 LR(Likelihood Ratio) 검정을 활용한다. LeSage and Pace(2009)는 SDM을 기본 모형으로 설정하고 LR 검정을 통해 가능하면 더 단순한 모형(SAR 혹은 SEM)을 활용해야 한다고 한다. 본 논문도 이 제안에 따라 식 (4)에 대해 다음의 두 가지 검정을 한다. 우선 SDM과 SAR 사이에 판단에서는 귀무가설을 $H_0 : \theta = 0$ 로 두고, 이 귀무가설을 기각할 수 없으면 더 단순한 SAR을 사용한다. 한편, SDM가 SEM 사이에 판단에서는 귀무가설로 $H_0 : \theta + \lambda\beta = 0$ 을 설정하고, 이를 기각할 수 없으면 더 단순한 SEM을 사용한다. 만일 두 개의 귀무가설을 동시에 기각하면 SDM이 데이터를 가장 잘 설명한다고 판단한다. LR검정 결과 첫 번째 귀무가설 $H_0 : \theta = 0$ 에 대해서 1%의 유의수준에서 기각하였다.¹⁸⁾ 또한, 두 번째 귀무가설 $H_0 : \theta + \lambda\beta = 0$ 에 대해서도 1%의 유의수준에서 기각하였다.¹⁹⁾ 즉 LR검정 결과 두 개의 귀무가설이 동시에 기각되어 주어진 패널데이터를 가장 잘 설명할 수 있는 공간계량 모형은 SDM으로 판단된다.

17) $\chi^2 = 27.20$, $p = 0.0013$

18) $\chi^2 = 20.82$, $p = 0.002$

19) $\chi^2 = 30.05$, $p = 0.0001$

IV. 분석결과

미세먼지농도에 대한 추정결과는 아래 <표 6>에 정리되어 있다. 여기서 SDM①과 SDM②는 ‘공간가중치 행렬’로 무엇을 활용했느냐의 차이이다. SDM①은 인접 여부의 공간가중치 행렬을 이용한 경우이며, SDM②는 거리(좌표)로 구성된 공간가중치 행렬을 이용한 경우이다. 비교를 위해, 패널데이터에 시간적 자기상관이 존재할 때 적용하는 패널일반화최소자승법(GLS, Generalized Least Square) 추정 결과도 함께 제시하였다. 또한, GLS와 일대일 비교를 위해 SDM들은 주효과만 정리하였다.

1. 편미 발생 가능성

우선, 종속변수의 공간종속을 나타내는 변수($W*\ln PM_t$)의 추정치(λ)가 SDM①에서는 0.231이고, SDM②에서는 0.166으로 두 값 모두 양(+)이며 통계적으로 유의한 결과가 도출되었다. 이는 미세먼지농도는 지역 간 상호영향을 미칠 수 있다는 것을 의미한다. 한 지역의 미세먼지농도가 올라가면 인접 지역의 미세먼지농도도 올라가며, 그 역(逆)도 성립한다.

<표 6> 미세먼지배출모형 추정결과

변수		GLS	SDM①	SDM②
주 효 과	$\ln Y_t$	82.240(10.38)***	70.952(2.68)***	80.231(3.96)***
	$(\ln Y_t)^2$	-8.263(-10.58)***	-7.052(-2.72)***	-7.771(-3.91)***
	$(\ln Y_t)^3$	0.276(10.76)***	0.233(2.75)***	0.251(3.86)***
	$\ln RD_t$	0.114(13.53)***	0.330(2.83)***	0.545(3.95)***
	$\ln PS_t$	-0.0001(-0.56)	-0.005(-0.79)	-0.005(-0.80)
	$\ln Yellow_t$	0.468(8.53)***	0.037(2.17)**	-0.012(-0.69)
	$\ln NGC$	0.003(9.06)***	-0.001(-0.80)	-0.0004(-0.24)
	상수	-269.68(-10.07)***	-	-
	$W*\ln PM_t$	-	0.231(2.73)***	0.166(1.70)*

주: ()에 있는 수는 t 값을 나타낸다.

*는 $p < 0.1$, **는 $p < 0.05$, ***는 $p < 0.01$ 를 나타낸다.

이제 미세먼지농도의 공간상관성이 존재하는 경우에 만일 이를 고려하지 않으면 나타날 수 있는 추정치의 편의 가능성을 검토해 보자. 우선 세 가지 추정방식 모두에서 미세먼지농도와 소득수준 간에는 N 자 형의 환경쿠즈네츠 곡선 가설이 검증되었다. <표 6>에서 보듯이 소득수준($\ln Y$)에 대해서 1차 항에서는 유의한 양의 추정치가, 2차 항에서는 유의한 음(-)의 추정치가 그리고 3차 항에서는 다시 유의한 양(+의 추정치가 도출되었다.

여기서 SDM①, SDM②의 추정치와 GLS의 추정치를 비교해 보면, 소득수준($\ln Y_t$)에 대해서는 GLS 추정치가 약간 과대 추정되었고, 제곱항($(\ln Y_t)^2$)에 대해서는 과소, 세제곱항($(\ln Y_t)^3$)에 대해서는 과대추정되었다는 것을 알 수 있다. 또한, 경유차 등록대수($\ln RD_t$)가 미세먼지농도에 미치는 영향은 GLS의 경우 과소 추정되었으며, 특히 SDM②와 비교할 경우 그 격차가 더 크게 나타난다.²⁰⁾ 즉 공간상호영향을 고려할 경우 경유 변수(혹은 경유소비량)가 미세먼지농도에 미치는 영향은 더 클 수 있다는 점이다. 한편, 황사의 영향($\ln Yellow$)도 GLS의 경우 SDM①에 비해 과대추정되었고, SDM②는 통계적으로 유의하지 않은 것으로 나타난다. 세 가지 계량모형 모두에서 석탄화력발전량($\ln PS_t$)의 추정치는 모두 통계적으로 유의하지 않았고, 화력발전 이외 석탄소비($\ln NGC$)의 경우 GLS는 통계적으로 유의하지만, SDM 모형들의 경우에는 그렇지 않은 것을 확인할 수 있다.

2. 직접효과와 간접효과

이번에는 (특정) 설명변수의 변화로 인한 인접 지역 종속변수(미세먼지농도)의 변화를 살펴보자. <표 7>에서는 공간터빈모형에 대해서만 추정결과를 보여준다, 단 주효과는 제외한다.

우선 가장 눈에 띄는 것은 SDM①의 경우에는 특정 지역의 소득수준 변화가 인접 지역의 미세먼지농도에 미치는 영향(θ_k)이 통계적으로 유의하게 영향을 미치지 않는 것으로 나타나지만(소득변수의 1차, 2차, 3차항 모두 통계적으로 유의하지 않음), SDM②의 경우에는 10% 유의수준에서 통계적으로 영향을 미치는 것으로 나타났다. 영향을 미치

20) 이러한 편의는 본 논문에서 '경유 소비량' 대신 '경유차 등록 대수'라는 대리변수를 사용했기 때문에 발생했을 수 있다. 즉 등록은 서울에서 하고 운행은 인접 지역인 경기나 인천에서 하는 경우가 있다.

는 것도 환경쿠르네츠 곡선 가설을 바탕으로 낮은 소득수준에서는 주변 지역의 미세먼지 농도를 높이다가 일정 수준 이상의 소득수준에서는 오히려 낮추고, 더 높은 소득수준에 대해서는 주변 지역의 미세먼지 농도를 다시 높이는 것으로 나타났다. 이러한 공간간 영향은 <표 7>의 간접효과에서도 확인할 수 있다. 즉 SDM①의 경우에는 소득수준의 간

(표 7) 미세먼지농도 모형 추정결과 (공간계량모형)

변수		SDM①	SDM②
θ_k	$W^* \ln Y_t$	14.373(0.28)	83.830(1.92)*
	$W^* (\ln Y_t)^2$	-1.263(-0.25)	-7.893(-1.84)*
	$W^* (\ln Y_t)^3$	0.036(0.22)	0.244(1.73)*
	$W^* \ln RD_t$	-0.135(-0.81)	-0.159(-0.84)
	$W^* \ln PS_t$	-0.037(-3.24)***	-0.070(-2.77)***
	$W^* \ln Yellow_t$	0.004(0.20)	0.049(2.65)***
	$W^* \ln NGC$	0.008(2.48)**	0.014(1.69)*
직접효과	$\ln Y_t$	73.548(2.50)***	84.820(4.06)***
	$(\ln Y_t)^2$	-7.300(-2.54)***	-8.209(-3.99)***
	$(\ln Y_t)^3$	0.241(2.57)***	0.265(3.94)***
	$\ln RD_t$	0.319(2.98)***	0.540(4.19)***
	$\ln PS_t$	-0.007(-1.21)	-0.007(-1.26)
	$\ln Yellow_t$	0.039(2.42)**	-0.009(-0.55)
	$\ln NGC$	-0.001(-0.50)	0.0001(0.06)
간접효과	$\ln Y_t$	32.582(0.55)	111.319(2.47)**
	$(\ln Y_t)^2$	-3.050(-0.53)	-10.511(-2.37)**
	$(\ln Y_t)^3$	0.094(0.50)	0.327(2.23)**
	$\ln RD_t$	-0.080(-0.44)	-0.090(-0.46)
	$\ln PS_t$	-0.043(-2.96)***	-0.081(-2.57)**
	$\ln Yellow_t$	0.013(0.78)	0.053(3.10)***
	$\ln NGC$	0.009(2.29)**	0.017(1.70)*
Log-likelihood		258.04	273.69

주: ()에 있는 수는 t 값을 나타낸다.

*는 $p < 0.1$, **는 $p < 0.05$, ***는 $p < 0.01$ 를 나타낸다.

접효과의 추정치가 통계적으로 유의하지 않지만, SDM②의 추정치는 5% 수준에서 유의하게 나타났고, 그 영향을 미치는 형태도 환경쿠즈네츠 곡선 가설을 따른다는 것을 확인할 수 있다. 앞서 본 바와 같이, SDM①은 공간가중치행렬(W)을 인접 유무에 대해 0 또는 1로 부여하고 열(列)을 기준으로 정규화시킨 것이고, SDM②는 공간가중치행렬을 거리에 반비례하도록 구성하고 이를 행(row) 기준 정규화시킨 것이다. 따라서 두 공간계량모형의 차이는 오로지 공간가중치행렬을 어떻게 구성했느냐에 달려 있다. SDM②만 유의한 추정치가 나온 것은 특정 지역 소득수준 변화가 시·도 경계를 접하고 있지 않지만, 거리상 가까운 지역들의 미세먼지농도에도 충분히 영향을 미쳤기 때문이라 예측된다. 혹은 그 반대로 시·도 경계는 접하지만 사실 거리상 먼 지역의 미세먼지농도에는 영향을 덜 미쳤을 것으로 추측된다. 결과적으로 공간상관관계가 존재하는 경우 미세먼지농도에 영향을 미치는 요인을 분석하는 경우에는 거리에 반비례하게 공간가중치 행렬을 구성하는 것이 특정 지역의 설명변수(요인)의 영향의 규모나 방향을 정확하게 분석하는 것뿐 아니라 인접 지역의 미세먼지농도에 대한 영향을 분석하는 데도 적합한 접근인 것으로 판단된다.

둘째, 인접 지역의 경유차 등록대는 특정 지역의 미세먼지농도에는 영향을 미치지 않는 것으로 분석되었다.²¹⁾ 또한, 이는 경유차 등록 대수는 직접효과만 존재하고 간접효과는 존재하지 않는 것으로도 확인할 수 있다. 이를 통해 경유차 등록대수(혹은 경유 소비량)는 해당 지역의 미세먼지농도만 높일 뿐, 인접 지역의 미세먼지농도에는 영향을 미치지 못하고 있다고 볼 수 있다. 그리고 앞서 본 바와 같이, 경유차 등록 대수가 해당 지역의 미세먼지농도에 미치는 영향은 공간상관성을 고려하지 않을 경우(GLS), 양(+)의 영향을 미칠 수 있으나 그 크기가 과소평가될 수 있다.

셋째, 석탄화력발전량은 인근 지역의 미세먼지농도를 낮추는 방향으로 영향을 미친다.²²⁾ 이는 직접효과는 통계적으로 유의하지 않고 간접효과만 존재하는 것으로 확인할 수 있다. 앞서 <표 5>에서 석탄화력발전이 미세먼지농도에 영향에 관한 GLS 추정치 가음(-)으로 나타났으나 통계적으로 유의하지 않은 부분과 연관지어 생각해 보자. 일반적으로 GLS 모형은 직접효과와 간접효과를 구분할 수 없다. 그러나 GLS모형과 달리

21) $W*\ln RD_i$ 의 추정치는 SDM①, SDM② 모두에서 통계적으로 유의하지 않다.

22) $W*\ln PS_i$ 의 추정치는 SDM①에서 -0.037로, SDM②에서 -0.07로 유의하게 추정된다.

SDM 모형들에서는 직접효과와 간접효과를 구분함으로써 석탄화력발전의 미세먼지농도에 대한 음(-)의 영향이 대부분 주변 지역에만 영향을 미치는 간접효과로 인한 것이라는 것을 명시적으로 파악하게 한다. 좀 더 정확하게는, 석탄화력발전량이 1% 증가할 때, 인접 지역의 미세먼지농도를 SDM①의 경우 0.043%, SDM②의 경우 0.081% 감소하는 것으로 해석할 수 있다. 이는 기존에 특정 지역의 석탄화력발전량이 해당 지역의 미세먼지농도를 높이거나 인접 지역의 미세먼지농도를 높일 수 있다는 주장의 설득력을 낮추는 결과이다. 기존 논문(배정환·김유선, 2017)에서도 한국의 석탄화력발전량이 유의하게 미세먼지농도에 영향을 미치지 못한다는 것과 일맥상통하며, 본 논문에서는 이러한 영향이 주 효과에서는 발견될 수 없지만, 간접효과에서는 음(-)의 영향이 나타나고 있다는 것을 발견하였다. 물론 이러한 결과 도출된 데는 본 논문에서 분석하는 데이터가 PM_{10} 으로, 만일 총먼지(TSP)나 초미세먼지($PM_{2.5}$)의 경우에는 다른 결과가 도출될 가능성도 있다. 이는 추후 연구에서 밝혀볼 필요가 있다.²³⁾

넷째, 인접 지역의 화력발전을 제외한 다른 석탄소비량(NGC)은 특정 지역의 미세먼지농도를 유의하게 양(+)으로 영향을 미치는 것으로 파악된다. 이는 직접효과는 없으나 간접효과는 존재하는 것으로 확인할 수 있다. 이는 특별히 <표 6>의 GLS 추정결과와 분명한 차이를 보인다. GLS 추정결과에서는 한 지역의 ‘화력발전을 제외한 다른 석탄소비량’은 해당 지역의 미세먼지농도만 높이는 것으로 추정된 데 비해, SDM 추정결과를 통해 얻은 결론은 해당 지역만 영향을 미치는 직접효과는 존재하지 않고, 간접효과를 통해 인근 지역의 미세먼지농도를 높이는 데 기여한 것으로 나타났다.

V. 결론 및 한계점

본 논문에서는 미세먼지농도에 대한 계량모형에 공간상관성을 고려할 경우 기존 논문과 어떤 차이가 발생하는지를 분석하였다. 추정결과 첫째, 공간상관성을 고려하지 않을 경우, 기존의 패널계량분석방법론으로는 추정치가 과소 혹은 과대 추정될 가능성이 존재한다는 것을 발견하였다. 이는 정부의 미세먼지농도 저감을 위한 정책의 규모나 방

23) 이처럼 석탄화력발전량의 추정치가 음(-)의 부호로 나타난 데는 기후변수나 일반사업장 변수 등의 설명변수가 모형에서 누락되었기 때문일 수 있다.

향에서 잘못된 판단을 내릴 수 있게 할 수 있으므로, 인접한 지역 단위의 미세먼지농도에 영향을 미치는 요인을 분석할 때에는 공간상관성을 반드시 고려해야 한다고 결론 내릴 수 있다.

둘째, 본 논문에서는 미세먼지농도 분석에서는 시·도 경계를 접하는 기준으로 공간가중치 행렬을 구성하기보다는 거리에 반비례하게 공간가중치 행렬을 구성하는 것이 미세먼지농도에 영향을 미치는 요인의 영향도를 분석하는 데 더 적절하다는 것을 발견하였다.

셋째, 소득수준은 해당 지역의 미세먼지농도를 증가시킬 뿐 아니라 인접 지역의 미세먼지농도에도 양(+)의 영향을 미치는 것으로 밝혀졌다. 이는 공간상관성을 고려하지 않는 일반적인 계량분석(본 논문에서는 GLS)에서는 밝혀낼 수 없었던 부분으로, 특히 공간가중치행렬(IV)을 지역 간 거리에 반비례하게 구성할 때 분명하게 나타나게 된다. 한 지역의 소득수준 증가는 그 지역의 소비량, 특히 제조업과 가정과 상업 분야의 난방용 보일러 연소량, 수송용 연료 사용량, 건설기계 사용량 그리고 비산먼지의 원인인 자동차 통행량을 증가시킨다. 결과적으로 미세먼지농도를 높일 가능성이 크다. 이렇게 높아진 미세먼지 농도는 인접한 지역의 미세먼지 농도에도 영향을 미칠 수 있다는 것을 밝혀냈다.

넷째, 경유차 등록 대수는 일반적으로 해당 지역의 미세먼지농도만 높이는 것으로 나타났다지만 공간상관성을 고려하지 않을 때는 그 효과가 과소평가될 위험성이 있었다. 이는 정책적으로 노후 경유차 관리의 필요성과 미세먼지 발생량이 적은 친환경자동차 확대의 필요성을 함의하고 있고, 이러한 정책이 지방자치단체별로 목표달성에 노력해야 함을 의미한다. 특히 이 과정에서 경유차 사용으로 인한 미세먼지의 영향을 과소평가하여 목표를 설정하지 않도록 해야 할 것이다.

다섯째, 화력발전용 이외의 유연탄 소비량은 해당 지역의 미세먼지농도를 변화시키는 부분보다 인근 지역의 미세먼지농도를 높이는 영향이 큰 것으로 나타났다. 이 부분은 공간상관성을 고려하지 않을 때는 직접효과와 간접효과를 구분할 수 없었다. 에너지경제연구원(2018)에 따르면, 화력발전용 이외 유연탄은 제철부와 시멘트 및 기타산업에서 소비하고 있다. 따라서 이들 산업이 입지하는 지역의 에너지 소비는 인근 주변 지역의 미세먼지농도를 높이는 데 큰 역할을 하는 것으로 나타났다. 따라서 이 부분은 인접 지역 간 주민의 미세먼지 피해구제에서 반드시 고려해야 하는 사안이며, 인접 지역 간 공동노력을 위해서도 고려해야 할 부분이다.

본 논문의 한계점은 다음과 같다. 첫째, 본 논문에서는 중국의 영향을 제외한 한국의 미세먼지농도의 발생요인을 분석하였다. 김진형·강성원(2018)에 따르면, 월별 미세먼지농도 데이터를 이용한 모형에서 중국의 영향을 포함했을 때 예측력이 높아짐을 보였다. 그러나 본 연구에서는 연도별 미세먼지농도 데이터를 이용하였기 때문에, 중국의 영향이 연간 평균값으로 모든 지역에 대해 동일하게 나타나게 되어 중국의 영향을 제외한 분석을 진행하였다. 둘째, 첫 번째 한계와 연관 지어 본 논문에서는 월별 데이터가 아닌 연간 데이터를 활용함으로써 월별·계절적 차이를 포착해내지 못했다. 본 논문이 미세먼지농도의 발생 요인으로 설정한 다양한 설명변수들이 대부분 연도별 데이터만 존재한다. 따라서 이들 설명변수와의 시간을 통일시키는 과정에서 월별 미세먼지농도 데이터의 12개월 평균값을 사용할 수밖에 없었다. 셋째, 본 논문에서 나타난 석탄 사용과 관련된 두 가지 특이한 현상은 추후 연구에서 밝혀봐야 할 부분이다. 추정 결과, 화력발전에서 사용된 석탄 사용의 증가는 인근 지역의 미세먼지농도를 낮추는 방향으로 영향을 미치고, 화력발전 이외의 석탄 사용의 증가는 해당 지역보다는 인근 지역의 미세먼지농도를 높이는 방향으로 영향을 미치는 것으로 나타났다. 이는 과학적 논리로 설명하기 다소 어려운 부분으로 미세먼지농도 데이터의 신뢰성에 관한 문제일 수도 있고, 각 영향이 너무 적어 부호 자체의 의미가 없을 수도 있다. 이러한 비논리적인 결과가 도출된 것은 첫째, 석탄화력발전소가 위치한 지역이 16개 시·도 중 소수에 불과해서 나타난 통계적 이상현상일 수 있으면, 둘째, 석탄화력발전소가 위치한 지역이 대체로 미세먼지 농도가 낮은 해안가에 위치하고 그 지역의 미세먼지 농도가 낮기 때문에 발생한 문제이거나, 셋째, 일별 혹은 월별이 아닌 연도별 데이터를 사용함으로써 석탄화력발전소의 미세먼지 농도에 미치는 영향을 보다 정확히 잡아내지 못한 한계일 수 있다. 이 부분에 대해서는 추후 연구를 통해 그 원인을 보다 정확히 파악해 볼 필요가 있다.

[References]

- 국토교통부 월간자동차등록현황, <https://www.data.go.kr/dataset/15024777/fileData.do>
 김진형·강성원, “국내 미세먼지 오염도에 영향을 미치는 요인에 대한 분석”, 「한국환경경

- 제학회 하계 학술대회논문집, 2018권 0호, 2018, pp. 779~791.
- 기상청 황사관측일수, <http://www.weather.go.kr/weather/asiandust/observday.jsp>
- 대한석탄협회, <http://www.kcoal.or.kr/info/info01.php>
- 배정환·김유선, “우리나라 대기오염배출 원인과 저감 정책 효과 분석”, 「자원·환경경제연구」, 제25권 제4호, 2016, pp. 545~654.
- 에너지경제연구원, 『2018 에너지통계 연보(yearbook of energy statistics)』, 2018.
- 윤인주·한상연, “우리나라 대도시의 총생산과 환경오염의 관계에 관한 실증연구”, 「한국정책연구」, 제10권 제1호, 2010, pp. 249~263.
- 이광훈, “국내 지역별 이산화탄소 배출에 관한 환경 쿠즈네츠 곡선 추정 및 비교”, 「환경정책연구」 제9권 제4호, 2010, pp. 53~76.
- 통계청 지역계정(Regional Account),
http://kosis.kr/statisticsList/statisticsListIndex.do?menuId=M_01_01&vwcd=MT_ZTITLE&pamTabl
- 한국전력 전력통계(2004~2015),
https://home.kepco.co.kr/kepco/KO/ntcob/list.do?boardCd=BRD_000099&menuCd=FN05030103
- 환경부(2004~2015), 에너지코리아 대기환경연보(Annual Report of Air Quality in Korea),
https://www.airkorea.or.kr/web/detailViewDown?pMENU_NO=125
- Anselin, L., “Model validation in spatial econometrics: a review and evaluation of alternative procedures”. *International Regional Science Review*, Vol. 11, No. 3, 1988, pp. 279~316.
- Anselin, L., “Spatial effects in econometric practice in environmental and resource economics”. *American Journal of Agricultural Economics*, Vol. 83, No. 3, 2001, pp. 705~710.
- Biles, J., “Using Spatial Econometric Techniques to estimate spatial multipliers: An assessment of regional economic policy in Ycantan, Mexico”, *The Review of regional Studies*, Vol. 33, No. 2, 2003, pp. 121~141.
- Cole, M. A., A. J. Rayner, and J. M. Bates, “The Environmental Kuznets Curve: An Empirical Analysis,” *Environment and Development Economics*, Vol. 2, No. 4, 1997, pp. 401~416.
- Drukker, D. M., “Testing for serial correlation in linear and panel-data models”, *The Stata Journal*, Vol. 3, No. 2, 2003, pp. 168~177.
- Elhorst, J. P., “Applied spatial econometrics: raising the bar”, *Spatial Economic Analysis*, Vol. 5, No. 1, 2000, pp. 9~28.
- Elhorst, J. P., *Spatial Economics from cross-sectional data to Spatial Panels*, Springer Brief

- in *Regional Science*, Springer, 2014.
- Green W. H., *Econometric Analysis*, 2000.
- Grossman, G. and A. Krueger, “Economic growth and the environment”, *Quarterly Journal of Economics*. Vol. 110, 1995, pp. 353~377.
- Hao, Y., and Y.-M. Liu, “The influential factors of urban PM2.5 concentrations in China: a spatial econometric analysis”. *Journal of Cleaner production*, Vol. 112, 2016, pp. 1443~1453.
- Kang, Y., T. Zhao and Y. Yang, “Environmental Kuznets curve for CO₂ emissions in China: A Spatial panel data approach”, *Ecological Indicators*, Vol. 63, 2016, pp. 231~239.
- Kelejian, Harry H., and Ingmar R. Prucha, “Estimation of spatial regression models with autoregressive errors by two-stage least squares procedures: A serious problem”, *International Regional Science Review*, Vol. 20, 1997, pp. 103~111.
- Keller, W., “International technology diffusion”, *Journal of Economic Literature*, Vol. 42, No. 3, 2004, pp. 752~782.
- LeSage, J., and R. K. Pace, *Introduction to Spatial Econometrics*. CRC press, 2010.
- Maddison, D. J., “Environmental Kuznets curves: a spatial econometric approach”. *Journal of Environmental Economics and Management*. Vol. 51, 2006, pp. 218~230.
- Mulligan, G. and L. Gibson, “Regression Estimates of Economic Base Multipliers for small communities”, *Economic Geography*, Vol. 60, No. 3, 1984, pp. 225~237.
- Myers, R. H., *Classical and modern regression application*. 2nd edition. Duxbury press. CA, 1990.
- Pinkse, J. and M. Slade, “The Future of Spatial Econometrics”, *Journal of Regional Science*, Vol. 50, No. 1, 2010, pp. 103~117.
- Poon, J. P., I. Casas, and C. He, “The impact of Energy, Transport, and Trade on Air pollution in China”, *Eurasian Geography and Economics*, Vol. 47, No. 5, 2006, pp. 568~584.
- Rothman, D. S., “Environmental Kuznets curves-real progress or passing over the buck?”, *Ecological Economics*, Vol. 25, 1998, pp. 177~194.
- Rupasingha, E., S. Goetz, D. Debertin, and A. Pagoulatos, “The environmental Kuznets curve for us countries: a spatial econometric analysis with extensions,” *Papers in*

- Regional Science*, Vol. 83, 2004, pp. 407~424.
- Selden, T., and D. Song, “Environmental Quality and Development: Is There a Kuznets Curve for Air Pollution Emissions?,” *Journal of Environmental Economics and Management*, Vol. 27, 1994, pp. 147~162.
- Smirnov, O., “Measuring Self-sustainability of Economic Development at th Country level’, paper presented at the 39th annual meeting of the Southern Region Science Association, Miami, FL, 2000.
- Stern, D. I., “Applying recent development in time series econometrics to the spatial domain”, *Professional Geographer*, Vol. 52, No. 1, 2000, pp. 37~49.
- Stern, D. I., “The Rise and Fall of the Environmental Kuznets Curve,” *World Development*, Vol. 32, No. 8, 2004, pp. 1419~1439.
- Suri, V., and D. Chapman, “Economic Growth, Trade, and Energy: Implications for the Environmental Kuznets Curve,” *Ecological Economics*, Vol. 25, No. 2, 1998, pp. 195~208.
- Xiangyu, G., Z. Zhou, Y. Zhou, X. Ye, and S. Liu, “A spatial panel Data Analysis of Economic Growth, Urbanization, and NO_x Emission in China”, *International Journal of Environmental Research and Public Health*, Vol. 15, No. 4, 2018.