존 데이터 기반 수단분담모형에 관한 연구

류시균^{1*} · 노정현² · 김지은¹

¹ 경기개발연구원 교통연구부, ² 한양대학교 도시대학원

A Study on the Modal Split Model Using Zonal Data

RYU, Sikyun^{1*} · RHO, Jeong Hyun² · KIM, Ji Eun¹

¹ Department of Transportation, Gyeonggi Research Institute, Gyeonggi 440-290, Korea

² Graduate School of Urban Studies, Hanyang University, Seoul 133-791, Korea

Abstract

This study introduces a new type of a modal split model that use zonal data instead of cost data as independent variables. It has been indicated that the ones using cost data have deficiencies in the multicollinearity of travel time and cost variables and unpredictability of independent variables. The zonal data employed in this study include (1) socioeconomic data, (2) land use data and (3) transportation system data. The test results showed that the proposed modal split model using zonal data performs better than the other does.

본 연구에서는 수단별 비용변수를 주요 설명변수로 활용하고 있는 현행 수단분담모형의 문제점으로서 설명변수간 높은 상관관계로 인한 다중공선성 문제와 버스노선의 가변성으로 인한 설명변수의 장래치 추정불가능성 문제를 지적하고 이와 같은 문제점을 극복할 수 있는 방안으로서 존을 설명하는 사회경제적 변수, 토지이용변수, 교통체계변수들을 설명변수로 하는 ‘존 데이터 기반 수단분담모형’의 활용가능성을 검증하였다. 장래교통수요추정모형으로서 수단분담모형의 설명변수가 갖추어야 할 조건으로서 목표연도별 설명변수의 추정가능성을 설정하고 이러한 조건을 만족하는 존 데이터를 설명변수로 한 수단분담모형을 구축하였으며 수단별 비용변수를 주요 설명변수로 하는 수단분담모형과의 비교를 수행하였다. 추정된 계수에 대한 통계적 유의성 검정에서 비용변수간 높은 상관관계로 인한 다중공선성 문제를 확인할 수 있었으며 적합도 평가(우도비의 비교)를 통해서 존 데이터 기반 수단분담모형이 수단별 비용변수를 설명변수로 한 수단분담모형에 비해서 설명력이 더욱 높다는 사실이 확인되었다.

Key Words

Modal Split Model, Mode Choice Model, Multicollinearity, Logit Model, PA Based Trip, OD Based Trip
수단분담모형, 로짓모형, 다중공선성, 사회경제지표, PA기반 목적통행, OD기반 목적통행

* : Corresponding Author
ryu@gri.re.kr, Phone: +82-31-250-3282, Fax: +82-31-250-3113

I. 서론

1. 연구의 배경 및 목적

4단계 교통수요추정모형은 여러 가지 단점에도 불구하고 가장 보편적인 교통수요추정모형으로 광범위하게 활용되고 있다. 수단분담은 통행분포와 통행배정의 사이에서 목적 OD를 수단 OD로 전환시켜주는 역할을 하며 수단별 비용변수(본 연구에서는 통행시간변수와 통행비용변수를 묶어서 비용변수라 한다)들을 핵심적 설명변수로 한 로짓모형이 주로 활용되고 있다.

그러나 수단별 비용변수는 이론적 측면과 현실적 측면에서 몇 가지 문제점을 안고 있는데 이론적 측면에서는 비용변수간 높은 상관관계로 인한 다중공선성 문제가 대표적이고 현실적 측면에서는 설명변수의 장래치 추정 불가능성 문제가 대표적이다. 비용변수간 높은 상관관계로 인한 다중공선성 문제는 추정된 파라메타의 신뢰도 저하 뿐만 아니라 파라메타의 값 자체도 왜곡시키는 것으로 알려져 있기 때문에 변수변환이나 대리변수의 활용 등, 다중공선성 문제를 완화시키기 위한 노력이 필요하다(류시균, 2008). 한편, 교통수단별 비용변수들의 장래치 추정 불가능성 문제는 특히 노선의 신설 및 폐지가 자유로운 버스서비스에서 두드러지게 나타난다. 현재의 버스서비스가 10~20년 후에도 유지될 것이라는 가정 자체가 성립되기 어려울 뿐만 아니라 대규모 신도시 개발계획을 대상으로 한 교통수요추정에 있어서는 버스서비스의 공급계획 자체가 불명확하기 때문에 기존점만 버스서비스의 비용추정은 불가능하다. 따라서 장래교통수요추정을 목적으로 한 수단분담모형의 설명변수로서 수단별 비용변수들은 적절치 않다고 할 수 있다.

본 연구는 수단별 비용변수들을 핵심적 설명변수로 채용하고 있는 장래교통수요추정 목적의 현행 수단분담모형의 문제점을 개선하기 위한 방안으로서 존의 사회경제적 변수, 토지이용변수 그리고 교통체계변수들을 핵심적 설명변수로 한 수단분담모형의 활용가능성을 검토하는데 목적이 있다.

2. 연구의 범위 및 방법

본 연구는 장래교통수요추정을 목적으로 한 수단분담모형, 엄밀한 의미에서는 로짓수단분담모형을 연구대상

으로 설정하고 기존 연구성과의 고찰을 통해서 수단별 비용변수들이 효용함수의 핵심적 설명변수로 활용되고 있음을 확인한다. 이어서 수단별 비용변수들이 안고 있는 문제점으로서 다중공선성 문제와 설명변수의 장래치 추정 불가능성 문제를 지적하고 존 데이터 기반 수단분담모형의 필요성을 제기한다.

본 논문의 제3장에서는 수도권가구통행실태조사자료를 토대로 대안집합을 구성하고 대안별 효용함수를 구축한다. 또한, 존을 설명하는 사회경제적 변수, 토지이용변수 그리고 교통체계변수들과 수단분담률간 상관관계를 검토함으로써 존 데이터 기반 수단분담모형의 설명변수로서 활용가능한 변수들을 선별한다.

제4장에서는 존 데이터 기반 수단분담모형과 비용변수 기반 수단분담모형을 구축하고 통계적 검정과 적합도 비교를 통해서 존 데이터 기반 수단분담모형의 활용가능성을 검증한다. 제5장에서는 존 데이터 기반 수단분담모형의 한계와 개선방향을 제시하고 향후 연구과제를 제시한다.

II. 기존연구 고찰

1. 현행 수단분담모형 고찰

본 연구에서는 현행 수단분담모형을 수단선택행태에 영향을 미치는 요인분석을 목적으로 한 모형(이하 해석적 모형이라 함)과 장래교통수요추정을 목적으로 한 모형(이하 실무적 모형이라 함)으로 구분해서 기존 연구성과를 고찰하였다. 해석적 모형은 개별 통행자들의 개인속성과 같은 장래교통수요추정에서는 활용이 불가능한 변수들이 설명변수로 활용되는 반면 실무적 모형에서는 개인속성변수와 같이 장래치 추정이 불가능한 변수들은 설명변수로 활용되지 않는 특징이 있다. 해석적 모형에 관한 기존 연구성과들을 정리하면 아래와 같다.

김순관(1999)은 개인속성과 교통수단속성 등을 설명변수로 한 네스티드로짓모형을 구축하여 제안하였다. 윤대식·육태숙·김상황(2006)은 대구시민의 통행수단선택행태에 관한 경험적 분석을 위해서 다항로짓모형을 구축하였으며 설명변수로는 개인속성과 교통수단속성을 활용하였다.

이순주(2007)는 수원시 출근통행자를 대상으로 교통수단속성(통행시간과 통행비용)과 개인속성(성별, 나이

등) 그리고 존의 사회경제지표를 설명변수로 한 다항로짓모형을 구축하였다.

전명진·백승훈(2008)은 조건부 로짓모형을 활용하여 통근통행의 수단선택요인을 추정하였다. 통행시간과 통행비용 외에도 통행자의 개인속성, 토지이용특성, 그리고 통근통행의 방향 등을 설명변수로 활용하였다.

이혜승·이희연(2009)은 각 교통수단의 효용에 의해 선택이 결정되는 조건부 로짓모형과 개개인의 사회·경제적 특성에 의해 교통수단이 결정되는 다항로짓모형을 결합한 혼합적 조건부 로짓모형을 구축하였다. 조건부 로짓모형에 사용된 독립변수로는 통행자 개개인의 사회·경제적 속성, 출발지와 도착지의 공간적 특성, 그리고 교통수단 속성(통행시간과 통행비용) 등이 활용되었다.

다양한 해석적 목적의 수단분담모형이 구축되었으나 모든 연구에서 공통적으로 교통수단별 비용변수들은 기본적인 설명변수로 활용되고 있음을 확인할 수 있다.

한편, 장래교통수요추정을 목적으로 한 실무적 모형으로서 서울시(2000)는 통행시간, 통행비용, 존의 인당

승용차 보유대수, 존의 지하철 역수, 존에서 지하철역간 거리, 출발·도착지 주차장 1급지, 출발·도착지 30만 이상 도시 등을 설명변수로 한 다항로짓모형을 활용하였다. 또한, 서울시(2004)의 '서울시 장래교통 수요예측 및 대응방안 연구'와 경기도(2004)의 '장래 교통수요 예측 및 대응방안 연구'에서는 비용변수, 도착지 종사자 수, 성별 변수를 이용한 다항로짓모형으로 수단분담모형을 구축하였으며 가장 최근에 완료된 수도권교통본부(2009)의 '수도권 장래교통 수요예측 및 대응방안 연구'에서는 통행목적별로 차내 통행시간, 차외 통행시간, 통행비용, 버스 환승횟수 등의 변수를 활용한 가중 네스티드 로짓모형, 가중 다항로짓모형, 일반 다항 로짓모형을 구축하였다. 마지막으로 박용수·류시균·오영택·김태호·노정현(2010)은 교통수단의 비용변수와 존의 특성을 반영할 수 있는 설명변수(토지이용, 교통서비스 특성)를 활용하여 수단분담모형을 개발하였다.

실무적 목적의 수단분담모형에서도 교통수단별 통행 비용변수들은 중추적 변수로서 활용되고 있으며 존을 설명하는 사회경제적 변수(승용차보유수준, 도착존의 종사자 수, 도시규모 등)나 교통체계변수(지하철역까지의 거리 등)들은 보조적 설명변수로 활용되고 있음을 확인할 수 있다. 한편, 장래치 추정이 불가능한 개인속성변수(성별, 나이 등)들이 실무적 모형에 도입된 사례도 있었으나 최근에는 발견되지 않고 있다.

한편, 수단 OD구축을 목적으로 한 미국의 대도시권에서는 네스티드 로짓모형이 폭넓게 활용되고 있음을 확인할 수 있다. 수도권교통본부(2009)는 미국의 올리언즈 모형과 불티모어 모형을 소개하면서 효용함수에 도입된 설명변수들을 제시하고 있다. 효용함수에 도입된 설명변수는 크게 통행시간변수(차내통행시간, 접근보행시간, 기종점에서의 소요시간, 대기시간, 환승시간), 통행비용변수(차량운영비용, 톨비, 주차비, 환승횟수), 존 특성(CBD, 존의 유형) 등이다.

동경도시권교통계획협의회(2001)는 일본 동경권의 퍼스트립조사에 기반한 장래교통수요예측에서 철도(4개 접근교통수단을 하부구조로 가짐), 자동차, 버스의 3개 대표교통수단을 대안집합으로 한 다항로짓모형을 제안하였다. 모형의 설명변수로서 개인속성변수(연령, 자동차 면허보유여부), 통행시간변수, 통행비용변수(운행비용, 주차비) 그리고 교통체계변수(존간 버스노선수)가 도입되었다. 한 가지 흥미로운 점은 자동차기여도라고 하는 변수를 비가정기반 목적통행을 대상으로 한 수단분담모

〈Table 1〉 Precedent Studies on the Modal Split Model of Analytical Approach

Researcher	Model	Variables
Kim, Soon-Gwan (1999)	Nested Logit Model	travel time, travel cost, vehicles per household, trip purpose dummy
Yun, Dae-Sic · Yuk, Tae-Suk · Kim, Sang-Hwang (2006)	Multinomial Logit Model	in-vehicle time, in-vehicle cost, out-of-vehicle time, out-of-vehicle cost, trip frequency, trip purpose, gender, age, job, household income, personal income
Lee, Soon-Joo (2007)	Multinomial Logit Model	travel time, travel cost, gender, age, geographic dummy, workers
Jun, Myung-Jin · Baek, Seung-Hun (2008)	Conditional Logit Model	travel time, travel cost, gender, age, householder, education level, job and residential location, residential population density, employment density, work-trip direction
Lee, Hye-Seung · Lee, Hee-Yeon (2009)	Joint Conditional Logit Model	age, gender, job, household income, development density, mixed use, bus accessibility, subway accessibility, travel time, travel cost

〈Table 2〉 Precedent Studies on the Modal Split Model of Practical Approach

Researcher	Model	Variables
서울시 교통정비 중기계획 (2000, 서울시)	Aggregate Multinomial Logit Model	travel time, travel cost, taxi dummy, bus dummy, vehicles per household, number of subway station, distance between subway stations, Origin · Destination parking lot, Origin · Destination 300,000 population city
서울시 장래교통 수요예측 및 대응방안 연구 (2004, 서울시)	Multinomial Logit Model	travel time, travel cost, destination workers, gender, age, other socioeconomic variables
장래 교통수요 예측 및 대응방안 연구 (2004, 경기도)	Multinomial Logit Model	travel time, travel cost, destination workers, gender, age, geographic dummy
수도권 장래교통 수요예측 및 대응방안 연구 (2009, 수도권교통본부)	Nested Logit Model, (Weighted) Multinomial Logit Model	in-vehicle time, out-of-vehicle time, travel cost, bus transfer availability, trip purpose
Park, Yong-Su · Ryu-Si -Kyun · Oh, Young-Taek · Ki m, Tae-Ho · Rho, Jeong-Hyun (2010)	Logit Model	area of office facilities, subway existence, highway IC existence

형의 설명변수로 도입하였다는 점이다. 자동차기여도 변수는 가정기반목적통행의 도착지 자동차분담률로 정의되는데 이는 가정기반 목적통행을 대상으로 한 수단분담모형의 추정결과를 비가정기반 목적통행을 대상으로 한 수단분담모형의 설명변수로 도입하고 있다는 점에서 특이하다고 볼 수 있다. 국외연구사례에서도 비용변수기반 수단분담모형이 주류를 이루고 있음을 확인할 수 있다.

2. 비용변수 기반 수단분담모형의 한계

1) 설명변수간 다중공선성 문제

현행 수단분담모형의 핵심변수인 비용변수들 사이에

〈Table 3〉 Correlation Coefficient between Travel Cost and Time Variables

Variables	Auto Cost	Auto Time	Transit Cost	Transit Time	Taxi Cost
Auto Time	0.8341				
Transit Cost	0.5775	0.7674			
Transit Time	0.8379	0.9826	0.7695		
Taxi Cost	0.8213	0.9586	0.8211	0.9420	
Taxi Time	0.8424	0.9898	0.7557	0.9960	0.9407

Note : 대중교통수단의 비용 및 시간은 버스의 일반화비용(시간×평균시간가치+비용)과 전철의 일반화비용을 비교해서 일반화비용이 낮은 교통수단의 시간 및 비용변수를 활용함

는 근본적으로 높은 상관관계가 존재하게 된다. 〈Table 3〉은 ‘수도권 장래교통 수요예측 및 대응방안 연구 (2009, 수도권교통본부)’의 수단분담모형 구축과정에서 설명변수로 활용된 교통수단별 비용변수간 상관관계를 분석한 결과로서 모든 설명변수들 사이에 극단적으로 높은 상관관계가 존재함을 확인할 수 있다.

다중공선성 문제는 회귀분석에서는 보편적으로 알려져 왔지만 수단선택모형으로 활용되고 있는 로짓모형과 관련해서는 많은 검토가 이루어지지 않았다. 최근 류시균 (2008)은 로짓모형에 있어서의 다중공선성 문제를 검토하였으며, 설명변수간 상관관계가 높을수록 추정된 파라메타의 신뢰도가 저하됨은 물론 파라메타의 값도 참값에서 멀어짐을 증명하였다. 실증적으로 본 연구의 수행과정에서도 교통수단별 비용변수간 높은 상관관계로 인한 다중공선성 문제를 확인할 수 있었다(본 논문의 제4장 참조).

2) 비용변수의 장래치 추정불가능성 문제

수단분담모형을 활용해서 장래 목표연도별 수단OD를 구축하기 위해서는 목표연도별 수단별 비용변수들이 추정되어야 한다. 승용차, 택시, 전철은 도로망과 전철망을 활용해서 추정하는 것이 가능하지만 노선의 신설과 폐지가 자유로운 버스의 비용변수는 현실적으로 추정이 불가능하다. 특히 비용변수의 장래치 추정불가능성 문제는 버스서비스의 구체적 공급계획이 제시되지 않는 대규모 도시개발계획에 대한 교통수요추정에서 더욱 두드러지게 나타난다. 즉, 교통수단별 비용변수를 활용해서 장래목표연도별 수단OD를 구축하는 것은 거의 불가능에 가깝다고 할 수 있으며 따라서 설명변수간 상관관계가 낮고 장래치 추정이 가능한 대안적 설명변수를 활용한 수단분담모형의 개발이 요구된다.

III. 대안집합 및 설명변수 선정

1. 대안집합 및 효용함수의 구축

1) 기초데이터 및 대안집합 구성

본 연구에서는 2006년도에 수행된 '수도권 가구통행 실태조사'의 원시 데이터(총 1,344,955통행으로 구성됨)를 활용해서 수단분담모형을 구축하였다. 상기 조사에서는 총 16개 교통수단을 대상으로 조사가 이루어졌으나 본 연구에서는 다음과 같이 16개 교통수단을 5개 대분류 교통수단으로 통합하고 수단간 대체효과가 미미한 지역간 교통수단과 도보를 제외한 3개 교통수단(승용차, 대중교통, 택시)을 대안집합으로 구성해서 수단분담모형을 구축하였다. <Table 4> 우측의 '우선순위'는 하나의 목적통행이 다수의 교통수단에 의해서 완료되었을 경우 대표교통수단을 선정하기 위한 순위이다(가령 승용차와 버스가 동시에 이용된 경우 버스(대중교통)가 대표교통수단으로 정의된다).

<Table 4> Classification Considering Similarity among Travel Modes

Category		Mode of Transportation	Priority
Walk		walk(1), bicycle(15), motorcycle(14)	6
Auto		auto drive(2), auto passenger(3)	5
Transit	Bus	commuter bus(4), inter-city bus(5), express-city bus(6), local bus(7), others(9)	3
	Metro/Subway	metro/subway(10)	2
Taxi		taxi(13)	4
Interregional Transportation		intra-city express bus(8), rail transit(11), KTX(12)	1

2) 효용함수의 구축

본 연구에서는 승용차, 대중교통, 택시를 대안집합으로 수단분담모형(식(1))을 구축하였으며 교통수단별 효용함수의 기본구조는 식(2) 및 식(3)과 같다.

- 수단분담모형의 기본형-

$$P_{ij}(k) = \frac{\exp(V_{ijk})}{\sum_{k' \in K_{ij}} \exp(V_{ijk'})} \quad (1)$$

여기서,

$P_{ij}(k)$: 출발존 i에서 도착존 j로의 통행자가 수단 k를 선택할 확률

K' : 기종점쌍 ij를 연결하는 교통수단대안집합 K_{ij} 의 원소(수단)

K_{ij} : 기종점쌍 ij를 연결하는 교통수단대안집합

V_{ijk} : 기종점쌍 ij를 연결하는 교통수단 k의 측정 가능한 효용

- 비용변수 기반 수단분담모형의 효용함수

$$U_{ij}(k) = f(c_{ijk}) + e_{ijk} = \alpha_k + \beta_k \cdot c_{ijk} + e_{ijk} \quad (2)$$

- 존 데이터 기반 수단분담모형의 효용함수

$$U_{ij}(k) = f(x_{ijk}) + e_{ijk} = \alpha_k + \gamma_k \cdot x_{ijk} + e_{ijk} \quad (3)$$

여기서,

$U_{ij}(k)$: 출발존 i에서 도착존 j로의 통행자가 교통수단 k에 대해서 인지하는 효용

c_{ijk} : 출발존 i, 도착존 j일 경우 교통수단 k의 비용변수(벡터)

x_{ij} : 출발존 i, 도착존 j 또는 기종점 존쌍 ij를 설명하는 존 데이터(벡터)

α_k : 대안고유더미(이하 상수항이라 한다)의 파라메타(벡터)

β_k : 비용변수의 파라메타(벡터)

γ_k : 존 데이터의 파라메타(벡터)

e_{ijk} : 기종점쌍 ij를 연결하는 교통수단 k의 측정 불가능한 효용

식(2)는 비용변수 기반 수단분담모형의 효용함수를, 식(3)은 존 데이터 기반 수단분담모형의 효용함수를 정의하고 있다.

2. 존 데이터 기반 모형의 설명변수 선정

1) 설명변수 대안검토

장래 목표연도별 수단 OD 구축을 목적으로 하는 수단분담모형의 설명변수가 갖추어야 할 첫 번째 조건은 장래목표연도의 설명변수값을 추정할 수 있어야 한다는 것이다. 본 연구에서는 시군별 통계연보, 통계청 통계연보, 대규모 개발계획 승인신청 조서 그리고 장래교통수

요예측에 활용된 변수들 중에서 장래치 추정이 가능한 다음 변수들을 설명변수 후보군으로 선정하였다. 광역행정구역특성, 도시화수준 및 기종점간 택시사업구역의 동질성은 범주형 변수(더미변수)로 표현되며, 가구당 승용차대수와 인구밀도는 연속형 변수로, 시가화면적비율(시가화면적/전체 행정구역면적), 종사자수비율(1/2/3차 종사자수/총종사자수), 수용학생수 비율(초/중/고/대학 수용학생수/총수용학생수), 연령대별 인구비율(연령대 인구/전체 인구)은 산정된 비율값(%값이 아님)으로, 지하철역까지의 거리는 Emme/2 네트워크에서 존 센트로이드 좌표와 지하철역 노드 좌표간 유클리디안 거리로 설명변수값이 산정되었다.

2) 수단분담률과 설명변수간 상관관계 검토

〈Table 5〉에서 제시된 변수들이 수단분담모형의 설명변수로 도입되기 위해서는 설명변수와 수단분담률간 인과관계가 존재해야 한다. 본 연구에서는 2006년 가구통행실태조사의 원시 데이터 중에서 통근목적의 통행을 대상으로 산정된 수단분담률과 설명변수간 상관관계에 관해서 검토하였으며, 본 논문에서는 지면제약을 고려해서 일부만을 수록하였다. 가령, 〈Figure 1〉과 〈Figure 2〉는 〈Table 5〉의 일부 변수들과 수단분담률간 관계를 플로팅한 것이다. 출발존이 서울시인 경우의 승용차분담

〈Figure 1〉 Relationship between Administrative District and Modal Split of Commuting Trips (Origin Zone)

〈Figure 2〉 Relationship between Urbanized Area Ratio and Modal Split of Commuting Trips (Destination Zone)

률은 매우 낮은 반면 대중교통분담률은 매우 높음을 알 수 있으며 결론적으로 존 데이터 기반 수단분담모형의 설명변수로서 활용가능성이 매우 높다고 할 수 있다.

Ⅳ. 존 데이터 기반 수단분담모형의 구축 및 평가

1. 모형의 추정 및 검정(OD 기반 통근통행)

본 연구에서는 2006년 수도권가구통행실태조사에서 구분한 통행목적(배웅, 귀가, 통근, 등교, 학원, 업무, 귀사, 쇼핑, 여가, 기타)별로 OD기반 수단분담모형을 구축하였다. 그러나 지면제약을 고려해서 본 논문에서는 OD기반 통근통행에 대한 수단분담모형의 추정결과만을 제시하고 나머지 9개 목적별 수단분담모형에 대해서는 적합도 비교결과만을 수록하였다.

1) 비용변수 기반 수단분담모형

비용변수기반 수단분담모형을 구축함에 있어서 사전적으로 고려해야 할 요소로서 모든 효용함수에 공통적으로 들어 있는 비용변수들(통행시간과 통행비용)에 대해

〈Table 5〉 Explanatory Variables of Modal Split Model Using Zonal Data

Type	Variables	Note
Category Variables	administrative district	Seoul/Incheon/Gyeonggi
	urbanization level	Eup/Myeon/Dong
	taxi district	
Continuous Variables	number of cars per household	vehicles
	population density	1,000/km ²
Proportion Variables	percent of urbanized areas	Gu/Eup/Myeon
	percent of workers	
	percent of students	
	percent of population into age groups	
Others	distance from a subway station	Emme/2 Coordinates of Network ¹⁾

1) 지하철역까지의 거리 = $\sqrt{[(x_1 - x_2)/1000]^2 + [(y_1 - y_2)/1000]^2}$ (여기서 (x_k, y_k) 는 Emme/2 네트워크에서 노드의 좌표값)

〈Table 6〉 Estimates of Modal Split Model Using Travel Cost Variables (Commute Trip)

Mode	Exploratory Variables	Expected Sign	Coefficient Estimate	Standard Error (STE)	t statistic
Auto	Auto Constant		3.6848	0.04997	73.7284
	Travel Cost (₩100)	-	-0.0624	0.00061	-101.58
	Travel Time (min)	-	-0.0511	0.00154	-33.041
Transit	Transit Constant		4.16200	0.04816	86.4080
	Travel Cost (₩100)	-	-0.0844	0.00120	-70.088
	Travel Time (min)	-	-0.0617	0.00104	-59.017
Taxi	Travel Cost (₩100)	-	-0.0336	0.00135	-24.945
	Travel Time (min)	-	-0.0191	0.00432	-4.4317
Sample Distribution	Mode	Total	Auto	Transit	Taxi
	Number of observations	193,161	99,381	92,077	1,703
	Distribution Ratio(%)	100.00	51.45	47.67	0.88
Log-Likelihood Ratio	$L(0)$			-212,209.0	
	$L(C)$			-142,321.7	
	$L(\hat{\alpha}, \hat{\beta})$			-128,621.6	
Log-Likelihood Ratio ³⁾	$\rho_1^2(0) = 1 - L(\hat{\alpha}, \hat{\beta}) / L(0)$			0.3938918	
	$\bar{\rho}_1^2(0) = (N-K) / N \times \rho_1^2(0)$			0.3938755	
	$\rho_2^2(C) = 1 - L(\hat{\alpha}, \hat{\beta}) / L(C)$			0.0962615	
	$\rho_3^2(\alpha) = 1 - L(\hat{\alpha}, \hat{\beta}) / L(\hat{\alpha})$			0.1402768	

서 공통의 파라메타를 갖도록 generic variable로 처리할 것인지 아니면 독립적으로 파라메타를 갖도록 처리할 것인지를 결정해야 한다. Generic variable로 처리하면 하나의 시간가치가 얻어지고 산정된 시간가치의 적정성 평가가 용이한 장점이 있는 반면 모형의 적합도가 낮아지는 단점이 있다. 반면 통행비용변수들에 대해서 자유롭게 파라메타를 취하도록 모형을 추정하면 수단별로 상이한 시간가치가 산정되는 단점이 있는 반면 모형의 적합도는 높아지는 장점이 있다²⁾. 본 연구에서는 존 데이터 기반 수단분담모형과 비용변수 기반 수단분담모형간 비교평가를 통해서 존 데이터 기반 수단분담모형의 활용

가능성을 검증하는데 목적이 있기 때문에 적합도 측면에서 유리한 모형이 구축되도록 모든 비용변수들이 자유롭게 파라메타를 취할 수 있도록 허용하였다.

〈Table 6〉은 OD기반 통근목적의 비용기반 수단분담모형 추정결과로서 수단별 비용변수 모두 부호검정 및 통계적 검정을 통과하였음을 보여주고 있다. 또한, 우도비($\rho^2(0)$) 역시 0.393을 상회하고 있어 모형의 전반적 설명력도 높다고 평가된다. 다만, 대안별 통행비용변수들의 설명력을 평가할 수 있는 지표인 새로운 우도비($\rho_3^2(\alpha)$, 각주 3) 참조)는 0.140으로 수단분담구조에 대한 통행비용변수들의 설명력이 높다고 평가하기는 쉽지 않다(모형의 우도비 0.393 중에서 상수항이 개선한 부분이 0.253인데 비해 대안별 비용변수가 개선한 부분은 0.140정도에 불과하기 때문이다).

2) 존 데이터 기반 수단분담모형

존별 또는 존간 사회경제지표, 토지이용변수 그리고 교통체계변수들로만 설명변수집합을 구성한 존 데이터 기반 수단분담모형(OD기반 통근목적)의 추정결과는 〈Table 7〉과 같다. 기대부호가 존재하는, 다시 말해서 논리적 추론이 가능한 파라메타를 대상으로 추정결과를 간략히 해석하면 아래와 같다.

통근통행의 경우 출발존이 인천<경기>서울 순으로 승용차 의존도가 높고 도착존은 경기<인천>서울 순으로 승용차 의존도가 높다. 이는 서울시의 대중교통서비스가 경기도나 인천시에 비해서 양호하고, 서울시의 경우 도로교통혼잡 및 높은 주차비 등으로 인해서 승용차 이용이 경제적이지 못하기 때문으로 해석할 수 있다.

출발존과 도착존 역시 공통적으로 면지역<읍지역>동지역의 순으로 승용차에 대한 의존도가 높는데 이 역시 대중교통서비스의 지역간 편차를 고려하면 납득할 수 있는 추정결과라 할 수 있다. 17세 이하 인구비와 승용차 분담률간 정(+)의 상관관계 역시 자녀의 등교를 지원하고 나서 출근하는 경우에는 승용차 이용이 불가피하다는

- 수단별로 추정된 통행시간 파라메타와 통행비용 파라메타를 이용해서 시간가치를 산정하면 비논리적인 수준의 시간가치가 산정되는 경우가 많기 때문에 시간가치가 여러 개일수록 추정된 모형의 논리적 정당성을 부여하기가 어려워지는 단점이 있다. 특히, 설명변수간 상관관계가 높은 통행비용변수들의 파라메타는 왜곡되어 있을 가능성이 높기 때문에 이들을 이용해서 산정된 시간가치 역시 왜곡되어 있을 가능성이 매우 높다.
- 일반적으로 우도비는 단순 우도비($\rho_1^2(0)$), 자유도로 보정된 우도비($\bar{\rho}_1^2(0)$) 그리고 상수항만을 설명변수로 도입했을 경우의 우도와 비교한 상수 우도비($\rho_2^2(C)$)가 활용되고 있다. 그러나 본 연구에서는 추정된 설명변수의 설명력을 평가하기 위해서 새로운 우도비($\rho_3^2(\alpha)$)를 추가로 도입하였다. 새로운 우도비는 대안고유터미(상수항)와 대안별 비용변수(또는 존 데이터)를 설명변수로 도입해서 모형을 추정한 다음 대안고유터미변수만을 이용해서 산정된 우도와 모형 전체의 우도를 비교한 지표다. 새로운 우도비를 통해서 대안고유터미를 제외한 설명변수들이 우도값 증대에 얼마만큼 기여했는지를 평가할 수 있다.

〈Table 7〉 Modal Split Model Using Zonal Data (Commute Trip)

Mode	Exploratory Variables	Expected Sign	Coefficient Estimate	Standard Error (STE)	t statistic
Auto	Auto Constant		3.18623	0.15652	20.3557
	Origin(O) Incheon Dummy	+	0.32957	0.02650	12.4362
	Gyeonggi Dummy (O)	+	0.13753	0.01522	9.03469
	Destination(D) Incheon Dummy	+	0.30112	0.02733	11.0169
	Gyeonggi Dummy (D)	+	0.33094	0.01846	17.9187
	Eup Dummy (O)	+	0.10346	0.04024	2.57126
	Myeon Dummy (O)	+	0.35588	0.04926	7.22362
	Eup Dummy (D)	+	0.20994	0.03567	5.88471
	Myeon Dummy (D)	+	0.63750	0.03743	17.0283
	Percent of Urbanized Areas (D)	-	-0.8116	0.04666	-17.395
	Percent of Tertiary Workers (O)		0.08475	0.03525	2.40386
	Percent of Tertiary Workers (D)		-0.5451	0.03004	-18.145
	Percent of Under 17 (O)	+	8.26108	0.13662	60.4656
	Percent of Under 17 (D)	+	3.72280	0.12496	29.7906
	Population Density (O)	-	-0.0050	0.00000	-11.650
	Population Density (D)	-	-0.0040	0.00000	-9.0027
	Distance from Subway Station (O)	+	0.02289	0.00248	9.20193
Transit	Transit Constant		5.15898	0.16050	32.1429
Taxi	Taxi District	+	1.36834	0.16264	8.41304
Sample Distribution	Mode	Total	Auto	Transit	Taxi
	Number of observations	193,161	99,381	92,077	1,703
	Distribution Ratio	100.00	51.45	47.67	0.88
Log-Likelihood Ratio	$L(0)$			-212,209.0	
	$L(C)$			-142,321.7	
	$L(\hat{\alpha}, \hat{\gamma})$			-126,449.1	
Log-Likelihood Ratio	$\rho_1^2(0) = 1 - L(\hat{\alpha}, \hat{\gamma})/L(0)$			0.4041294	
	$\bar{\rho}_1^2(0) = (N - K)/N \times \rho_1^2(0)$			0.4040896	
	$\rho_2^2(C) = 1 - L(\hat{\alpha}, \hat{\gamma})/L(C)$			0.1115227	
	$\rho_3^2(\alpha) = 1 - L(\hat{\alpha}, \hat{\gamma})/L(\hat{\alpha})$			0.4525156	

점을 감안하면 역시 논리적 설명이 가능하다. 출발존과 도착존 공통으로 인구밀도와 승용차분담률은 부(-)의 상관관계가 있는데 이 역시 인구밀도가 높은 지역일수록

양질의 다양한 대중교통서비스가 공급되기 때문으로 해석할 수 있으며 지하철역까지의 거리에 비례해서(떨어질수록) 승용차 분담률이 높아지는 현상 역시도 논리적으로 추론이 가능하다.

택시구역의 동질성이란 택시요금의 시계할증이 없는 경우(동질적 택시구역)와 시계할증이 있는 경우(이질적 택시구역)를 구분하기 위한 범주형 변수로서 추정된 파라메타는 출발지와 도착지가 동일 택시사업구역인 경우 택시분담률이 높아지는 현상을 기술하고 있다.

우도비($\rho^2(0)$)는 0.404로 산출되었으며 따라서 비용변수 기반 수단분담모형에 비해서 소폭이긴 하지만 설명력이 높은 것으로 평가된다. 더욱이 대안별 통행비용변수들의 설명력을 평가할 수 있는 새로운 우도비($\rho_3^2(\alpha)$)는 모형의 우도비(0.404)보다 높은 0.452로 산출되었는데 이를 통해서 수단분담구조의 변동 대부분을 존 데이터가 설명하고 있음을 확인할 수 있다.

2. 비용변수 기반 모형과의 비교평가

1) 적합도 평가

OD 기반 10개 목적별 비용변수 기반 수단분담모형과 존 데이터 기반 수단분담모형의 적합도를 비교한 결과는 〈Figure 3〉~〈Figure 5〉로 정리되어 있다. 〈Figure 3〉은 상수항을 포함한 일체의 설명변수가 도입되지 않은 모형의 우도값($L(0)$)과 추정된 모형의 우도값($L(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$)을 비교한 우도비($\rho^2(0)$)로서 비용변수 기반 수단분담모형에 비해서 존 데이터 기반 수단분담모형의 적합도가 다소 높음을 확인할 수 있다.

〈Figure 4〉는 상수항만으로 구축된 모형의 우도값($L(C)$)과 추정된 모형의 우도값($L(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$)을 비교한 상수 우도비($\rho^2(C)$)로서 상수항 이외의 설명변수의 우도비 증진 기여도를 간접적으로 평가할 수 있는 준거가 된다. 전목적에서 존 데이터 기반 수단분담모형의 우도비가 비용변수기반 수단분담모형의 상수 우도비보다 높음을 확인할 수 있으며 그 격차는 〈Figure 3〉에서 나타난 격차보다 크다는 사실을 확인할 수 있다.

〈Figure 5〉는 추정된 모형에서 상수항만을 이용해서 산출한 우도값($L(\hat{\alpha})$)과 모형 전체의 우도값($L(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$)을 비교하고 있으며 10개 전목적에 있어서 존 데이터의 설명력이 대안별 비용변수에 비해서 수단분담구조를 더욱 잘 설명하고 있음을 보여주고 있다.

〈Figure 3〉 Likelihood Ratio($\rho_1^2(0)$) Comparison 〈Figure 4〉 Likelihood Ratio($\rho_2^2(C)$) Comparison 〈Figure 5〉 Alternatives Likelihood Ratio($\rho_3^2(\alpha)$) Comparison

특히, 통근목적과 여가목적의 존 데이터 기반 수단분담모형에서는 수단분담구조의 변동 대부분을 존 데이터가 설명하고 있음을 알 수 있다.

2) 통행비용변수의 신뢰도 평가

존 데이터(존의 사회경제지표, 토지이용변수 그리고 교통체계변수 등)는 수단분담모형에 반드시 포함되어야 할 당위성이 없기 때문에 부호검정이나 통계적 검정을 통과한 설명변수만으로 모형을 구축해도 문제가 발생하지 않는다. 그러나 비용변수 기반 수단분담모형에 있어서는 수단별 비용변수들이 핵심적 역할을 하기 때문에 수단별 비용변수들을 최대한 모형에 포함시키려는 노력이 시도되며 실제로 통계적으로 유의하지 않더라도 부호검정만 통과하면 설명변수로 활용하는 경우가 적지 않다. 〈Table 8〉은 목적별 비용변수 기반 수단분담모형의 추정결과에서 비용변수들에 대한 부호검정 및 통계적

〈Table 8〉 Results of t-test and sign-test of Travel Time and Cost Variables

Variable Description	Auto		Transit		Taxi	
	travel cost	travel time	travel cost	travel time	travel cost	travel time
see-off	○	△	○	△	○	▽
return home	○	○	○	○	○	▽
work	○	○	○	○	○	○
school	○	○	○	○	○	▽
institute	○	○	○	△	○	▽
business	○	▽	○	×	○	▽
back to work	○	▽	○	×	○	▽
shopping	○	▽	○	▽	○	▽
recreation	○	▽	○	△	○	▽
other	○	▽	○	▽	○	▽

Note) ▽ : sign-test failure

△ : t-test failure

× : sign-test and t-test failure

검정 결과를 정리한 것이다.

추정된 60개의 파라메타(목적 10개 × 목적별 효용함수 3개 × 효용함수별 비용변수 2개) 중에서 16개의 파라메타가 부호검정을 통과하지 못하였고 4개의 파라메타는 95%의 신뢰구간에서 t-검정을 통과하지 못하였으며 2개의 파라메타는 부호검정과 t-검정을 동시에 통과하지 못하였다.

비용변수들에 대한 상관분석에서 통행시간간 상관계수가 모두 0.98 이상(〈Table 3〉의 음영처리된 부분)이라는 점을 감안하면 이러한 현상은 오히려 당연한 결과로 보아야 할 것이다. 일반적으로 표본이 증가하면 파라메타의 신뢰도 역시 증가하는데, 설명변수간 높은 상관관계로 인한 다중공선성 문제가 대량의 데이터에도 불구하고 이와 같은 결과를 야기하였다고 해석할 수 있다.

V. 결론 및 향후연구과제

1. 결론

본 연구에서는 비용변수를 핵심적 설명변수로 활용하고 있는 현행 수단분담모형의 이론적·현실적 문제점(설명변수간 높은 상관관계로 인한 다중공선성 문제와 버스, 지하철의 통행비용변수들에 대한 장래치 추정불가능성 문제)을 극복하기 위한 방안으로서 장래치 추정이 가능한 존의 사회경제적 변수, 토지이용변수 그리고 교통체계변수를

설명변수로 하는 '존 데이터 기반 수단분담모형'의 활용 가능성을 검증하였다. 비용변수 기반 수단분담모형의 추정과정에서 설명변수간 높은 상관관계로 인한 다중공선성 문제를 확인할 수 있었으며 모형의 적합도 측면에서 존 데이터 기반 수단분담모형이 비용변수 기반 수단분담모형에 대해서 상대적으로 우월할 수 있다는 점도 확인할 수 있었다.

도시교통체계에서 존 데이터는 대중교통서비스 수준과 높은 상관관계를 갖고 있고 따라서 대중교통서비스, 특히 버스서비스의 장래치 추정불가능성 문제를 회피할 수 있는 대리변수로서 활용할 수 있을 것이라는 추론이 본 연구의 출발점이었다. 가령, 인구밀도가 높은 도시화 지역의 대중교통서비스수준이 인구밀도가 낮은 농어촌지역의 대중교통서비스수준에 비해서 높은 것은 경제학의 수요-공급법칙으로도 충분히 설명이 가능하다.

본 연구에서 제안하고 있는 존 데이터 기반 수단분담모형은 통행비용변수 기반 수단분담모형에 비해서 몇 가지 장점을 갖고 있다. 첫째, 장래교통수요추정에서 장래치 추정이 불가능할 뿐만 아니라 다중공선성 문제를 안고 있는 수단별 비용변수를 필요로 하지 않는다는 점이다. 둘째, 대규모 택지나 신도시와 같은 대규모 개발계획에 따른 교통수요추정에 적용이 가능하다는 점이다. 대규모 개발계획 조서를 통해서 존 데이터(사회경제적 변수, 토지이용변수, 교통체계 변수)를 추출할 수 있기 때문에 버스와 전철간 분할은 불가능하지만 대중교통 통합수단OD의 추정은 가능하다. 셋째, 정책적 필요에 대응해서 장래치 추정이 가능한 일부 수단의 통행비용변수들을 선택적으로 활용할 수 있다는 점이다. 본 연구를 통해서 구축된 존 데이터 기반 수단분담모형에는 수단별 통행비용변수를 일질 설명변수로 포함시키지 않았으나 승용차의 통행시간이나 통행비용, 전철의 통행시간이나 통행비용과 같이 장래치 추정이 가능한 비용변수들을 추가적으로 도입하는 것은 분석가가 자유롭게 선택할 수 있는 사안이다.

마지막으로 본 연구성과물의 활용영역을 명확히 해두고자 한다. 본 연구에서 제안된 존 데이터 기반 수단분담모형은 교통수단별 비용변수를 설명변수로 채용하고 있지 않기 때문에 수단별 비용변수에 변화를 야기하는 교통정책의 효과분석에는 활용할 수 없다는 점이다. 교통수단별 비용변수를 변화시키는 교통정책을 평가하기 위해서는 교통수단별 비용변수를 주요 설명변수로 하는 전통적 수단분담모형을 이용하기를 권장한다.

2. 향후연구과제

본 연구를 통해서 제시된 존 데이터 기반 수단분담모형을 활용해서 장래 목표연도의 수단 OD를 추정하기 위해서는 보완해야 할 사항이 적지 않다.

첫째, 목적 OD를 수단 OD로 전환시키기 위해서는 1차적으로 전체 통행에서 수단분담모형의 적용대상 통행과 비적용대상 통행을 분할하기 위한 방법론을 개발하여야 한다. 가령, 본 연구에서 도보통행(자전거, 오토바이 포함)과 지역간 수단통행(시외·고속버스, 일반철도, KTX 포함)은 모형의 적용대상에서 제외되어 있다. 목적 OD로부터 도보통행과 지역간 통행을 어떻게 합리적으로 추정할 것인가에 대한 추가적 검토가 필요하다.

둘째, 본 연구에서는 전철과 버스를 대중교통수단으로 통합해서 대안집합을 구성하였다. 그러나 버스 OD는 통행배정단계에서 공로상에 배정되지만 전철 OD는 전용의 철도 네트워크에 배정되기 때문에 버스 OD와 전철 OD는 반드시 독립적 수단 OD로 구축되어야 할 필요성이 높다. 대중교통 OD에서 버스 OD와 전철 OD를 식별하는 방안에 대한 추가적 연구가 필요하다.

셋째, 복합교통수단을 이용한 통행, 특히 접근교통수단의 고려방안에 대한 추가적 연구 역시 수단 OD 구축을 위해서는 반드시 수행되어야 할 분야에 해당된다. 본 연구에서는 대표교통수단만을 대안으로 한 수단분담모형을 구축하였기 때문에 접근교통수단은 제외되어 있다. 즉, 환승이 고려되지 못함으로 인해서 실질적으로 총수단통행량의 과소추정이 발생하게 되므로 이러한 문제점을 개선하기 위한 추가적 연구가 필요하다.

REFERENCES

1. 경기도(2004), 장래 교통수요 예측 및 대응방안 연구, 경기도.
2. Kim, Soon-Gwan(1999), Models of Transportation System Analysis and Travel Characteristics for Seoul Metropolitan Area, Seoul Development Institute.
3. 김종형(2005), 인천시 교통수요 추정, 인천발전연구원.
4. 도철웅(2006), 교통공학원론(하), 청문각.
5. RYU, Sikyun(2008), "Effects of Multicollinearity in Logit Model", Journal of Korean Society of Transportation, Vol.26, No.1, Korean Society

- of Transportation, pp.113~126.
6. Park, Yong-Su · Ryu, Si-Kyun · Oh, Young-Taek · Kim, Tae-Ho · Rho, Jeong-Hyun(2010), "Development of Modal Split Models Introducing Zonal Land Use and Zone-pair Transportation Service", Journal of Transport Research, Vol.17, No.4, pp.49~61.
 7. 서울시(2000), 서울시 교통정비 중기계획, 서울시.
 8. 서울시(2004), 서울시 장래교통 수요예측 및 대응방안 연구, 서울시.
 9. 수도권교통본부(2009), 수도권 장래교통 수요예측 및 대응방안 연구, 수도권교통본부.
 10. 원재무(1997), 도시교통론, 박영사.
 11. YUN, Dae-Sic·YUK, TAE SUK·KIM, Sang-Hwang(2006), "A Study on the Intercity Mode Choice Behavior of Daegu Citizens According to the Introduction of Gyeongbu High-Speed Railway", Journal of Korean Society of Transportation, Vol.24, No.1, Korean Society of Transportation, pp.29~38.
 12. Lee, Soon-Joo(2007), The Development of Mode Choice Model by RP Data: Focused on Analysis of the City of Suwon, Master's Thesis, University of Seoul.
 13. Lee, Hye-Seung · Lee, Hee-Yeon(2009), "Differential Changes in Commuter's Mode Choice after the Intergrated Public Transit System in Seoul Metropolitan City", Journal of Korean Geographical Society, Vol.44, No.3, pp.323~338.
 14. 임강원(1997), 도시교통계획 -이론과 모형-, 서울대학교 출판부.
 15. Jun, Myung-Jin · Baek, Seung-Hun(2008), "An Analysis on Determinants of Commuter's Mode Choice in the Seoul Metropolitan Area Using Conditional Logit Models", Journal of Korea Planners Association, Vol.43, No.4, pp.9~19.
 16. Jun, Eun-Ha · Lee, Seong-Woo(2007), "A Study of Mode Choice for Commuters in Seoul Metropolitan Area : 1995-2000", Seoul Studies, Vol.8, No.4, pp.107~125.
 17. Fertal, Martin J. and others(1966), Modal Split: Documentation of Nine Methods for Estimating Transit Usage, Bureau of Public Roads, Washington, D.C., pp.54~68.
 18. Forbord, R.J.(1966), Twin Cities Modal Split Model, Minnesota Highway Department.
 19. K. Moskowitz(1956), California Method of Assigning Diverted Traffic to Proposed Freeways, HRB Bulletin 130.
 20. M.E. Campbell(1951), Assignment of Traffic to Expressways(as determined by diversion studies), Southwestern Association of State Highway Officials : Proceedings.
 21. Southworth F(1978), "A highly disaggregated modal-split model-some tests" Environment and Planning A, Vol.10, No.7, pp.795~812.
 22. 동경도시권교통계획협의회(2001), 평성10년 동경도시권 퍼슨트립조사(종합도시교통체계조사) 보고서 -장래추계편-, 동경도시권교통계획협의회.

✎ 주 작 성 자 : 류시균
 ✎ 교 신 저 자 : 류시균
 ✎ 논문투고일 : 2011. 5. 2
 ✎ 논문심사일 : 2011. 8. 1 (1차)
 2011. 12. 6 (2차)
 ✎ 심사판정일 : 2011. 12. 6
 ✎ 반론접수기한 : 2012. 6. 30
 ✎ 3인 익명 심사필
 ✎ 1인 abstract 교정필