

공간 데이터 재구축을 통한 음식점종 매출액 영향 요인 분석* : 이종 공간 데이터의 집계단위 변환을 중심으로

Analyzing Influence Factors of Foodservice Sales by Rebuilding Spatial Data : Focusing on the Conversion of Aggregation Units of Heterogeneous Spatial Data

노은빈¹⁾ · 이상경²⁾ · 이병길³⁾

Noh, Eunbin · Lee, Sang-Kyeong · Lee, Byoungkil

Abstract

This study analyzes the effect of floating population, locational characteristics and spatial autocorrelation on foodservice sales using big data provided by the Seoul Institute. Although big data provided by public sector is growing recently, research difficulties are occurred due to the difference of aggregation units of data. In this study, the aggregation unit of a dependent variable, sales of foodservice is SKT unit but those of independent variables are various, which are provided as the aggregation unit of Korea National Statistical Office, administration dong unit and point. To overcome this problem, we convert all data to the SKT aggregation unit. The spatial error model, SEM is used for analysing spatial autocorrelation. Floating population, the number of nearby workers, and the area of aggregation unit effect positively on foodservice sales. In addition, the sales of Jung-gu, Yeongdeungpo-gu and Songpa-gu are less than that of Gangnam-gu. This study provides implications for further study by showing the usefulness and limitations of converting aggregation units of heterogeneous spatial data.

Keywords: Foodservice Sales, Aggregation Unit, Spatial Autocorrelation, Floating Population

초 록

이 연구에서는 서울연구원 제공 집계구 자료를 이용하여 유동인구와 공간 특성, 공간자기상관이 음식점종의 매출액에 미치는 영향을 분석한다. 최근 공공 분야의 빅데이터가 증가하고 있지만 자료집계 단위가 달라 연구에 어려움을 야기하는 경우가 많다. 본 연구에서도 종속변수인 매출액은 SKT 집계구 단위로, 공간 특성은 통계청 집계구와 행정동 단위로, 그리고 유동인구는 Point 정보로 구축되어 있다. 분석을 위해 먼저 SKT 집계구 단위로 모든 독립변수의 집계단위를 변환하였다. 공간자기상관 효과를 검정할 수 있는 모형인 SEM을 이용하여 회귀 분석을 수행한 결과 매출액에 대한 공간자기상관의 영향이 확인되었으며, 또한 유동인구와 주변 종사자수, 집계구 면적이 매출액에 정의 영향을 주는 것으로 나타났다. 또한, 강남구에 비해 중구, 영등포구, 송파구 매출액이 더 적은 것으로 나타났다. 이 같은 결과는 유동인구가 많고 주변 종사자수가 많은 지역, 그리고 음식점종 점포들이 몰려있는 지역이 창업에 유리하다는 것을 시사한다. 이 연구는 이종 공간 데이터의 집계단위 변환의 유용성과 함께 지도상의 면적 비례로 집계단위를 변환하는 것의 한계를 보여줌으로써 후속 연구에 시사점을 제공한다.

핵심어 : 음식점종 매출액, 집계구 단위, 공간자기상관, 유동인구

Received 2017. 11. 30, Revised 2017. 12. 12, Accepted 2017. 12. 29

* This paper is based on Noh's Gachon University Master's Thesis.

1) Dept. of Urban Planning, Gachon University (Email: sjgneb05@naver.com)

2) Corresponding Author, Member, Dept. of Urban Planning, Gachon University (Email: skylee@gachon.ac.kr)

3) Member, Dept. of Civil Engineering, Kyonggi University (Email: basil@kgu.ac.kr)

1. 서 론

최근 은퇴자를 중심으로 진입장벽이 낮고 특별한 기술이 없어도 창업이 가능한 음식업종에 대한 창업이 증가하고 있다. 그러나 상권에 대한 이해가 부족한 상태에서 음식업의 특성을 고려하지 않는 창업으로 인해 폐업률이 타 업종에 비해 상당히 높은 실정이다. 폐업률을 낮추기 위해서는 창업자들에게 도움을 줄 수 있는 다양한 정보 제공이 필요한데, 음식업종의 매출액 영향 요인에 대한 정보도 그 중의 하나라고 할 수 있다. 이 같은 정보가 실용적 의미를 갖기 위해서는 대규모 공간 자료를 활용한 연구가 필요한데, 개인정보 보호 등의 이유로 인해 관련 자료에 대한 접근이 쉽지 않은 실정이다. 최근 들어 빅데이터 활용의 중요성이 강조되면서 공공분야를 중심으로 이 같은 상황이 개선될 여지가 보이고 있는데, 이와 관련하여 서울시와 서울연구원에서는 소매업종 점포 매출액 자료를 집계구 단위로 집계하여 연구자들에게 제공하고 있다. 이를 이용할 경우 음식업종 매출액의 공간분포도 분석할 수 있으며, 기존의 공간 자료와 결합할 경우 매출액에 영향을 주는 요인도 분석할 수 있을 것으로 보인다.

음식업종 매출액에 영향을 주는 요인으로는 먼저 유동인구를 고려해 볼 수 있다. 유동인구는 주변 상업시설들의 매출액 및 공간 생산성에 영향을 미치기 때문에 유동인구가 많은 곳에 점포가 입지하는 경우가 많다. 하지만 유동인구가 음식업종 매출액에 미치는 영향을 다룬 연구는 거의 찾아보기 어려운 실정이다. 이와 달리, 유동인구가 여타 소매업종의 매출액에 미치는 영향을 분석한 연구는 몇 편 발견되고 있는데, 상업시설(Choi and Shin, 2001), 편의점(Lee *et al.*, 2010), 커피전문점(Shin and Moon, 2011), 소매업(Kim *et al.*, 2015), 상가(Yu *et al.*, 2016) 등을 다룬 연구들이 있다. 또한 유동인구가 소매업 입지와 밀도에 미치는 영향을 분석한 연구도 있는데 Shin *et al.*(2001), Kim and Jung(2013), Lee and Jung(2014), Kim and Ko(2015) 등이 있다. 유동인구 외에도 점포가 위치하고 있는 주변의 환경도 매출액에 많은 영향을 준다. 가령, 주변에 사업장들이 많아 종사자들이 많거나 음식업종 점포들이 밀집할 경우 매출액에 긍정적 영향을 줄 수 있다. 또한 Jung *et al.*(2015)의 의류업종 연구에서 나타났듯이 공간자기상관(spatial autocorrelation)으로 대변되는 매출액이 큰 점포들이 군집하여 입지하는 경향이 확인될 수도 있다. 이 같은 현상을 분석하기 위해서는 대규모 공간 데이터베이스를 구축하는 것이 필요한데, 음식업종의 매출액 자료와 개별 공간 자료의 구축 단위가 서로 다르다는 점이 데이터베이스 구축에 어려움을 야기할 수 있다. 가령, 서울연구원 제공 매출액 자료는

SKT 집계구 단위로 구축되어 있는 반면 상권 공간특성 자료의 상당수는 통계청 집계구와 행정동 단위로 구축되어 있다. 이 외에도 유동인구 자료는 매시 형태의 Point 정보로 구축되어 있다. 이 같이 구축단위가 서로 다른 자료를 변환하여 하나로 통합할 경우 원자료의 왜곡이 나타날 수 있다는 점에서 주의가 필요하다.

본 연구에서는 유동인구와 공간 특성, 그리고 공간자기상관이 음식업종 매출액에 미치는 영향을 분석하고자 한다. 분석방법으로는 공간자기상관을 고려할 수 있는 공간회귀모형(spatial regression model)을 이용하고자 한다. 한편, 데이터베이스 구축과정에서 Point 정보로 구축된 유동인구 자료와 통계청 집계구 자료를 SKT 집계구 자료로 변환하게 되는데, 이를 통해 이중 데이터의 집계단위 변환을 통한 공간통계분석의 가능성과 함께 그 한계를 살펴보고자 한다.

2. 분석 모형

본 연구의 분석대상이 되는 음식업종 매출액은 집계구 자료이기 때문에 공간계량경제학(spatial econometrics)에서 말하는 공간자기상관을 내포할 가능성이 매우 높다. 공간자기상관은 일반적으로 양의 자기상관을 의미하는 데, 이는 인접할수록 유사한 특성을 지니게 되고 상관관계가 높아지게 되는 것을 의미한다(Lim and Lee, 2013). 공간자기상관의 존재는 OLS(ordinary least square) 회귀모형의 추정치를 최소분산 추정치(minimum variance)로 확신할 수 없게 한다는 점에서 대안을 필요로 한다. 공간자기상관이 관측될 경우 일반적으로 공간회귀모형을 분석에 이용한다. 공간회귀모형으로는 공간종속변수(spatial dependent variable)를 OLS 모형에 포함하는 SLM(spatial lag model)과 공간오차(spatial error)를 삽입하는 SEM(spatial error model)이 많이 사용된다(Kim *et al.*, 2010).

OLS 회귀모형은 종속변수로 사용되는 음식업종 집계구 매출액의 형태에 따라 다양한 모형이 가능하지만 본 연구에서는 로그변환을 한 매출액을 종속변수로 하는 선형모형을 이용하고자 하며 이는 Eq.(1)과 같이 정리된다.

$$Y = X\beta + \epsilon \quad (1)$$

where Y is the log-transformed sales of aggregation unit, X is the vector of floating population and locational characteristics, β is a parameter, and ϵ is an error term.

공간종속변수를 OLS모형에 삽입한 SLM은 Eq.(2)과 같다.

$$P = \rho WP + X\beta + \epsilon \tag{2}$$

$$\epsilon \sim N(0, \sigma^2 I_n)$$

where W is a spatial weights matrix, and ρ is a parameter for a space dependent variable.

SEM은 공간오차(spatial error)를 OLS모형에 포함하는 것으로 Eq.(3)과 같이 표현된다.

$$P = X\beta + u, \quad u = \lambda Wu + \epsilon \tag{3}$$

$$\epsilon \sim N(0, \sigma^2 I_n)$$

where λ is a parameter for spatial error.

OLS 회귀모형 대신 SLM이나 SEM을 이용하는 것이 타당하기 위해서는 OLS 추정 잔차에서 공간자기상관이 검정되어야 한다. 일반적으로 LM Lag, LM Error 검정을 통해 공간자기상관의 존재 유무를 검정하며 그 결과를 이용하여 SLM과 SEM 중 더 우수한 모형을 선택하게 된다. LM Lag 통계치만 유의하다면 SLM을, LM Error 통계치만 유의할 경우에는 SEM을 선택한다. 둘 다 유의할 경우에는 Robust LM Lag과 Robust LM Error 검정을 통해 더 우수한 모형을 선택하게 된다(Lim and Lee, 2013).

3. 분석 자료

3.1 매출액 집계구 자료

1) 집계구 특성

본 연구에서는 서울연구원에서 제공한 SKT Geovision의 음식점중 매출액 추정 자료를 분석에 이용한다. SKT Geovision에서는 매출액 정보를 점포 또는 필지 단위가 아닌 집계구 단위로 제공하고 있기 때문에 본 연구에서도 불가피하게 집계구 단위로 자료를 구축할 수밖에 없는 한계를 가지게

되었다. Fig. 1은 통계청 집계구와 SKT 집계구의 차이를 보여주고 있는데, 통계청 집계구는 야간 상주인구 1,000명을 기준으로 설정하기 때문에 상업, 업무지역의 경우 야간 상주인구가 주거지역에 비해 적어 집계구의 크기가 커지는 반면 SKT Geovision의 매출액 집계구는 통계청 집계구의 상업지역을 업소, 사무실 밀집지역으로 세분화하기 때문에 집계구 크기가 더 작아진다는 것을 알 수 있다.

Table 1에서 통계청 집계구와 SKT Geovision 집계구의 크기를 비교해 보면, 통계청 집계구의 평균이 48,194m²로 SKT 집계구 평균 29,503m²보다 약 1.6배 정도로 더 크며, 최소값, 중간값, 최대값, 표준편차도 통계청 집계구가 SKT 집계구보다 더 큰 것으로 나타났다. Table 2는 SKT 집계구 자료의 특성을 좀 더 파악하기 위해 면적 크기순으로 집계구를 배열한 후 각 분위별 면적을 제시한 것이다. 집계구의 모양이 불규칙하기 때문에 면적 정보가 제시되었음에도 불구하고 크기를 가늠하는 것이 어려워 표준화 작업의 일환으로 집계구 면적과 동일한 가상의 원을 가정한 후 그 반경을 분석하였다. 그 결과를 보면, 5%에 해당하는 반경은 28.1m, 50% 중간값의 반경은 59.8m, 95%는 144.9m로 나타났다. 95% 이상에서는 원의 크기가 급격하게 커지는 것으로 나타났는데 최대 반경은 1,025.9m로 95%에 비해 7.1배 정도로 큰 것으로 나타났다.

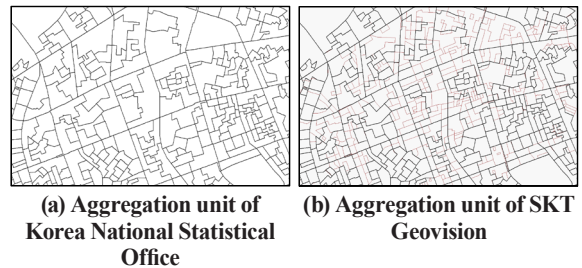


Fig. 1. Comparison of the aggregation unit of Korea National Statistical Office with SKT Geovision

Table 1. Basic statistics of aggregation units in 6 gus* of Seoul

| Aggregation unit | N (count) | Mean (m ²) | Std. Deviation (m ²) | Minimum (m ²) | Median (m ²) | Maximum (m ²) |
|-----------------------------------|-----------|------------------------|----------------------------------|---------------------------|--------------------------|---------------------------|
| Korea National Statistical Office | 3,569 | 48,194 | 195,503 | 1,065 | 15,475 | 5,613,087 |
| SKT Geovision | 5,974 | 29,503 | 125,202 | 96 | 11,234 | 3,306,741 |

*: 6 gus are Jongro-gu, Jung-gu, Yeongdeongpo-gu, Gangnam-gu, Seocho-gu, and Songpa-gu

Table 2. Percentile of the area of SKT aggregation unit

| Percentile | Area of aggregation unit(m ²) | Radius(m)* |
|------------|---|------------|
| 5 | 2,480 | 28.10 |
| 10 | 3,678 | 34.22 |
| 15 | 4,571 | 38.15 |
| 20 | 5,460 | 41.69 |
| 25 | 6,279 | 44.71 |
| 30 | 7,162 | 47.74 |
| 35 | 8,102 | 50.78 |
| 40 | 9,070 | 53.73 |
| 45 | 10,029 | 56.50 |
| 50 | 11,234 | 59.80 |
| 55 | 12,497 | 63.07 |
| 60 | 13,907 | 66.53 |
| 65 | 15,473 | 70.18 |
| 70 | 17,299 | 74.20 |
| 75 | 19,831 | 79.45 |
| 80 | 23,065 | 85.69 |
| 85 | 27,770 | 94.02 |
| 90 | 37,098 | 108.67 |
| 95 | 65,958 | 144.90 |
| 100 | 3,306,741 | 1,025.95 |

* This radius assumes that aggregation unit is a circle

2) 집계구 매출액 특성

본 연구에서는 SKT 집계구 음식점중 매출액을 종속변수로 설정하며, 분석에 이용하는 자료는 2013년 4월 월간 매출액이다. 매출액에 대한 기초통계분석 결과, 평균은 250,758천원인데 표준편차는 514,721천원으로 평균의 2배가 넘는 것으로 나타나 집계구의 매출액 차이가 큰 것으로 나타났다.

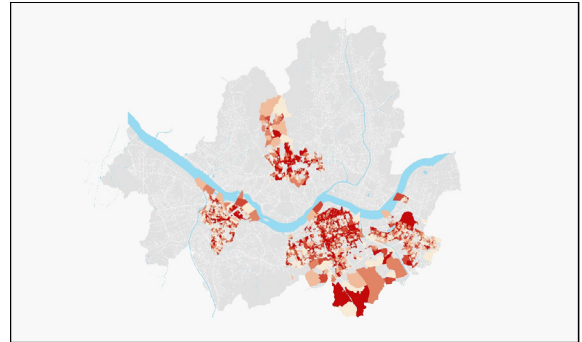


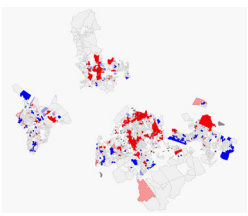
Fig. 2. Distribution of foodservice sales by SKT aggregation unit

음식업종 집계구 매출액의 공간자기상관을 분석하기 위해 Moran's I를 산정한 결과 0.3146으로 비교적 강한 양의 자기상관성이 나타났다. 이는 음식 업종에서 유사한 수준의 매출액을 보이는 지역들이 공간적으로 서로 인접해 있다는 것을 의미한다. Moran's I 값이 전역적 공간자기상관을 보여주는 지표라면 Anselin(1995)이 제안한 LISA(local indicator of spatial association)는 국지적인 차원에서 공간자기상관을 보여준다. LISA는 특정 지역과 인접한 지역들이 갖는 값의 가중 평균값으로 산출되며 가중 평균값이 서로 유사하게 나타나면 정적인 자기상관으로, 차이가 크게 나타나면 부적의 자기상관으로 나타난다. 음식 업종의 LISA 분석 결과, 매출액이 높은 집계구 주변에 매출액이 높은 HH유형(high-high)이 468개, 매출액이 낮은 집계구 주변으로 매출액이 낮은 LL유형(low-low)이 170개, 매출액이 낮은 집계구 주변으로 매출액이 높은 LH유형(low-high)이 63개, 매출액이 높은 집계구 주변으로 매출액이 낮은 HL유형(high-low)이 40개로 분석되었다. 매출액이 높은 지역들이 이웃하고 있는 HH유형과 매출액이 낮은 지역들이 이웃하고 있는 LL유형의 분포를 통해 음식업종의 매출액이 국지적인 차원에서 공간적으로 군집을 이루고 있음을 알 수 있다.

Table 3. Basic statistics of foodservice sales by SKT aggregation unit

| N (count) | Mean (1000 KRW) | Std. Deviation (1000 KRW) | Minimum (1000 KRW) | Maximum (1000 KRW) |
|-----------|-----------------|---------------------------|--------------------|--------------------|
| 2,962 | 250,758 | 514,721 | 77 | 11,522,861 |

Table 4. LISA output

| LISA Index | Count | LISA Image Output |
|--------------------|-------|---|
| Not significant | 2,191 |  |
| HH(Red) | 468 | |
| LL(Blue) | 170 | |
| LH(Pink) | 63 | |
| HL(Purple) | 40 | |
| Neighborless(Grey) | 30 | |

Note: Moran's I is 0.3146 and significant at the 0.05 significance level

LISA 지표를 살펴보면 종로구와 중구의 업무밀집지역, 영등포구 여의도동 일대, 강남대로와 압구정동 일대, 삼성역 인근, 송파구의 올림픽공원 인근의 음식점 매출액이 HH유형을 나타내고 있다. 이는 지역의 성격이 업무밀집지역이거나 상업지역의 성격을 가질 때, 음식점은 밀집할수록 매출액이 증가한다고 볼 수 있다. 반면에 비교적 주거지역의 성격이 강한 지역에 위치할수록 음식점은 밀집할수록 매출액이 낮아진다고 해석할 수 있다.

3.2 유동인구 자료

1) SKT Geovision 유동인구 추정자료

본 연구에서는 서울연구원이 제공한 서울시 6개구(종로구, 중구, 영등포구, 강남구, 서초구, 송파구)의 2013년 4월 평균 유동인구수를 분석 대상으로 한다. 이 자료는 SKT Geovision이 제공한 것으로 SKT 통신횟수(통화, 문자 등) 자료를 50m*50m 단위로 집계한 것으로 Table 5에서는 구별로 유동인구를 정리하였다. 구별로 보면, 강남구, 송파구, 서초구 순으로 나타나고 있어 강남 3구에 많은 유동인구가 발생하고 있다는 것을 알 수 있다. Fig. 3에서 유동인구 분포를 이미지

로 보면, 종로구와 중구의 업무밀집지역, 영등포구의 여의도동, 강남구와 서초구의 대로변 일대에 유동인구가 많은 것으로 나타났다. 이는 SKT의 유동인구 추정자료가 단순 보행량만이 아닌 실제 통화횟수를 기반으로 유동인구를 추정한 자료로서 차량을 이용한 유동인구를 포함하기 때문에 풀이된다. 차량 효과는 올림픽대로가 지나는 한강변에도 유동인구가 많은 것에서 그 근거를 찾을 수 있다.

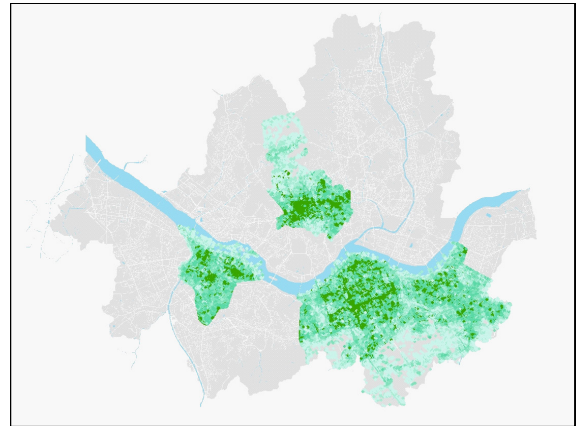


Fig. 3. Distribution of floating population of SKT

2) 1차 가공

본 연구에 사용된 유동인구 추정자료는 해당 셀에서 해당 시간대(한 시간 단위)에 통신한 휴대전화 수로서 동일 시간대에 대한 중복은 유동인구의 측정에서 제외하였다. SKT의 유동인구 자료는 기지국을 이용하여 통신횟수를 측정된 뒤 자체 알고리즘을 이용하여 추정한 자료로서 기지국의 측정 오차를 포함하고 있을 가능성이 높다. 방송통신위원회는 기지국을 활용한 위치정보의 오차범위를 150m ~ 수km로 발표한 바 있다(Korea Communications Commission, 2015).

Table 5. Basic statistics of floating population of SKT

| Gu | N(points) | Mean | Std. Dev. | Minimum | Maximum |
|-----------------|-----------|-------|-----------|---------|----------|
| Jongro-gu | 5,802 | 296.9 | 579.0 | 0.1 | 21,156.3 |
| Jung-gu | 3,753 | 598.5 | 765.1 | 1.6 | 8,507.6 |
| Yeongdeongpo-gu | 7,969 | 338.9 | 496.0 | 0.2 | 22,663.2 |
| Gangnam-gu | 11,785 | 423.9 | 657.0 | 0.2 | 30,025.7 |
| Seocho-gu | 10,181 | 317.2 | 613.2 | 0.5 | 20,499.9 |
| Songpa-gu | 10,787 | 278.6 | 338.7 | 0.0 | 7,019.8 |

SKT의 유동인구 추정자료는 차량을 이용한 유동인구를 포함하는 데, 집계구 매출액에 영향을 주는 유동인구는 차량을 이용하는 유동인구가 아닌 도보로 이동하는 유동인구일 가능성이 높다. Table 2에서 SKT 매출액 집계구의 면적 분포를 보면 집계구 면적의 95% 이하가 6.6만m² 이하로 나타나고 있는데, 집계구를 원으로 가정할 경우 반경 145m 이하에 해당한다. 이 같은 결과와 Korea Communications Commission(2015)에 근거하여 본 연구에서는 기지국의 측정 오차를 최소화하면서 집계구에 들어와 있을 가능성이 높은 반경 150m 이내 유동인구를 해당 집계구의 유동인구로 가정한다.

유동인구 자료의 재구축은 Arcmap 10.0을 이용하며, 먼저 Buffering 기능을 이용하여 모든 Point로부터 150m 반경을 설정한 후 Spatial Join 기능을 사용하여 150m 반경 내에 들어오는 Point의 값을 기준 Point에 더하였다. 이 과정에서 150m 일 경우 일부 Point가 포함되지 않을 수 있는 상황이 발생하여 실제로는 155m로 설정하여 누락되는 Point가 없게 하였다. 그 결과 29개의 Point의 유동인구가 합산되며 Table 6은 그 합산한 유동인구의 기초 통계량을 정리한 것이다. Fig. 4는 155m 보다 작은 55m Buffering을 통해 Point 유동인구를 합산하는 과정을 보여주고 있다. 155m Buffering도 개념적으로 동일하며 이를 확장한 것이라고 할 수 있다.

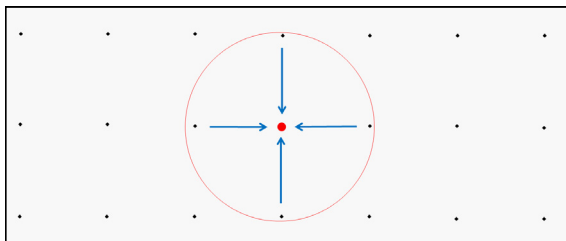


Fig. 4. Floating population of step 1 processing data

Table 6. Floating population of point made by step 1 processing

| Radius (m) | Points (count) | Mean (persons) | Std. Dev. (persons) | Minimum (persons) | Median (persons) | Maximum (persons) |
|------------|----------------|----------------|---------------------|-------------------|------------------|-------------------|
| 155 | 29 | 13,053 | 10,665 | 0 | 10,460 | 106,284 |

3) 2차 가공

본 연구에서는 1차 가공된 데이터를 GIS Spatial Join 기능을 통해 해당 매출액 집계구에 포함된 유동인구를 분석하여 그 평균을 이용한다. 2차 가공 집계구 유동인구는 Fig. 5와 같이 집계구에 포함되는 유동인구 포인트 값(1차 가공 유동인구)의 합을 유동인구 포인트 개수로 나눈 값이 된다.

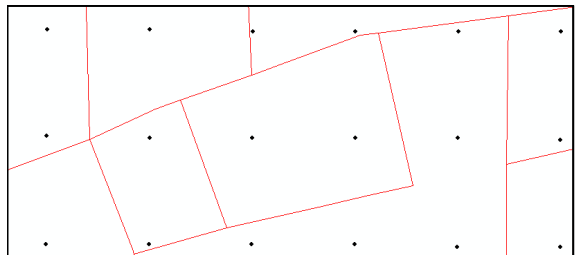


Fig. 5. Floating population of step 2 processing data

1차 가공을 통해 해당 Point를 기준으로 150m 반경을 설정해 해당 반경 안에 들어오는 값을 합하고, 2차 가공을 통해 해당 집계구에 들어오는 유동인구를 평균값으로 적용함으로써 하나의 집계구 안에 들어오는 유동인구만이 아닌 집계구 주변의 유동인구의 값도 포함하게 된다. 결과적으로 집계구의 크기와 무관하게 집계구별로 29개의 Point 유동인구가 합산된 결과가 산출된다. Table 7에서 보이듯이 집계구별 평균은 1차 집계결과와 비슷한 수준의 평균을 가지게 된다.

Table 7. Floating population of aggregation unit made by step 2 processing

| Mean | Std. Dev. | Minimum | Maximum |
|--------|-----------|---------|---------|
| 15,385 | 12,136 | 204 | 106,283 |

Table 8. Specification for variables

| Variables | Specification | Year | Source |
|-------------------------|---|------|---------------------|
| Ln(sales) | Monthly sales of SKT aggregation unit | 2013 | SKT geovision |
| Ln(floating population) | Rebuilding floating population data made by SKT | 2013 | SKT geovision |
| Ln(workers) | Number of workers around SKT aggregation unit | 2010 | The Seoul Institute |
| Ln(area) | Area of SKT aggregation unit | 2013 | SKT geovision |
| Ratio of foodservice | Area of foodservice business divided by dong area | 2013 | The Seoul Institute |
| Gu dummy variables | Jongro-gu, Jung-gu, Yeongdeongpo-gu, Seocho-gu, Songpa-gu (reference group: Gangnam-gu) | 2013 | SKT geovision |
| Number of buses | Number of bus stop in SKT aggregation unit | 2012 | Biz-Gis |
| Subway station distance | Nearest distance from subway station | 2012 | Biz-Gis |

3.3 독립변수 공간자료

Table 8에서 볼 수 있듯이, 본 연구에서는 음식점업 집계구 매출액 영향 요인으로 유동인구와 함께 공간특성 요인을 설정한다. 공간특성 요인으로는 집계구 주변 종사자수, 집계구 규모, 음식점업의 비율, 정차 버스수, 지하철역 거리를 설정한다. 집계구 주변 종사자수 변수(workers)는 통계청 집계구 단위 종사자수를 SKT 집계구의 면적 비율로 재구성한 값으로 배후지역의 종사자수가 아닌 집계구의 종사자수를 의미한다. 이를 수식으로 나타내면 통계청 집계구 단위 종사자수×(SKT 매출액 집계구면적/통계청 집계구 면적)로 표현될 수 있다. 집계구의 종사자 수가 많다는 것은 해당 집계구에 일정시간 머물면서 해당 지역의 소비지출에 영향을 주는 인구가 많다는 것으로 해석할 수 있다. 곧 유동인구와는 상반되는 개념으로 집계구의 상업시설을 주기적으로 이용하는 인구를 의미한다. 종사자수 변수의 재구성은 GIS Spatial Join 기능을 이용한다.

집계구 규모 변수(area)는 집계구 폴리곤의 면적을 이용한다. 집계구 면적이 클수록 많은 점포가 있을 수 있다는 점에서 매출액에 정의 영향을 미칠 것으로 예상된다.

음식업종의 비율 변수(ratio of foodservice)는 행정동을 기준으로 한 건물용도별 전용면적 비율을 의미한다. 본 연구는 음식점업의 상권 규모를 대변하기 위해 행정동 전체 건물의 전용면적 대비 음식점 전용면적의 비율을 사용한다.

지역상권의 발달정도에 따른 지역별 차이를 분석하기 위해 서울시 6개구(종로구, 중구, 강남구, 영등포구, 서초구, 송파구)를 강남구를 참조집단으로 하는 지역터미널수 형태로 투입한다. 교통접근성이 매출액에 미치는 영향을 분석하기 위해 정차 버스수(number of buses)와 지하철역까지의 거리(subway station distance)를 분석에 이용한다. 정차 버스수는

집계구에 속한 버스정류장에 정차하는 버스의 수를 산정하며, 정차 버스수가 많을수록 집계구는 버스 접근성이 높다는 것을 의미한다. 지하철역까지의 거리는 GIS near tool을 이용하여 각 집계구에서 최단 거리에 위치한 지하철역까지의 거리를 측정하였다. 지하철역까지의 거리 측정은 집계구와 지하철역 입구의 최근접 거리로 측정하며, 지하철역 입구가 집계구에 포함되어 있을 경우 집계구와 지하철역까지의 거리는 0으로 하였다. 회귀분석에는 100m 단위로 단위를 변환하여 투입한다.

독립변수가 매출에 미치는 영향을 탄력성으로 파악하기 위해 연속변수의 경우 양변에 로그를 취한다. 종속변수인 매출액과 독립변수인 유동인구, 주변 종사자수에 각각 로그를 취함으로써 음식점업 매출액의 유동인구 탄력성과 주변 종사자수 탄력성을 파악할 수 있다. Table 9는 이들 독립변수들의 기초 통계량을 정리한 것이다.

4. 분석 결과

Table 4에서 볼 수 있듯이, 종속변수로 사용되는 매출액 변수의 Moran's I는 0.3146으로 이는 매출액에서 공간자기상관이 존재한다는 것을 의미한다. 이 같은 결과는 공간자기상관을 고려하지 않는 OLS 회귀모형이 매출액을 분석하는 데에 적합하지 않다는 것을 의미한다. 대안으로 공간회귀모형을 사용하기 위해서는 OLS 회귀 잔차에 대한 통계적 진단을 수행하여야 한다. Table 10을 보면, OLS 회귀잔차의 비정규성을 진단하는 Jarque-Bera 통계량이 483.31, 이분산성을 진단하는 Breusch-Pagan 통계량이 94.73으로 모두 1% 수준에서 유의하게 나타나고 있다. 이는 잔차 분포의 비정규성과 이분

Table 9. Basic statistics of variables

| Variables | Mean | Std. Dev. | Minimum | Maximum |
|-------------------------|-------|-----------|---------|---------|
| Ln(sales) | 18.13 | 1.82 | 11.26 | 23.17 |
| Ln(floating population) | 9.40 | 0.71 | 5.32 | 11.57 |
| Ln(workers) | 5.56 | 1.27 | 1.33 | 9.28 |
| Ln(area) | 9.60 | 0.99 | 6.09 | 15.01 |
| Ratio of foodservice | 0.01 | 0.01 | 0 | 0.05 |
| Jongro-gu dummy | 0.09 | 0.28 | 0 | 1 |
| Jung-gu dummy | 0.06 | 0.25 | 0 | 1 |
| Yeongdeongpo-gu dummy | 0.17 | 0.38 | 0 | 1 |
| Seocho-gu dummy | 0.20 | 0.40 | 0 | 1 |
| Songpa-gu dummy | 0.19 | 0.39 | 0 | 1 |
| Number of buses | 1.50 | 4.86 | 0 | 117 |
| Subway station distance | 3.80 | 4.00 | 0 | 47.35 |

산성을 보여주는 것으로 회귀분석의 기본 가정인 잔차 분포의 정규성과 등분산성에 위배된다. 결과적으로 OLS 회귀모형 대신 공간회귀모형이 필요하다는 것을 시사한다.

공간회귀모형을 대표하는 SEM과 SLM 중에서 더 적합한 모형을 식별하기 위해서는 LM 검정을 수행해야 한다. LM-lag과 LM-error의 통계치가 모두 1% 유의수준에서 유의하여 선택을 어렵게 하고 있지만 Robust LM Lag 통계치는 10%, Robust LM Error 통계치는 1% 유의수준에서 유의하여 SEM이 SLM보다 더 적합하다는 것을 보여주고 있다. 이에 본 연구에서는 SEM을 최종모형으로 하여 회귀분석을 수행한다. 분석 결과, SEM의 Loglikelihood가 OLS보다 더 크며 AIC(Akaike Info Criterion)과 SC(Schwarz Criterion)은 더 작게 나타났다. 이 결과는 SEM이 음식점종의 집계구 매출액 분석에서는 OLS 회귀모형보다 적합도가 높다는 것을 의미한다.

SEM 적용 결과를 보면, 유동인구(floating population), 종사자수(workers), 집계구 면적(area), 지역더미변수(중구, 영등포구, 송파구)가 유의한 것으로 나타났다. 반면, 정차 버스 수(number of buses)와 지하철역까지의 거리(subway station distance)는 유의하지 않은 것으로 나타났다. 계량경제학에서는 종속변수와 독립변수가 모두 로그를 취할 경우 변수의 계수는 탄력성으로 해석한다. 이를 유의한 변수들에 적용하면 다음과 같다. 유동인구가 1% 증가할 때 음식점업 종의 매출은 0.76% 증가하는데, 이는 음식점업 매출이 유동인구에 민감하게 반응한다는 것을 의미한다. 종사자수는 전 사업체의 총 종사자수를 집계구 단위로 재구성한 변수로 종사자수가

증가한다는 것은 해당 집계구에 일정시간 머물면서 해당 지역의 소비지출에 영향을 주는 인구가 증가한다는 것을 의미한다. 음식점종의 매출액은 지역의 종사자수가 1% 증가할 때, 0.55% 증가하는 것으로 나타났다. 유동인구와 종사자수는 그 특성이 다른데, 유동인구가 일회성 소비자라면 종사자수는 주기적 성격을 가지는 소비자로 해석할 수 있다. 음식점종의 매출액은 유동인구와 종사자수에 모두 정(+)의 영향을 받지만 유동인구의 탄력성이 종사자수보다 커 유동인구에 더욱 민감하다는 것을 알 수 있다.

집계구 면적은 정의 영향을 주는 것으로 나타났는데, 집계구 면적이 음식점종이 위치한 해당 사업지의 크기를 의미하진 않지만 집계구 면적이 클수록 더 많은 음식점종이 있을 가능성이 있고 그 결과로 매출액에 정의 영향을 준 것으로 추론된다. 음식점종의 매출액은 집계구 면적이 1% 증가할 때 0.26% 증가하는 것으로 해석된다. 집계구가 속한 자치구를 의미하는 구(Gu) 변수는 더미변수이고 추정 계수들이 모두 음의 값을 보이고 있기 때문에 계량경제학에 따르면 계수값 β 는 해당 구의 음식점종 매출액이 강남구와 비교하여 $(e^{\beta} - 1) \times 100\%$ 더 낮은 것으로 해석된다. 음식점종 집계구 매출액은 강남구에 비해 중구가 55%, 영등포구가 37%, 송파구가 21% 더 낮은 것으로 나타났다. 하지만 종로구와 서초구는 통계적으로 유의한 차이를 보여주지 않았다.

공간효과(spatial effect)를 의미하는 λ 가 1% 수준에서 유의한 것으로 나타나 잔차들에서 공간자기상관이 존재하는 것으로 확인되었다. 공간승수효과는 $\frac{1}{1-\lambda}$ 로 산정되는 데, λ 값

Table 10. Results of regression

| | | OLS | SEM | |
|-------------------------|-------------------------|-----------------|-----------|----------|
| Spatial effect | λ | | 0.30*** | |
| Variables | Constant | 6.60*** | 5.86*** | |
| | Ln(floating population) | 0.69*** | 0.76*** | |
| | Ln(workers) | 0.55*** | 0.50*** | |
| | Ln(area) | 0.22*** | 0.26*** | |
| | Ratio of foodservice | 2.57 | 1.34 | |
| | Gu | Jongro-gu | 0.0012 | -0.017 |
| | | Jung-gu | -0.77*** | -0.80*** |
| | | Yeongdeongpo-gu | -0.44*** | -0.46*** |
| | | Secho-gu | -0.19** | -0.16 |
| | | Songpa-gu | -0.22** | -0.24* |
| | Number of buses | 0.0026 | 0.0040 | |
| Subway station distance | 0.0030 | -0.0029 | | |
| (Pseudo) R ² | | 0.33 | 0.38 | |
| Goodness of Fit | Log likelihood | -5371.74 | -5291.07 | |
| | AIC | 10767.5 | 10606.10 | |
| | SC | 10839.4 | 10678.10 | |
| Normality(Jarque-Bera) | | 483.31*** | | |
| Homoscedasticity | Breusch-Pagan | 94.73*** | | |
| | Kosenker-Bassett | 60.88*** | | |
| Likelihood ratio | | | 161.34*** | |
| LM-Lag | | 68.25*** | | |
| Robust LM-Lag | | 3.50* | | |
| LM-Error | | 178.58*** | | |
| Robust LM-Error | | 113.83*** | | |

Note: ***, ** and * are significant at 0.01, 0.05, 0.1 significance levels, respectively.

이 0.30이기 때문에 공간승수효과는 1.429가 된다. 이는 잔차 값이 공간자기상관에 따른 공간승수효과로 인해 43% 정도 증가한다는 것을 의미한다.

5. 결론

이 연구는 서울연구원이 제공한 음식점종의 집계구 매출액 자료를 이용하여 음식점종 매출액에 영향을 주는 요인들을 분석하였다. 분석모형으로 공간회귀모형의 일종인 SEM을 이용하였으며, 자료 구축은 SKT 매출액 집계구를 기준으로 통계청 집계구와 행정동 단위로 구축된 공간특성 자료, 그리고 Point로 구축된 유동인구 자료를 변환하였다.

음식업종 집계구 매출액 분석 결과, 공간자기상관을 고려한 SEM이 OLS 회귀모형보다 적합도가 더 좋은 것으로 나타

났으며 이를 통해 공간자기상관이 집계구 매출액에 정의 영향을 준다는 사실을 확인할 수 있다. 이는 매출액이 클수록 군집하는 경향이 있다는 것을 의미한다. 독립변수 중 유의한 변수를 보면 유동인구가 많을수록, 주변 종사자수가 많을수록, 집계구의 면적이 클수록 음식점종의 매출액이 커지는 것으로 나타났다. 탄력성을 비교한 결과, 유동인구의 탄력성이 주변 종사자수나 집계구 면적에 비해 상당히 높게 나타나 입지 결정 시 유동인구에 대한 가중치를 높게 할 필요가 있는 것으로 나타났다. 강남구를 기준으로 종로구, 중구, 영등포구, 송파구의 집계구 매출액을 비교한 결과, 종로구와 서초구를 제외한 나머지 3개구가 강남구보다 매출액이 적은 것으로 나타났다. 이 같은 결과는 매출액을 기대한다면 강남구에서 유동인구가 많으며 주변 종사자가 많고 음식점종 관련 점포들이 밀집해 있는 장소를 선택하는 것이 유리하다는 것을 의미한다.

분석 과정에서 확인된 이중 자료의 집계단위 변화에 대한 시사점을 정리하면 다음과 같다. 첫째, 이중 자료의 집계단위 변환에 Spatial Join 기능을 활용할 때 자료의 성격에 따라 자료에 대한 전처리가 필요한 것으로 나타났다. 본 연구의 유동 인구와 같이 주변지역의 영향을 받거나 위치정확도가 충분치 않은 데이터의 집계단위의 변환에는 주변 Point 정보를 활용한 자료 전처리가 필요하다. 둘째, 변수들 중 음식점종 비율변수의 경우 가설의 성립 가능성에도 불구하고 행정동 단위 자료를 변환한 변수가 유의하지 않게 나타났는데, 이는 통계청 집계구 기준 정보를 SKT 집계구로 변환하는 과정에서 단순히 지도상 면적 비례에 따라 집계단위를 변환한 것이 영향을 미쳤을 가능성이 있다.

본 연구는 데이터 제공의 한계로 해당 업종의 점포 기준이 아닌 점포가 입지한 집계구 단위로 분석이 이루어졌다. 따라서 업종별로 점포 자체의 특성이 고려되지 못했다는 점과 개별 점포의 분석결과가 아닌 집계구 단위의 매출액 합계가 분석에 사용되었다는 점에서 한계가 있다. 향후 연구과정에서 유동인구의 특성을 구분하여 성별, 연령별 정보에 대한 분석을 추가한다면 음식점종 창업자들에게 좀 더 유용하게 활용될 수 있을 것으로 기대된다.

References

- Anselin, L. (1995), Local indicators of spatial association-LISA, *Geographical Analysis*, Vol. 27, pp. 93-115.
- Choi, M. and Shin, S. (2001), An empirical analysis of the effect of pedestrian volume on retail sales, *Journal of Korea Planning Association*, Vol. 36, No. 2, pp. 75-83. (in Korean with English abstract)
- Jung, E., Sung, H., and Rho, J. (2015), Analysis on influence factors for apparel retail sales considering spatial autocorrelation, *Journal of Korea Planning Association*, Vol. 50, No. 5, pp. 215-231. (in Korean with English abstract)
- Kim, C. and Jeong, S. (2013), A study on the factors affecting hotel & restaurant location in Seoul, *Korean Management Consulting Review*, Vol. 13, No. 4, pp. 297-314. (in Korean with English abstract)
- Kim, D., Song, Y., and Lee, S. (2010), The spatial multiplier effect of office rent by the location of a business conglomerate, *Seoul Studies*, Vol. 11, No. 2, pp. 35-49. (in Korean with English abstract)
- Kim, S., Kim, T., Im, H., and Choi, C. (2015), Pedestrian volume and built environmental factors on sales of convenience stores, cosmetic shops and coffee shops in Seoul, *Journal of Korea Planning Association*, Vol. 50, No. 3, pp. 299-318. (in Korean with English abstract)
- Kim, T. and Ko, J. (2015), The impacts of potential consuming population on the store density - the case of food and beverage service activities in Seoul, *Urban Design*, Vol. 16, No. 3, pp. 101-116. (in Korean with English abstract)
- Korea Communications Commission (2015), Providing 119 Wi-Fi location information platform service for emergencies, *Press Releases of Korea Communications Commission*. (in Korean)
- Lee, I., Lee, C., and Kang, S. (2010), The study of site factors affecting the sales of convenience stores, *Journal of the Korea Real Estate Analysts Association*, Vol. 16, No. 4, pp. 53-77. (in Korean with English abstract)
- Lee, K. and Jung, C. (2014), The effect of time period pedestrian volume on store location - focused on the suwon's retail stores and restaurants, *Journal of the Architectural Institute of Korea Planning & Design*, Vol. 30, No. 8, pp. 47-55. (in Korean with English abstract)
- Lim, P. and Lee, S. (2013), Estimation of prices and rents of knowledge industrial centers in seoul metropolitan area considering spatial autocorrelation, *Journal of the Korea Real Estate Analysts Association*, Vol. 19, No. 2, pp. 5-20. (in Korean with English abstract)
- Shin, W. and Moon, S. (2011), A study on the effects of locational characteristics on the sales of a coffee shop franchise, *Journal of the Korea Real Estate Analysts Association*, Vol. 17, No. 2, pp. 111-123. (in Korean with English abstract)
- Shin, W., Jung, C., and Lee, S. (2001) A study on the changing patterns of the commercial power based on retail trade area analysis, *Journal of the Korea Real Estate Analysts Association*, Vol. 8, No. 1, pp. 1-11. (in Korean with English abstract)
- Yu, K., Jang, S., and An, Y. (2016), An analysis on the impact of commercial street factors and floating population on commercial sales by region using the structural equation model, *SH Urban Research & Insight*, Vol. 6, No. 2, pp. 137-152. (in Korean with English abstract)