

## 중력모델(Gravity Model)을 통한 상권분석

김주영\*, 김정수\*\*

빅데이터와 AI를 활용하여 산업은 빠르게 변화하고 있으며, 코로나로 인한 팬더믹은 이러한 변화를 가속시키고 있다. 특히 유통산업에서는 배달 등을 통한 비대면 거래가 급증하는 모습을 보였다. 이러한 상황에서 점포를 중심으로 한 판매상권보다는 소비자 거주지 중심의 소비상권이 더 중요해질 수 있다. 즉 유통산업의 변화를 정확히 파악하기 위해서는 판매상권과 더불어 소비상권이 동시에 고려되어야 할 것이다.

본 연구에서는 중력모델을 활용하여 소비상권과 판매상권을 통합적으로 분석하고자 하였다. 먼저, 신용카드 패널 데이터를 사용하여 소비상권과 판매상권을 추출하였다. 이를 바탕으로 판매상권과 소비상권에서의 업종 및 소비자의 나이에 따른 구매력의 차이와 거리에 따른 민감성을 분석하였다. 두 번째로, 추출된 판매상권과 소비상권별로 각각의 영향력을 분석하여 상권의 특성에 따라 상권의 영향력이 달라지는 것을 알 수 있었다. 세 번째로, 대형몰이 새롭게 진입하였을 때 상권의 영향력과 업종별 거리에 따른 민감성을 분석하였다. 마지막으로 이러한 연구결과를 바탕으로 시사점과 향후 연구방향에 대하여 제시하였다.

주제어 : 상권, 판매상권, 소비상권, 상권분석, 중력모델

### I. 서론

코로나로 인한 팬더믹 현상은 빅데이터와 인공지능(AI)을 통하여 발전되어 가던 산업의 변화 속도를 훨씬 더 빠르게 만들었다. 특히 유통산업은 이러한 영향을 직접적으로 받게 된 대표적인 산업으로 유통산업을 구성하는 기업들과 임직원은 새로운 변신을 과감하게 시도하고 있다. 유통산업에서의 변화 중 가장 두드러지는 현상은 배달을 통한 비대면 거래의 활성화라고 볼 수 있다. 과거에도 물론 온라인 주문을 통한 배달이 있었지만, 많은 유통업체들은 실제 매장이나 점포를 기반으로 발달하여왔다. 심지어 온라인 전문 유통업체들도 오프라인 매장을 열어 소비자들에게 다양한 경험을 주고자 노력해

왔다.

하지만 최근 배달의 비중이 매우 커짐에 따라 배달을 통해 효용이 증가하는 유통과 배달이 효용을 증가시키지 않는 유통으로의 재편을 예견해 볼 수 있다. 온라인쇼핑 운영업체를 대상으로 조사한 통계청의 ‘온라인쇼핑동향조사’에 따르면 온라인쇼핑 거래액은 2019년 약 136조에서 2020년 약 159조로 빠르게 성장하고 있다(통계청 2021a). 특히 배달 등의 음식 서비스 부분의 매출액은 약 17조 규모로 전년 대비 78.0% 증가하였으며, 음식료품의 매출액은 약 19조로 전년 대비 46.4% 증가하였다. 또한 외식업체를 대상으로 조사한 ‘외식업체경영실태조사’에서는 2020년 배달앱을 이용하는 업체가 19.9%로 전년 대비 8.7%p 증가한 모습을 보이고 있다(통계청 2021b).

이러한 추세는 상권에 대한 시각에도 변화를 요구

\* 서강대학교 경영대학 경영학부 교수 (jkimsg@sogang.ac.kr), 제1저자

\*\* 서강대학교 대학원 경영학과 박사과정 (esko445@sogang.ac.kr), 교신저자

하고 있다. 지금까지의 상권은 주로 판매상권의 관점에서 생각되어 왔다. 하지만, 배달을 통한 구매가 보편화 된다면 소비자의 거주지를 중심으로 한 소비상권이 중요해질 수 있다. 소비상권이란 다른 지역과 구분되는 구매력을 가진 소비자가 소비할 수 있는 곳을 의미한다(조춘한, 정강욱, 정환, 김주영 2020). 소비상권의 관점에서 보면, 방문이나 체험이 필요한 식당, 오락과 쇼핑을 중심으로 한 업종들은 소비상권과 매우 근접한 곳 또는 차량 등을 활용하여 이동소요 시간이 적거나 이동편리성이 높은 곳에 위치하게 될 것이다. 하지만 물류창고 및 온라인 주문 전용 픽업 장소로 사용되는 다크스토어(Dark-store)나 공유주방과 같은 곳의 경우 소비상권의 주변에 같이 형성되겠지만, 과거의 판매상권과 같은 모습은 아닐 수 있다. 나아가 매출자체도 해당지역이 아닌 본사나 배달전문앱, 혹은 기타 결제대행사 등 다른 곳에서 발생할 수 있기 때문에 추후 연구 분석도 쉽지 않을 수 있다.

따라서 소비자의 소비 및 거주 위치를 중심으로 한 소비상권과 판매상권의 움직임을 같이 연구해야 지역에서의 유통산업의 변화를 제대로 파악할 수 있게 될 것이다. 본 연구에서는 중력모델(Gravity Model)을 통하여 소비상권의 영향력과 판매상권의 영향력을 모두 분석하고자 한다. 또 소비상권과 판매상권간 거리의 민감도를 파악하고, 업종별로 이러한 소비상권 및 판매상권의 영향력과 민감도가 달라지는지 분석하고자 한다. 나아가, 상권연구가 많이 활용되고 있는 상권영향평가분야에서 해당지역에 대형유통점이 입점 전후로 상권의 영향력과 민감도 변화를 평가하는 방안에 대해 제안하고자 한다.

## II. 이론적 고찰

### 1. 상권연구

상권은 기업이 제공하는 제품 또는 서비스를 구매

할 가능성이 있는 잠재고객이 존재하는 지역이라고 정의할 수 있다(Huff 1964). 또한 특정기간 동안 점포에서 판매되는 제품이나 서비스를 구매하기 위해 소비자가 이동하는 영역이라고 할 수 있다(Applebaum and Cohen 1961). Applebaum and Cohen(1961)은 상권의 영역이 점포의 판매방식, 판매하는 제품이나 서비스 유형, 소비자의 쇼핑 습관, 기존 및 신규 경쟁자에 의해 영향을 받으며, 이러한 영역은 구매 패턴의 변화에 따라 유연하게 결정된다고 하였다. 즉 상권은 고정되어 있는 영역이 아니며, 상권을 구성하는 요소에 따라 변화하는 모습을 보이고 있다.

상권과 관련하여 주로 신규업체의 진입이 상권에 미치는 영향에 대해 연구되어왔다. 권태구, 성낙일(2014)은 소매업종별 사업체 폐널자료를 활용하여 대형할인점과 기업형 슈퍼마켓이 소매업종 사업체수에 미친 영향을 분석한 결과, 대형할인점과 기업형 슈퍼마켓이 신규로 입점할 경우 소규모 슈퍼마켓과 식료품 소매업체의 사업체의 감소에 영향을 미치는 것을 밝혔다. 서진형, 조춘한(2020)은 초저가 프랜차이즈인 다이소의 입점이 주변 상권에 미치는 영향을 분석하여, 다이소의 입점이 집객을 유도하는 효과를 확인하였다. 음식점과 편의점의 매출은 다이소 매출액과 함께 증가하였고, 화장품과 의류의 매출액은 감소하였으며, 문구점과 생활잡화의 매출의 경우는 입지와 거리에 따라 상이한 결과를 보였다. 조춘한 등(2020)의 연구에서는 대규모점포의 출점을 영향을 판매관점과 소비관점으로 나누어 분석한 결과, 대규모점포의 출점으로 인한 영향은 슈퍼마켓, 식당 등 업종과 대규모점포와의 거리에 따라 긍정적인 영향과 부정적인 영향이 복합적으로 발생한다는 것을 밝혔다. 조영재, 최동욱(2021)은 편의점이 일정수준 증가할 경우 지역의 고용률을 증가시키고 매출액을 증가시키며, 경쟁수준이 증가할 경우 개별 편의점의 고용과 매출이 감소하는 것을 확인하였다.

그밖에도 상권과 관련하여 다양한 연구가 진행되었다. 권용걸, 강양석(2002), 이태정, 조성현(2007)은 인구학적 특성, 지가, 역세권, 소득, 거리 등의 요인이 대형할인점의 매출에 영향을 미치는 요인이라는 것을 밝혔다. 박성용(2003)은 소비자를 대상으로 대형할인점, 슈퍼마켓, 재래시장에 대한 인식조사를 통해 대형할인점은 슈퍼마켓 및 재래시장과 경쟁관계에 있다는 것을 확인하였다. 서용구, 한경동(2004)은 시장 영업면적, 백화점 영업면적 등 소매 경쟁구조가 심화될수록 포화지수는 높아지며, 인구 거주 수, 이동교통 특성이 대형할인점의 입지 여건 향상시키는 등 대형할인점의 적정 점포수와 포화수준에 대하여 연구하였다.

## 2. 중력모델

중력이란 뉴턴의 물리학의 중력의 법칙(Newton's Law of Universal Gravitation)에서 유래한 것으로 서로 떨어져 있는 물체들 간에 존재하는 힘을 의미한다. 중력의 기본원칙은 서로 다른 위치에 존재하는 물체들의 사이에는 각자의 질량에 비례하여 서로 끌어당기는 힘이 있다는 것이다. 사회과학에서는 중력이란 개념을 활용하여 지리적 공간에 위치한 사람들 간의 상호작용효과(Interaction Effect)를 설명하는데 많이 활용하고 있다(Fotheringham 2001; Haynes and Fotheringham 1984). 유통분야에서는 Reilly(1931)의 소매인력법칙(Law of Retail Gravitation)에서 소매점의 판매 잠재력을 측정하는 것부터, 시장경제를 추정하거나 점포나 몰 등의 수요예측, 최적의 위치선정 등 많은 곳에서 활용되어 왔다(Davies 1976; Ghosh 2001; Mayo, Jarvis and Xander 1988).

Huff(1964)는 중력모델을 활용하여 점포의 효용성은 점포의 규모에 비례하고 점포와 고객 간의 거리에는 반비례한다는 확률적 상권모형을 제안하였

다. Stanley and Sewell(1976)은 점포의 매출액과 선호도가 점포의 규모와 거리에 의해서만 영향을 받는다는 Huff의 모형을 비판하였다. 점포에 대한 소비자의 이미지도 점포의 선호도에 영향을 미칠 수 있기 때문에, 이를 반영하여 Huff의 중력모델을 개선하였다. Nakanishi and Cooper(1974)는 점포의 효용성은 규모뿐만 아니라 상품의 구색, 가격, 점포 이미지, 종업원의 서비스, 점포 접근성 등 점포의 매력도도 영향을 줄 수 있으며, 이를 변수로 활용하여 공간상호작용모형인 다중경쟁 상호작용모델(Multiplicative Competitive Interaction)을 제안하였다. 이외에도 지역적인 특성에서의 거리가 아닌 심리적인 거리나 선호의 차이 등을 활용한 개인수준의 제품선택모델에서도 활용되어지고 있다(김주영 1994; DeSarbo, Kim, Choi, and Spaulding 2002; Timmermans 2001).

Reilly의 중력모델은 두 개의 쇼핑물(A, B)의 중간지점에 소비자가 위치해 있을 때, 두 곳에 대한 매력도가 동일하다고 가정한다. 이때 두 개의 쇼핑물에서 소비자까지 거리( $d_A, d_B$ )의 비율은 두 개의 쇼핑물 규모( $S_A, S_B$ )의 비율의 제곱근 값과 같다고 나타낼 수 있다(수식 1). 이를 조금 풀어쓰면, 특정 쇼핑물에서 해당지점까지의 거리는 아래 수식 (2)와 같이 계산되어 진다. 따라서 중력모델은 두 개 혹은 그이상의 쇼핑물의 규모에 따라서 추정할 수 있는 지역적인 상권의 모습이나 크기를 찾는 데 활용되고 있다.

$$\frac{d_A}{d_B} = \sqrt{\frac{S_A}{S_B}} \quad (1)$$

$$d_A = \frac{d_A + d_B}{1 + \sqrt{\frac{S_B}{S_A}}} \quad (2)$$

본 연구에서는 중력모형을 활용하여 소비자가 점포에서 구매하는 구매량을 점포가 속해있는 판매상권과 소비자가 속한 소비상권의 크기 및 소비자와 상권 간의 지리적 거리로 설명하는 모형을 만들고자 한다. 이는 Reilly의 중력모형이 응용한 뉴턴의 기본적인 중력법칙을 조금 다른 방법으로 유통분야에 적용한 것이라고 할 수 있다.

기본적인 뉴턴의 중력법칙은 수식 (3)과 같이 두 물체( $i, j$ )간 중력( $F_{ij}$ )은 각 물체의 질량( $M_i, M_j$ )과 그들 간의 거리 ( $D_{ij}$ ) 그리고 중력상수( $G$ )에 의하여 결정된다는 것이다.

$$F_{ij} = G \times \frac{M_i \times M_j}{D_{ij}^2} \quad (3)$$

본 연구에서는 이를 응용하여, 두 개 중의 하나의 물체는 소비상권( $C$ )이고, 다른 물체는 판매상권( $S$ )으로 정의하여 아래 수식 (4)를 만들었다. 여기에서 중력이란, 소비상권이 클수록, 또 상대가 되는 판매상권이 클수록 서로 끌어당기는 힘이 커지는 것을 의미하여, 이것은 서로 간에 발생하는 구매 혹은 매출이 크게 된다는 것으로 추론할 수 있다. 물론, 중력이 클지라도, 소비상권과 판매상권 사이에 장애물, 예를 들면 국경이 존재하여 직접방문 또는 쿼터, 관세절차 등의 영향을 받는다면 이를 감안해야 할 것이며, 강이나 고속도로 및 다리가 있는 것이라면, 물리적 거리가 아닌 도달할 때까지의 시간으로 거리를 대신 할 수도 있다. 또 소비상권의 크기가 같더라도 특정 소비상권은 더 큰 구매력을 보일수도 있고, 역시 같은 판매상권 크기라 할지라도 더욱 큰 판매력을 보이는 상권도 있을 수 있겠다.

소비상권의 영향력을 위한 모수( $\alpha$ )와 판매상권의 영향력을 위한 모수( $\beta$ )를 추가하여 어느 것이 더 큰 영향을 미치는지 파악하고자 했다. 거리는 자승하지 않았는데, 이는 거리를 가변수(Dummy)화하여

거리의 영향이 자승의 관계인지 아닌지 쉽게 알 수 있기 때문이다.

$$F_{ij} = \frac{C^\alpha \times S^\beta}{D_{ij}} \quad (4)$$

### III. 연구모델 설정

#### 1. 자료 설명

연구를 위해서, 대형몰을 비롯해 다양한 유통점포들이 있는 인구수 30-35만 정도의 혁신도시들 중 한곳을 선택하였다. 종적으로 대략 30km, 횡적으로 대략 30km에 해당하는 지역에서 2014년 8월부터 2019년 7월까지 매달 소비자들이 지역의 점포에서 구매한 구매액, 소비자들의 점포 이용건수, 및 해당 점포에서의 판매건수 등을 연구변수로 선정했다.

기본 자료는 신용카드 패널 데이터를 사용하였다. 해당 자료는 개인정보보호법의 권고에 따라 500m 단위로 지역이 구분되어 있다(조춘한 등 2020). 신용카드 패널 데이터의 지역들을 실제 지도를 보면서 해당 소집단(237개)의 위치를 확인하였다. 카드소유자가 구매를 한 점포의 위치도 같은 방법으로 점포들 소집단(160개)의 위치를 확인하였다. 이때 거주지집단의 위치와 점포집단의 위치는 정확하게 일치하지는 않는다. 거주지집단과 점포집단은 개인정보보호법의 권고에 따라 익명성을 위해서 집단으로 묶은 것이며, 이를 군집분석방법을 통하여 합하여 소비상권과 판매상권으로 구분하였다.

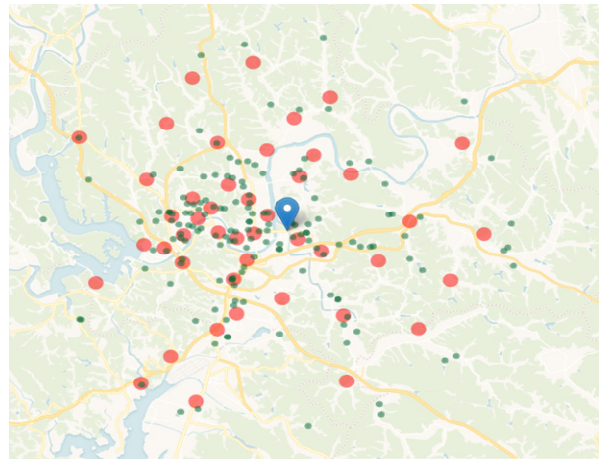
추가적으로 동일한 점포집단의 점포들을 4가지의 업종(편의점, 음식점, 슈퍼/할인점, 패션점, 대형몰)으로 구분하였으며, 해당기간 중에 새로 오픈한 점포와 기존점포를 구분하였다. 동일한 거주지집단의 소비자들은 나이를 30세 미만, 45세 미만, 45세

이상 총 3가지로 구분하였다. 각 거주지집단의 소비자들의 구매액은 각 점포집단의 점포에서 소비자들이 구매한 금액과 방문기록을 한 달 단위로 합한 값을 활용하였다.

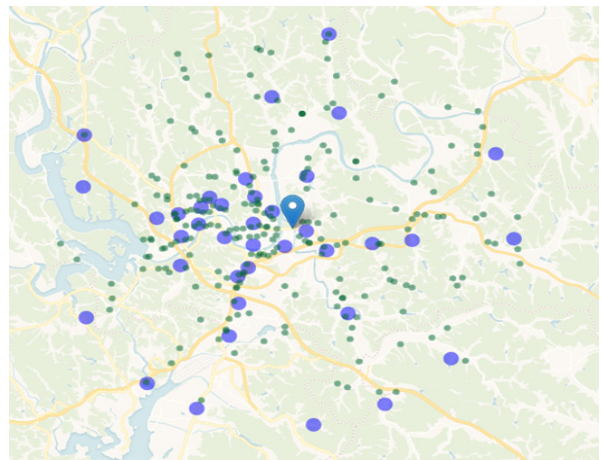
따라서 실제 자료는 업종별(5) x 소비자 나이별(3) x 점포 신규여부(2) x 월별기간(59) x 거주지집단 수(237) x 점포집단 수(160) 별 거래기록으로 되어있다. 이때, 해당조건에 맞는 거래건수 등이 5건이 안 되면, 개인정보유출의 위험이 있는 것으로 판단되어 거래기록이 없는 것으로 처리하였으며, 거주지집단과 점포집단의 거리가 너무 멀 경우 거래가 거의 형성되지 않기에 16km이내의 것들만 분석대상으로 선정하였다. 최종적으로 사용된 총 거래의 기록은 1,984,846개로 집계 되었으며, 거래기록에는 거래금액, 소비자들의 이용건수, 점포의 판매건수가 포함되어 있다.

## 2. 소비상권과 판매상권의 추출

본 연구에서 활용할 중력모델을 적용하기 위해 상권을 기준으로 개별적인 소비자 및 판매 데이터를 재구성하였다. 소비상권과 판매상권을 추출하기 위해 K-평균 군집화(K-Means Clustering)방법을 사용하였다. 자료가 빅데이터이기에 파이썬(Python)을 활용한 사이킷런(Sklearn) 분석프로그램(<https://scikit-learn.org/stable/>)의 K-평균 군집화 방법을 이용하였다. 이 방법은 기본적으로 뭉치고자 하는 집단의 수를 정하고, 각 집단 내에 속한 것들의 제곱합(sum of square)을 최소화시키기 위해 집단 안에 있는 것들을 다른 집단으로 이동시키는 방법이다. 간단하고 효과적인 방법이기에 때문에, 머신러닝(machine learning)의 비지도(unsupervised) 학습에 사용되는 대표적인 방법이다 (Arthur and Vassilvitskii 2007).



<그림 1> 소비상권과 소비자 거주지집단



<그림 2> 판매상권과 점포집단

본 연구에서는 소비 및 판매상권의 적절한 개수를 도출하기 위하여 엘보우(Elbow) 기법을 활용하였다. 엘보우 기법은 각 클러스터의 중심과 각 데이터들 간의 거리를 나타내는 관성(Inertia)이라고 하는 적합도가 집단의 수가 커지면서 그때마다 추가적으로 좋아지는 적합도가 줄어들게 되는데, 더 이상 적합도가 별로 증가하지 않게 되기 직전의 집단의 수를 선택하는 것이다. 이때, 가중치는 각 거주지집단과 점포집단이 가지고 있는 구매금액 및 판매금액의 로그변환값을 표본 가중치로 삼았다. 기본적으로

로 K-평균 군집화 방법은 탐색적 방법이기 때문에 초기값에 따라서 약간의 변화가 있어서 2-3번 시도한 결과 최적 소비상권의 개수는 45개, 판매상권은 38개로 도출되었다. 물론 연구목적에 따라서 적정 상권을 인위적으로 나눌 수도 있다. 본고에서도, 판매상권중 대형몰에 속한 점포들을 별도로 한 개의 판매상권으로 배정하였다.

<그림 1>을 보면 소비자들의 거주지 집단과 이들을 묶은 소비상권의 중심점(총 45개)을 볼 수 있고, <그림 2>에서는 점포집단과 이를 묶은 판매상권의 중심점(총 38개)을 볼 수 있다. 이때 각 그림의 초록색원은 소비자와 점포집단이고, 빨간색원은 소비상권, 파란색원은 판매상권이며, 지도 가운데 표시는 대형몰을 의미한다. 각 소비상권 혹은 판매상권별로 포함된 거주지집단과 점포집단의 수는 각각에서 나오는 구매액이나 판매액에 따라 달라질 수 있다.

### 3. 소비상권의 규모와 판매상권의 규모 (Buying Mass and Selling Mass)

소비 및 판매상권의 크기는 전통적인 중력모델의 특정물체의 질량에서 유추한 것이다. 즉, 질량이 커질수록 서로 당기는 힘이 커지는 것과 마찬가지로, 이를 바탕으로, 소비상권의 규모는 소비상권을 형성하고 있는 소비자들의 숫자와 평균구매액으로 정의할 수 있다. 판매상권의 규모는 판매점들의 숫자 및 평균매출액으로 판매상권의 규모를 정의할 수 있을 것이다.

소비상권의 규모를 측정하기 위해서는 해당지역에 거주하는 사람들의 숫자 및 인구분포 혹은 가능하다면 소득수준과 연간 구매액 등을 활용할 수 있다. 이런 정보 중 가장 중요한 것은 구매액이며, 다른 정보들은 구매액을 추정하기 위한 정보들이라고 할 수 있다. 판매상권의 규모를 측정하기 위해서도 판매점포의 수와 매장평수, 업종의 종류와 판매제

품의 가격대나 점포들의 경쟁력과 업계의 트렌드 등도 파악하면 좋을 것이다. 하지만, 궁극적으로는 해당 판매상권에서 발생하는 판매액을 파악하는 것이 정확한 판매상권 규모를 측정하는데 효과적일 것이다. 물론 소비상권과 판매상권 규모는 연구목적에 따라 다른 정보로 구성할 수 있다.

본 연구에서는 아래 수식 (8)과 같이 종속변수로 구매액을 활용하기 때문에, 독립변수의 역할을 하는 소비상권의 규모와 판매상권의 규모를 소비상권의 구매액이나 판매상권의 판매액으로 하는 것은 종속변수와 같은 정보를 활용하는 것이라고 판단하였다. 따라서 소비상권의 규모는 소비상권의 소비자들의 해당지역 점포들의 이용건수의 총합, 판매상권의 규모는 판매상권을 구성하는 점포들의 판매건수의 총합으로 하였다.

위에서 언급한 바와 같이 동일한 지리적인 위치의 소비상권이라 하더라도, 그 안에 구매력이 높은 집단과 낮은 집단이 있을 수 있다. 학생들이 많은 지역과 중장년층이 많은 지역은 판매상권을 형성하는 점포들의 업종에 따라서 다른 구매력을 보일 것이라고 판단된다. 따라서, 동일한 지리적 위치의 소비상권을 다시 세분화하여, 나이 대에 따라 3가지의 소비상권(소비자 거주지 집단 특성에서 나눈 것과 동일하게)으로 구분하여, 각각의 나이별 소비상권의 구매력의 차이, 즉 중력에 대한 영향력의 차이를 보고자 하였다.

판매상권의 규모의 중력에 대한 영향력도 판매력이라고 본다면, 신규점포의 판매력이 강하다고 할 수 있다. 이러한 논리는 신규점포의 상권은 변화하는 소비자들 및 업의 추세를 따르기 때문이라고 볼 수 있고, 반대로 기존점포의 판매력이 강할 수 있다는 논리는 기존 고객군을 이미 확보하고 있고 오랫동안 고객경험 업무역량이 최적화 되어 있기 때문이라고 볼 수 있다. 하지만, 실질적인 검증을 해볼 필요는 있다고 판단되었기에, 판매상권도 그 안에

있는 점포들을 신규점포와 기존점포로 나누어 각각의 판매력 차이를 보고자 한다.

#### 4. 모델의 구축

연구모델에 사용하는 기본적인 기호들은 아래와 같다.

- $i$  : 소비자집단 번호 (1 -  $i$ : 위치에 따라 구분됨)
- $j$  : 점포집단 번호 (1 -  $j$ : 위치에 따라 구분됨)
- $k$  : 점포집단 업종번호 (해당 위치에 특정 업종이 없을 수 있음)  
(1: 편의점 2: 음식점 3: 할인점 4: 패션점 5: 대형몰-2016년 9월 오픈)
- $c$  : 소비상권 번호
- $s$  : 판매상권 번호
- $l$  : 소비자집단 나이 구분(1=30세미만, 2=45세미만, 3=45세 이상)
- $w$  : 점포집단 신규여부(1= 기존, 2= 2014년 8월 이후 신규)
- $t$  : 시간 (2014년 8월 - 2019년 7월: 월단위)
- $F_{ijkt}$  : 시간  $t$ 에서 소비상권  $c$ 에 속한 소비자집단  $i$ 와 업종이  $k$ 인 판매상권  $S$ 에 속한 점포집단  $j$ 간의 Gravity
- $C_{c,t}$  : 시간  $t$ 에서 나이대  $l$ 인 소비상권  $C$ 의 구매규모(이용건수의 합)
- $D_{(ij)}$  : 소비자집단  $i$ 와 점포집단  $j$ 간의 거리
- $ID_{(ij)}$  : 소비자집단  $i$ 와 점포집단  $j$ 간의 거리를 500m 단위로 나눈 구간 가변수
- $S_{skwt}$  : 시간  $t$ 에서 업종  $k$ 일 때 신규여부( $w$ )에 따른 속한 판매상권의 판매규모(판매건수의 합)
- $CN_{ijklkut}$  : 시간  $t$ 에서 소비자집단  $i$ 의 나이대  $l$ 이 점포집단  $j$ 가 업종  $k$ 일 경우 소비자집단  $i$  중 실제 이용건수
- $SN_{ijklkut}$  : 시간  $t$ 에서 소비자집단  $i$ 의 나이대  $l$ 이 점포집단  $j$ 가 업종  $k$ 일 경우 점포집단  $j$ 에서 일어난 판매건수

- $DUR_t$  : 2014년 9월부터  $t$ 시점까지의 누적월수
- $SA_{ijklkut}$  : 시간  $t$ 에서 소비자집단  $i$ 의 나이대  $l$ 이 점포집단  $j$ 가 업종  $k$ 일 경우 구매한 총 금액

모델의 모수는 아래와 같다.

- $\alpha_{kl}$  : 나이대에 따른 소비자집단의 구매력이 중력에 미치는 영향력 조정값
- $\beta_{kw}$  : 업종과 신규여부에 따라 점포집단의 판매력이 중력에 미치는 영향력 조정값
- $\gamma_{(ij)k}$  : 업종별 소비자집단과 점포집단의 거리구간  $ID_{(ij)k}$ 가 중력에 미치는 영향력 가중치
- $a_k$  : 업종  $k$  별 기본매출
- $\delta_k$  : 시간흐름이 매출에 미치는 영향 (시간의 흐름에 따른 영향을 제거하기 위한 통제역할)

기본적인 중력모델은 위의 수식 (4)인데, 소비상권과 판매상권의 규모를 설정하면서, 중력에 차별적으로 영향을 미칠 것이라고 예상되는 소비상권의 특징과 판매상권의 특징을 감안하여 아래 수식 (5)와 같이 모델을 구축하였다.

$$F_{i(c)j(s)kt} = \frac{\exp(\sum_{l=1}^3 \alpha_{kl} \times C_{c,t}) \times \exp(\sum_{w=1}^2 \beta_{kw} \times S_{skwt})}{\gamma_{(ij)k} \times ID_{(ij)}} \quad (5)$$

수식 (5)에서 보면, 업종이  $k$ 이고, 해당 개월이  $t$ 인 경우, 소비상권  $C$ 에 속한 소비자 거주지 집단, 혹은 소비자집단  $i$ 와 판매상권  $S$ 에 속한 점포집단  $j$ 간의 중력  $F_{i(c)j(s)kt}$ 는 지수함수(Exponential Function)로 표시된  $i$ 가 속한 소비상권의 규모와 그 영향력  $\alpha_{kl}$ 와 역시 지수함수로 표시된  $j$ 가 포함된 판매상권의 규모와 그 영향력  $\beta_{kw}$ 을 곱한 것이 분자에 있다. 소비상권의 규모는 나이대별로, 판매상권은 점포가 신규인지의 여부에 따라서 규모를 다르

게 산정하였다. 또, 영향력  $\alpha_{kl}$ ,  $\beta_{kw}$ 는 업종별로 같은 규모라 할지라도 영향력이 다를 수 있음을 파악하기 위한 것이다.

본모에는 소비자집단과 점포집단의 거리를 500m 단위로 구분한 가변수와 각각의 가변수의 영향력  $\gamma_{(ij)k}$ 를 곱한 것이 있다.

거리를 그냥 쓰지 않고, 가변수화 한 것은 거리의 영향력이 선형 혹은 거리의 자승 혹은 다른 형태의 함수인지 사전에 알 수 없기 때문에 가변수화하여 각 구간별 영향력인  $\gamma_{(ij)k}$ 의 추정값을 보면 보다 정확하게 알 수 있기 때문이며, 역시 업종별로 거리의 영향력이 다를 것이라고 예상하고 있다. 즉, 편의점은 거리가 멀어지면 급격하게 중력이 작아질 것이지만, 패션점은 그보다는 거리의 영향력이 완만할 것이라고 예상할 수 있다. 소비상권과 판매상권간의 거리를 쓰지 않은 것은 보다 정확한 거리자료를 쓰는 것이 더 바람직하다고 생각했기 때문이다.

구체적으로 소비상권의 규모와 판매상권의 규모는 아래 수식 (6)과 (7)과 같다.

$$C_{clt} = \sum_{t'=t-2}^t \sum_{w=1}^2 \sum_{k=1}^5 \sum_{j=1}^J \sum_{i \in c} CN_{ijklwt'} \quad (6)$$

$$S_{swt} = \sum_{t'=t-2}^t \sum_{l=1}^3 \sum_{k=1}^5 \sum_{i=1}^I \sum_{j \in s} SN_{ijklwt'} \quad (7)$$

소비상권은 소비자집단이 지역 내의 모든 점포집단에서 이용한 건수를 나이별로 더한 것을 기본으로 하였다. 여기에 매달 변동이 있을 수 있기에, 3달간의 이용건수의 합을 소비상권의 규모로 하였다. 판매상권도 역시 해당지역에 거주하는 모든 소비자집단이 특정 점포집단에서 찾아가서 발생한 판매건수를 신규점포와 기존점포를 구분하여 모두 더한 후에, 역시 3달간의 판매건수를 더한 것을 활용하였다.

$$SA_{i(c)j(s)klwt} = f(F_{i(c)j(s)kt}, DUR_t) \quad (8)$$

본고에서 제안하는 중력모델(수식 5)을 활용해서 실제 총 구매액과의 예측모형은 수식 (8)과 같이 개념적으로 설명할 수 있다. 이때, 시간이 경과함에 따라 발생할 수 있는 경기의 흐름 등이나 지역발전과 같은 영향력을 경과되는 개월 수( $DUR$ )를 추가로 생각하였다.

이와 같은 독립변수들과 구매액과의 관계를 선형으로 할 수도 있지만, 중력의 공식자체가 가지고 있는 비선형속성 때문에 추정값의 불안전성과 편차로 인하여 기존의 연구들은 적지 않은 노력을 해왔다. 초기의 중력모델들은 종속과 독립변수를 모두 로그 변환을 하여서 선형으로 만들어서 연구를 진행하여 왔다. 하지만 중력모델의 실제 적용과 추정에 있어서는 문제점이 지적되고 있다. 국제경제에서는 두 개의 국가나 지역 간의 거래에서 국가별 차이도 크고, 국가 간 수출량과 수입량 등이 한쪽이 많은 경우도 있고, 교역량이 기록되지 않는 등 이론적으로도 많은 문제가 나타나서 프로빗(Probit)을 활용한 2단계 측정법을 제안하기도 하였다(Anderson and Van Wincoop 2003; Helpman, Melitz, and Rubinstein 2008).

또 Santos and Tenreiro(2011)는 로그변환을 하는 모델은 종속변수가 실제로 0인 경우에 적용할 수가 없고, 추정치에 편향(bias)이 존재할 수 있어서, 포아송 유사 최우도(Poisson pseudo-maximum-likelihood; PPML) 방법을 활용할 것을 주장하였다(Santos and Tenreiro 2006; Santos and Tenreiro 2011). 그들의 연구에 따르면, 로그변환 후 일반적인 최소자승법(OLS)을 사용하면, 이분산성(heteroskedasticity)이 존재할 경우 시뮬레이션(Simulation) 방법에 의하여 계산해보니, PPML을 통한 추정치보다 거의 2배 이상의 추정치를 보여주고 있음을 보여주었다. 하지만, 그들이 제안하고 있는 PPML방법은 포아송 오차



항에 대한 독립성 가정 등에 대한 문제가 있어 실제 적용에서의 문제점도 지적되고 있으며, 다른 방법으로 제안되고 있는 비선형최소자승(Non-linear Least Square; NLS) 방법도 극소값(Local-minimum) 등의 문제로 최적값을 적절히 찾을 수 있는지에 대한 우려가 존재하고 있다.

본고의 자료는 일정지역에서 동일한 측정방법을 활용한 자료원천에서 나오는 자료이기에 자료가 가진 측정방법의 문제는 훨씬 적다고 생각된다. 또한 종속변수인 거래가 없는 경우도 실제로 거의 발생하지 않고, 거래량을 결정하는 변수들 중 고려 못한 변수들이 특정 상관에만 다르다고 보기가 어려워, 로그변환에 따른 최소자승법을 활용하기로 하였다. 따라서 종속변수와 독립변수들을 모두 로그변환을 하며, 절편을 만들고 필요한 단위를 맞추기 위한 모수를 추가하면, 아래 수식 (9)가 된다. 여기서  $\gamma_{(ij)k}$  앞의 부호는 수식 (5)로 보면 음수로 되어야 하지만, 회귀식에서 음수 또는 양수를 알아서 추정할 것으로 생각되어서 양수로 표현하였다.

$$\ln SA_{i(c)j(s)klwt} = a_k + \left[ \sum_{l=1}^3 \alpha_{kl} \ln C_{clt} + \sum_{w=1}^2 \beta_{kw} \ln S_{swt} + \gamma_{(ij)k} ID_{(ij)} \right] + \delta_{kt} \times DUR_t \quad (9)$$

전체 모델을 로그변환을 하면서  $CN_{ijklwt}$  를 활용해서 만든  $C_{clt}$  의 값과  $SN_{ijklwt}$  을 활용해서 만든  $S_{swt}$  두 값이 모두 크고 한쪽으로 심하게 기울어져 있어 로그변환을 하였다. 이를 통해 이러한 왜도 등이 사라지는 효과도 있었다. 이는 종속변수도 마찬가지이다. 또,  $C_{ilt}$  나  $S_{jkt}$ , 또 소비상권과 판매상권 간의 거리가 업종에 따라서 변하는 것을 보기 위해 수식 (9)는 업종별로 추정하고자 한다. 종속변수와 독립변수를 모두 로그변환을 하는 경우에, 추정된

모수는 탄력도(Elasticity)로도 해석될 수 있다.

정리해보면, 입력 자료는 종속변수로  $SA_{ijklwt}$ , 월별, 나이대가  $l$ 인 소비자집단  $i$ 가 업종  $k$ 의 새로 생겼거나 혹은 기존에 있던 점포집단  $j$ 에서 구매한 구매액과, 독립변수로 이때 소비자집단에서 해당점포집단에서 구매한 사람 수,  $CN_{ijklwt}$ 와 해당점포집단에서 매출이 생긴 건수  $SN_{ijklwt}$  그리고 소비자집단  $i$ 와 점포집단  $j$ 까지의 중심점간의 거리가 활용되었다. 모델의 모수는  $a_k$ (5개),  $\delta_k$ (5개),  $\alpha_i$ (3x5개),  $\beta_{kw}$ (2x5개),  $\gamma_{(ij)k}$ (약 150개 x 5)가 된다.

## IV. 모델적용 및 결과

### 1. 기본 모델의 적정성 검토

본고에서 활용하는 자료는 단적(cross-sectional) 자료의 특성을 포함한 시계열자료이다. 따라서 내생성(Endogeneity)이 발생할 수 있다. 내생성은 모델의 오류항과 독립변수가 모델의 오류항과 독립변수간의 상관관계가 존재하여서 추정되는 모수가 편향(Bias)되거나 통계적 유의성검증이 잘못되는 경우이다(Abdallah et al. 2015; Daranto 2020). 가장 많은 경우는 중요한 독립변수가 포함되지 않는 경우이며, 또 다른 유형은 시계열의 경우, 과거시점의 변수에 따라 현재 시점의 변수가 영향을 받는 경우라고 할 수 있다. 이를 해결하기 위하여 2SLS, 3SLS, GMM, 시차회귀(Lagged regression) 등 다양한 방법과 하우스만 검정(Hausman specification test)과 그와 유사한 내생성 검증방법들이 제시되고 있다(Abdallah et al. 2015; Daranto 2020; Hausman 1978).

하지만, 본고에서는 하우스만 검정을 위해서 필요한 도구(Instrument) 변수를 가지고 있지 못하여 검

&lt;표 1&gt; 업종별 회귀식의 요약

구분	R (모델추정 표본)	R (hold-out 표본)	R-square	Adj. R-Square	자유도	F	유의 수준	DW (모델추정 표본)	DW (hold-out 표본)
대형점	0.233	0.199	0.054	0.051	(36, 11445)	18.300	0.000	1.009	1.169
편의점	0.400	0.396	0.160	0.159	(37, 373884)	1929.286	0.000	1.556	1.548
음식점	0.339	0.340	0.115	0.115	(37, 697500)	2450.943	0.000	1.687	1.624
할인점	0.313	0.320	0.102	0.102	(37, 236257)	723.270	0.000	1.412	1.415
패션점	0.260	0.261	0.067	0.067	(37, 71123)	138.894	0.000	1.469	1.510

증을 수행할 수가 없었다. 이러한 경우, 내생성의 원인이 될 수 있는 변수가 아닌 독립변수를 선택해야 한다. 본 연구는 종속변수가 판매액이며, 독립변수는 이를 소비자 및 판매상권집단으로 묶은 후 방문자수 및 이용건수를 활용하여서 내생성문제가 큰 문제가 되지 않도록 하였다. 이를 해결하는 방법은 전체 집단을 한꺼번에 분석하기보다는 성격이 다른 세부집단으로 구분해서 분석을 한다면 내생성문제를 야기하는 집단이라는 변수를 분석차원에서 포함하는 것이다(Abdallah et al. 2015). 본고에서는 업종별로 구분하여 분석을 진행하였다.

또 시계열로 인한 내생성은 사업자들이 전년도의 결과나 미래의 예측에 따라 알아서 업종이나 메뉴 등을 변화시키기 때문에 완전하게 내생성을 제거하기는 어렵다. 참고로 내생성의 검증은 아니지만 오차항들 간의 계열상관(Serial Correlation)이 있는지를 검증하는 더빈-왓슨 검정(Dubin-Watson test)은 <표 1>과 같이 유의하지 않은 것으로 나타났다. 따라서 본 연구는 내생성 문제를 포함하고 있을 수 있으며 연구결과의 해석에 한계점을 가지고 있다고 할 것이다.

로그변환을 한 후에는 일반적인 선형회귀분석을 통하여 모수 추정값들을 얻었다. 업종별로 5개의 회귀식을 추정했다. <표 1>에서 보듯이 R제곱으로 나온 설명력은 적지만, 사용된 설명변수가 각각의 거래량이 발생한 소비상권과 판매상권에서의 이용건

수와 판매건수와 시간변수라는 것을 감안하면 아주 나쁜 것은 아니며, 회귀식 자체의 유의성은 통계적으로 적절하다고 할 수 있다.

또 다른 모델의 타당성에 대한 검증방법은 홀드아웃(Hold-out) 표본을 따로 가지고 모델의 적합성을 다시 평가하는 방법이다. 따라서 전체 자료의 30%를 무작위로 선별해서 적합도를 보니 <표 1>과 같이 모델추정에 사용된 자료의 결과가 크게 다르지 않음을 볼 수 있었다. 또 두 개의 표본으로 추정한 계수별 차이에 대하여는 <표 2>에 추가적으로 설명되어 있다. 추정추후 분석에서는 모든 표본을 다 사용한 결과를 보여주고 있다.

## 2. 기본 모델의 결과

<표 2>는 업종에 따라 소비자의 나이, 기준 혹은 신규점포, 경과 시간에 따른 차이를 나타낸다. 추정된 모수들을 보면, 업종별로 차이가 있다고 할 수 있다. 먼저, 상수항은 각 업종별 기본 거래규모라고 볼 수 있다. 대형몰과 패션점이 비교적 큰 편이며, 음식점, 할인점 편의점의 순서로 되어있다. 그 외의 모수들은 모두 표준화된  $\beta$ 값을 보여주고 있다. 앞에서 언급된 바와 같이  $\beta$ 값은 기본적인 회귀분석모델로 생각할 때는 탄력도로도 해석될 수 있다. 마지막으로 거리에 대한  $\gamma_{(ij)k}$  값은 가변수화 된 추정치들이기에 0.05 유의수준에 미치지 못한 개수를 보여주고

<표 2> 업종별 모형추정값 및 유의수준

구분	대형몰	편의점	음식점	할인점	패션점
상수*	14.520	11.205	13.157	12.819	13.899
$\alpha_{k1}$ (30세미만)	-0.060	0.174	0.114	-0.369 <sup>^</sup>	-0.179
$\alpha_{k2}$ (45세 미만)	0.089	-0.067 <sup>^</sup>	-0.104 <sup>^</sup>	0.029 <sup>^</sup>	0.356
$\alpha_{k3}$ (45세 이상)	-0.062	-0.120	-0.038	0.292 <sup>^</sup>	-0.280
$\beta_{k1}$ (기존점포)	-0.003	-0.011 <sup>^</sup>	0.001	-0.018	0.002
$\beta_{k2}$ (신규점포)	0.010	-0.009 <sup>^</sup>	-0.035 <sup>^</sup>	0.023	0.020
$\delta_k$ ( <i>DUR</i> , 시간)	0.006	0.055 <sup>^</sup>	-0.007 <sup>^</sup>	-0.098 <sup>^</sup>	0.001
$\gamma_{(ij)k}$ ***	3/30	0/31	0/31	1/31	3/31

\* : 상수만 비표준화값이며, 나머지 모수들은 모두 표준화값(베타값)임  
 \*\* : 모든 모수는 유의수준 0.01에서 통계적으로 유의함  
 \*\*\* : 유의수준 0.05 이상의 추정치 개수  
 ^ : 홀드아웃표본과 모델추정표본으로 추정된 계수들을 비교검토 한 결과 유의수준 0.05이상의 차이를 보이는 추정값들임

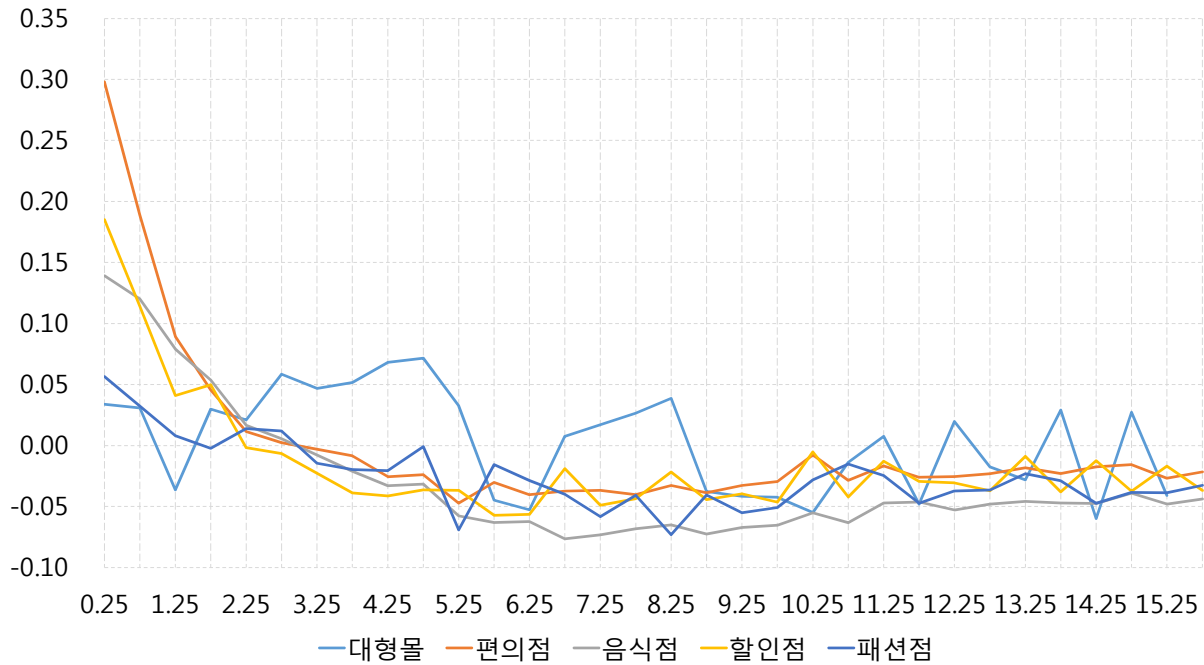
있다. 대형몰과 패션점이 약간 유의미하지 않은 게 있고, 나머지는 적절하다고 볼 수 있다. 유의미하지 않은 경우는 해당 거리에 자료의 수가 적거나 실제 값이 상당한 차이를 보이는 경우라고 판단된다. 이때, 중요모수 추정값들은 전체표본에서는 통계적으로 유의했고, 시간 가변수는 유의하지 않은 것들이 있어서 <표 2> 밑에 표시되어 있다. 예를 들어 대형몰의 (*DUR*) 가변수 30개중 3개는 통계적으로 유의하지 않다. 또 홀드아웃표본으로 추정된 계수와 모델추정표본으로 추정된 계수의 차이를 통계적으로 검증한 결과 편의점의 45세미만 구매력 계수(-0.104)는 통계적으로 유의하게 차이가 났다. 이런 경우 전체표본으로 추정된 것이 유의하다고 해도, 일반화할 때는 주의해야 할 것이다.

대형몰의 경우, 30-45세의 구매력(0.089)이 다른 나이에 비하여 높은 것으로 나타나고 있는 반면, 편의점은 오히려 젊은 나이대인 30세 미만(0.174)이 구매력이 높은 것으로 나타났다. 음식점도 역시 30세 미만인 젊은 층의 구매력(0.114)이 높은 것으로 나타났다. 할인점은 슈퍼 등을 포함하고 있어 새로 생긴

점포(0.023)의 흡입력이 아무래도 기존점포보다는 높은 것을 볼 수 있다. 패션점은 역시 구매력을 갖춘 30-45세의 구매력(0.356)이 높은 것을 알 수 있다.

또 통제변수처럼 생각하고 있는 다른 추정치인 시간의 흐름에 따른 효과를 보면, 편의점과 대형몰, 그리고 패션점은 시간이 흐름에 따라 더 판매금액이 늘고 있고, 할인점과 음식점은 줄고 있는 것을 볼 수 있다. 연구의 중요 관심 추정치인 소비상권과 판매상권간의 거리에 따른 영향력 모수인  $\gamma_{(ij)k}$ 를 보면 아래 <그림 3>과 같다.

<그림 3>에서 보면, 모든 업종에서 가까울수록 더 많이 구매를 하는 것으로 나왔는데, 역시 편의점은 그중에서도 거리에 훨씬 더 민감함을 보여주고 있다. 그 다음으로는 할인점과 음식점이 거리에 민감함을 보여주고 있다. 패션점은 멀리 이동해서도 구매가 이루어지기에 좀 덜 민감한 것으로 나타났다. 흥미 있는 것은 역시 대형몰로서 가까운 소비상권뿐만 아니라 15km 가 넘는 거리의 소비지역에서도 구매를 하는 것으로 나타났으며, 가장 거래가 많이 일어나는 거리는 3-5km 정도이고, 7.5-9km에서도 일어나는 것



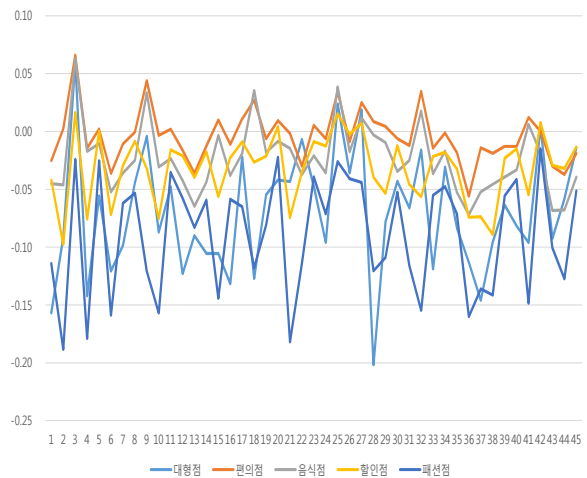
<그림 3> 소비상권과 판매상권간의 거리의 영향

으로 나타났다. 이런 것은 연구대상 지역에 따라 소비상권의 크기가 크면 발생하는 것으로 거리의 영향력이 일정하지는 않을 듯하다. 하지만, 대형물은 거리의 민감성은 거의 없다고 할 수 있겠다.

## 2. 상권별 영향력 분석

소비상권 45개와 판매상권 38개를 실제로 확인하여 모델을 추정하였다. 각 상권별로 판매량에 얼마나 영향을 미치는지를 파악하기 위해서, 수식 (9)를 수정하여, 수식 (10)과 같이 모수  $\alpha$ 와  $\beta$ 를 상권별로 추정하여 보았다. 대신 나이와 신규점포여부와 관련된 모수는 구분하지 않았다.

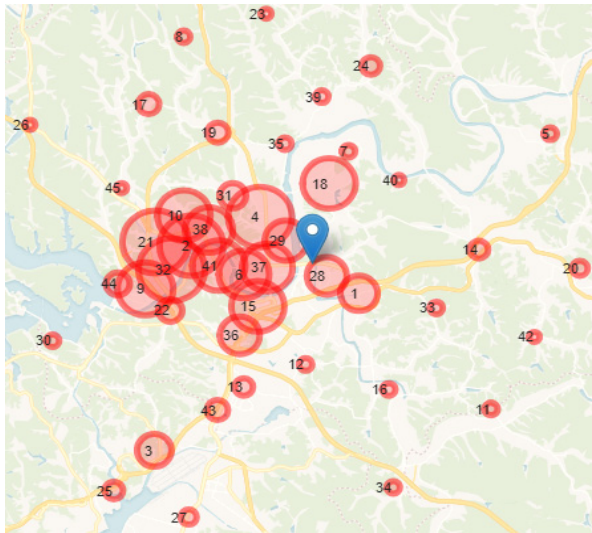
$$\ln SA_{i(c)j(s)kwt} = a_k + \left[ \sum_c \alpha_{ck} \sum_l \ln C_{clt} + \sum_s \beta_{sk} \sum_w \ln S_{swt} + \gamma_{(ij)k} ID_{(ij)} \right] + \delta_{kt} \times DDUR_t \quad (10)$$



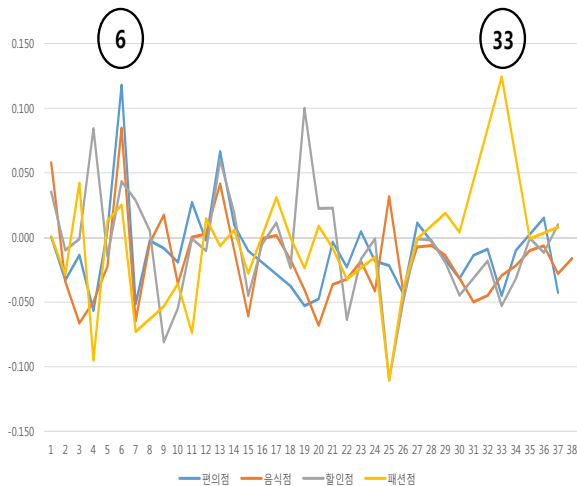
<그림 4> 소비상권 영향력 모수 추정치

<그림 4>와 <그림 5>에는 에서는 실제 소비상권들의 모수  $\alpha_{kl}$ 와 상권 규모,  $C_{clt}$ 의 로그값들을 나이별로 다 더한 후에 그 상권의 영향력 모수  $\alpha_{kl}$ 를 곱한 상권 영향력을 보여주고 있다. 마찬가지로 <그림 4>와 <그림 5>에는 판매상권의 영향력 모수  $\beta_{kw}$ 와

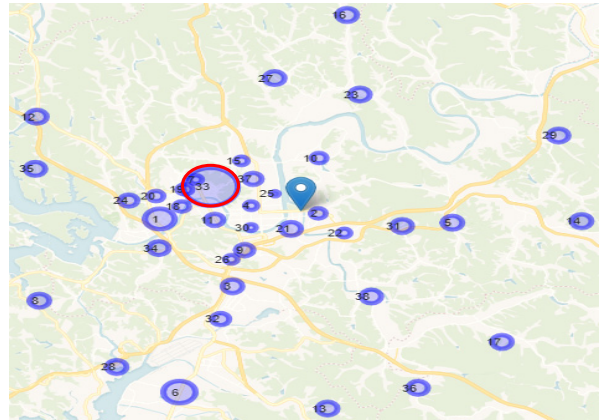
실제 판매상권들의 규모,  $S_{skut}$ 를 로그변화 후 신규 여부와 상관없이 더한 후에 상관별 영향력 모수  $\beta_{kw}$ 를 곱한 판매상권 영향력값을 보여준다. 판매상권에서 대형몰은 별도의 판매상권에서만 판매되기에 해당 회귀식은 추정할 수 없어 제외되었다.



<그림 5> 소비상권들의 영향력(모수 x 상권 규모)



<그림 6> 판매상권 영향력 모수 추정치



<그림 7> 판매상권들의 영향력(모수 x 상권 규모)

그림들에서 보듯이 각 상권별로 상당한 차이들을 보여주고 있다. <그림 4>에서 보면, 수평축에 소비상권이 나타나있고, 수직축에 해당 소비상권이 업종별 모수의 크기가 있는데, 모수의 크기가 클수록 해당상권에서 해당 업종의 구매가 많이 발생한다고 볼 수 있다. 패션점의 영향력 모수가 대부분 0보다 낮고, 차이가 큰 것을 보면 소비상권들 중 패션점을 가는 경향이 상대적으로 적고, 차이도 큰 것을 알 수 있다. 마찬가지로 대형몰도 잘 안 가는 것을 볼 수 있다. 나아가서 소비상권별로 편의점과 음식점의 값이 같이 움직이는 것을 보면, 편의점을 자주 가는 소비상권은 역시 음식점도 즐겨 찾는다고 판단할 수 있다. <그림 5>에는 각 소비상권 영향력의 크기가 해당 소비상권이 위치한 점의 원의 크기로 표시되고 있다. 실제로 로그변환 규모를 곱한 상권별 크기를 보면 실제 소비자들의 거주지역의 소득이나 구매력과 유사한 결과를 보일 것이라고 유추해 볼 수 있다. 이런 분석을 통해서 소비상권별 특성을 좀 더 세밀하게 파악할 수 있을 것으로 보인다.

<그림 6>과 <그림 7>도 유사하게 판단할 수 있다. 먼저 <그림 6>의 판매상권의 영향력 크기를 보면, 백화점이 있는 판매상권(33번)이 크게 형성되어 있음을 볼 수 있다. <그림 7>을 보면 33번 판매상권의

경우 패션점의 영향력이 매우 큰 것으로 나타나고 있으며, 업종별 판매상권들의 모수들을 보면, 패션점이 영향력이 큰 상권에서는 할인점의 영향력이 작은 것으로 나타나고 있다. 그림 상으로는 잘 보이지는 않지만, 6번 판매상권 같은 경우에는 거의 모든 업종에서 영향력을 발휘하는 것으로 보인다. <그림 5>의 비슷한 지역에 소비상권이 형성되어 있지 않음을 보면, 인근지역 소비상권의 소비자들이 도심과 멀리 떨어진 지역상권에서 모든 것을 해결하는 곳임을 알 수 있다. 이러한 분석을 통해서 판매상권들의 특성차이도 쉽게 알아볼 수 있으며, 추후 매장전략에서도 활용할 수 있을 것이다.

### 3. 신규 대형몰의 영향 분석

현재의 모델구조 하에서, 대형몰이 진입한 후의 상권의 변화 등을 비교할 수도 있다. 대형몰이 진입한 후의 업종별 매출의 변화나 지역 판매점포들이 대형점포의 거리에 따라 매출의 변화가 있는지 혹은 소비자들의 대형몰과 얼마나 떨어져 있는지에 따라서 지역 판매점포에서의 구매감소 등은 기존의 연구들에서 연구가 진행되었다(조춘한 등 2020). 그들의 연구에 따르면, 대형몰이 들어오면서 지역 판매점포들의 매출은 별로 감소하지 않았으며, 대형몰과의 업종이 겹치는 경우에는 판매 감소가 나타

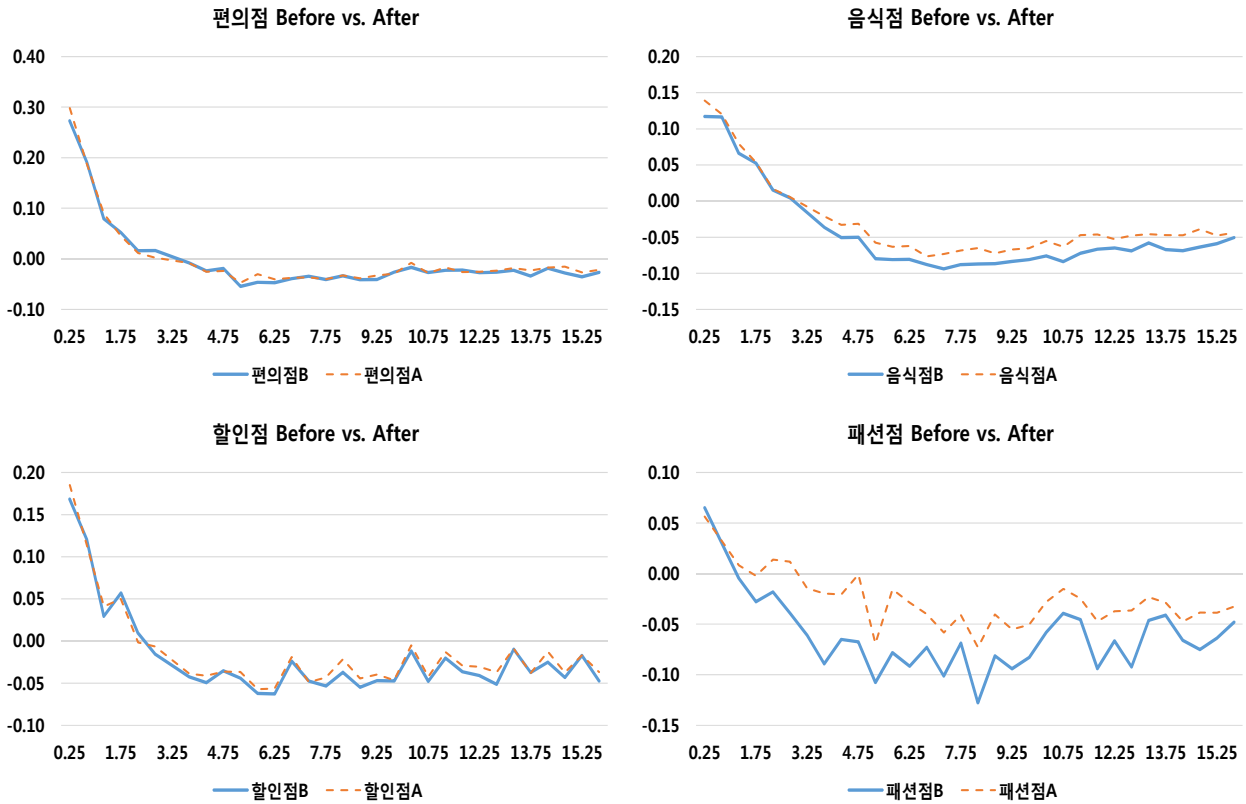
났다. 또 대형몰의 영향이 미치는 거리는 업종별로 차이가 나는데, 이러한 차이는 관심대상 대형몰뿐만 아니라 중형 경쟁점포의 등장에 따라서 얼마든지 변화될 것으로 판단되었다.

본고에서는 소비상권과 판매상권의 영향력이나 판매상권과 소비상권의 거리의 영향력이 어떻게 변화했는지를 보고자 한다. 즉, 반드시 정책을 입안하거나 규제를 위해서 하는 분석이 아니라 지역의 올바른 발전과 개별 기업들의 전략수립에도 도움이 되기 위한 상권의 변화를 파악하는데 도움이 되는 분석을 해보고자 하였다. 이를 위해서 수식 (9)의 모델을 활용하는데, 자료는 대형몰이 진입하기 전 자료들과 진입 후 자료를 구분하여 대형몰을 제외한 업종별 회귀식, 총 8개를 추정하였다.

<표 3>에 나타난 바와 같이 대형몰의 진입 이전과 이후는 약간의 차이를 보이고 있다. 표본의 수가 크기 때문에 모든 모수는 통계적으로 유의하게 나타나고 있다. 유의미한 차이는 굵은 폰트로 표시된 것과 같이 대형몰 진입이후에 판매상권에서 기존점포의 영향력보다는 신규점포의 영향력이 커진다는 것이다. 아무래도 대형몰이 들어오면 경쟁력이 적은 점포는 폐점하고, 새로운 전략으로 새로운 점포가 입점하기 때문으로 해석된다. 또 할인점의 경우에는 장년층인 30-45세의 소비상권의 영향력이 더 커졌다는 것을 알 수 있다. 이는 새로운 해당 층의 유입

**<표 3> 대형몰 진입 전과 후의 상권들의 영향력 모수들**

구분	대형몰 진입 전(Before)				대형몰 진입 후(After)			
	편의점B	음식점B	할인점B	패션점B	편의점A	음식점A	할인점A	패션점A
αk1 (30세미만)	0.213	0.136	-0.415	-0.435	0.164	0.107	-0.363	-0.003
αk2 (45세 미만)	-0.083	-0.130	-0.241	0.424	-0.065	-0.103	0.128	0.335
αk3 (45세 이상)	-0.148	-0.052	0.579	-0.125	-0.106	-0.022	0.194	-0.413
βk1 (기존점포)	0.007	0.006	-0.015	0.018	-0.049	-0.027	-0.023	-0.059
βk2 (신규점포)	-0.012	-0.006	0.004	-0.022	0.033	0.005	0.023	0.084
δk (DUR, 시간)	0.085	0.014	-0.043	-0.006	0.026	0.010	-0.058	0.026



<그림 8> 대형몰 진입 전과 후의 거리의 영향

인구가 소비능력을 갖추고 늘어났다는 것을 반영하였다고도 볼 수 있다.

그 다음으로는 아래 <그림 8>과 같이 소비상권과 판매상권간의 거리가 미치는 영향이다. 편의점과 할인점의 경우에는 소비상권과 판매상권간 거리의 영향이 거의 없다고 할 수 있다. 즉, 항상 가던 곳에 가서 구매를 한다고 생각할 수 있다. 음식점의 경우에는 거리가 멀수록 대형몰 이후에 오히려 더 판매금액이 증가했다. 패션점의 경우에는 15km이전까지 거의 모든 거리에서 판매금액이 증가했다. 이것은 기본적으로 차를 타고 이동하는 것이 늘었다고 판단할 수 있으며, 소득이 높고 이동이 편한 인구의 유입에 따른 효과라고 볼 수 있다. 이는, 대형몰의 패션점포에서 구매하는 것은 포함되지 않은 결과이기 때문이다.

따라서 이러한 분석들을 통하여 대형몰의 진입과 주변의 인구이동 및 소비상권의 특성들로 인하여 지역상권들의 영향력이 변화되는 것들을 파악할 수 있다. 인구이동이나 소득수준의 변화 및 판매점포들의 개별정보가 있다면 좀 더 세밀한 분석들이나 대안가설들도 개발해서 검증할 수도 있을 것이다.

## V. 결론 및 논의

### 1. 연구의 요약

기존의 상관분석은 주로 판매상권을 중심으로 연구되었으나, 배달과 배송이 보편화됨에 따라 소비상권의 중요성이 부각되고 있다. 이에 본 연구에서

는 중력모형을 활용하여 판매상권과 소비상권의 영향력을 통합적으로 분석하였다. 현재의 유통환경에서는 판매상권이 지역적으로 확인이 되고, 매출자료도 지역적으로 확인이 가능하다. 하지만, 무점포에서의 배달이 증가하고 판매상권의 매출이 본사에서만 확인이 되고 지역적 매출정보를 얻을 수 없는 상태가 된다면, 전통적인 매출상권을 통한 분석은 의미가 없어질 수 있다. 오히려 구매상권만을 기반으로 한 연구가 더 설득력이 있게 되며, 그렇게 된다면 소비자가 찾아가는 판매상권의 영향력이 어느 정도가 되는지를 추후에는 확인할 수 있을 것이다. 본고에서는 신용카드 패널 데이터를 활용하여 소비상권과 판매상권의 업종별 거리에 따른 민감도를 파악하고, 대형몰 진입 후 업종별 상권의 영향력과 민감도 변화를 분석하였다.

첫째, 소비상권과 판매상권의 영향력을 분석한 결과 30대 미만 소비자는 편의점, 30세 이상 45세 미만 소비자는 패션점, 45세 이상 소비자는 할인점에서 높은 구매력을 보이는 등 소비자의 나이에 따라 업종별 구매력이 다른 것을 확인하였다. 소비상권과 판매상권의 거리에 따른 민감도 또한 업종별로 상이한 모습을 보였다. 대체로 가까운 거리에서 구매가 많이 발생하였으며, 특히 편의점의 경우 거리에 따른 민감도가 가장 높은 것으로 나타났다. 상대적으로 패션점은 거리에 따른 민감도가 낮게 나타났으며, 대형몰의 경우 15km 이상 떨어진 곳에서도 구매가 발생하는 모습을 보였다.

둘째, 상권을 45개의 소비상권과 38개의 판매상권을 분류하여 각각의 영향력을 분석한 결과, 상권에 따라 구매량 및 판매량의 차이가 나타나는 것을 확인하였다. 분석결과, 특정 소비상권은 패션몰과 대형몰은 거의 이용하지 않는 상권 경우도 존재하였다. 판매상권의 경우 패션점의 영향력이 클 때 할인점의 영향력이 작은 경우가 존재하였으며, 업종에 관계없이 영향력이 높은 상권도 존재하였다.

셋째, 신규 대형몰의 진입 이후 상권의 영향력에 대해 분석하였다. 진입 후 새로 개점한 점포의 영향력이 더 증가하는 것으로 나타났는데, 이는 대형몰이 진입하면서 기존 경쟁력이 낮은 점포는 소멸하고 경쟁력 높은 신규점포가 유입되었다고 볼 수 있다. 할인점의 경우 30~50세의 소비상권 영향력이 증가하였는데 이는 해당 연령층의 신규유입으로 인해 소비력이 커진 것으로 볼 수 있다. 거리에 따른 업종별 민감도를 분석한 결과 편의점의 경우 대형몰 진입 후에도 거리에 따른 영향이 거의 존재하지 않았으며, 음식점의 경우 대형몰 진입 후 거리가 멀어질수록 판매금액이 증가하였다. 패션점의 경우 거의 모든 거리에서 판매금액이 증가하였다. 즉, 대형몰의 진입, 인구이동, 소비상권의 특성 등으로 지역상권의 영향력이 변화하는 것을 보여주고 있다.

## 2. 연구의 시사점

본 연구의 시사점은 다음과 같다. 첫째, 상권분석을 위해 중력모형을 기반으로 하여 소비 및 판매상권의 규모추정에 다양한 변수의 사용가능성을 검증하여 기존 중력모형을 확장시켰다. 상권의 규모를 추정하기 위하여 소비상권에서는 구매액을 중심으로 연령층, 구매건수 등을 활용하였으며, 판매상권에서는 매출액을 중심으로 업종, 매출건수 등을 활용하였다. 본 연구에서 활용한 변수 이외에도 거주하는 소비자들의 인구학적 특성, 점포의 수, 판매하는 제품의 가격, 인근 점포와의 경쟁강도 등 다양한 변수들도 적용시킬 수 있는 유연한 모델을 구축하였다는 점에서 의의가 크다고 할 수 있다.

둘째, 본 연구에서는 상권의 유형을 소비상권과 판매상권으로 구분하여 각 상권의 관점에서 거리에 따른 민감도를 분석하였으며, 업종별로 이러한 민감도의 변화를 분석하였다. 이를 통해 판매상권과 소비상권은 각 상권에 따라 상이한 특징이 존재하는 것을



분석하였다. 이를 통해 상권을 정확하게 분석하기 위해서는 판매상권과 소비상권을 동시에 고려해야 한다는 것을 보여주었다. 특히 배송업과 배달업이 크게 성장함에 따라 도보 등 소비자의 방문을 기준으로 형성된 판매상권 보다, 배달권을 기준으로 한 소비상권의 중요성이 점차 증가하고 있기 때문이다.

셋째, 상권의 영향력을 평가하는 대안을 제시하고 있다. 상권영향평가는 대규모점포의 출점에 따른 주변 상권의 매출액 변화를 예측하는 것으로 대규모점포의 출점을 위해서는 상권영향평가서를 제출하도록 하고 있다. 본 연구의 상권분석모델은 단순히 규제나 정책을 위한 분석뿐만 아니라 지역 상권의 발전과 기업의 전략수립에 활용할 수 있다. 단순히 소비상권 또는 판매상권의 단적인 측면만이 아닌 복합적인 분석을 통해 점포입점 전략을 수립할 수 있다. 예를 들어 특정 상권을 종합 판매상권으로 발전시키기 위해 주변 소비상권의 영향력과 상대적으로 판매력이 약한 업종을 분석하여 약소업종의 성장가능성을 예측할 수 있으며, 나아가 상권의 전반적인 성장을 예측할 수 있을 것이다. 또는 기업에서 새로운 유통 신사업을 계획하거나, 기존 점포를 입점을 위한 전략에도 활용할 수 있을 것이다.

### 3. 연구의 한계 및 향후 연구방향

모델을 활용한 연구의 진행방향은 연구의 목적을 정하여 얻을 수 있는 자료를 확인하고, 그에 맞는 모델의 구조를 구축한다. 구축된 모델을 바탕으로 모수를 추정하고 추정결과를 해석하여 연구목적에 확인하는 것이다. 상권연구의 경우에서 지역적 소비상권은 존재하지만 지역적 판매상권이 줄어들고, 다크스토어와 배달 같은 판매방식으로 매출이 일어나는 경우, 다크스토어의 위치와 그 매출액을 알 수 있으면 기존과 같은 상권연구 모델의 구조로 연구가 가능하다.

본 연구의 모델에서 활용되는 자료는 단면적인

(Cross-sectional) 시계열 자료로서 다양한 내생성문제를 가질 수 있다. 도구변수가 좀 더 있다면, 이런 문제의 심각성을 통계적으로 해결하거나 심각성을 알 수 있었을 것이다. 추후 연구에서는 이러한 문제를 해결할 수 있는 자료로 연구를 해야 할 것이다.

또 다른 자료의 필요성의 측면에서 보면, 본 모델은 소비상권과 판매상권의 거래량이 모두 자료로 확인되어야 적용 가능한 모델로, 최소한 지역적인 위치가 없더라도 매출정보를 확보해야 모델의 적용이 가능하다는 한계점이 존재한다. 향후 배송업 및 배달업이 더욱 성장하게 된다면 소비상권의 영향력은 같지만, 지리적으로 확인되는 판매상권의 영향력은 감소할 것으로 예상된다. 혹은 소비상권은 존재하지만 판매상권은 없는 경우도 나타날 수 있다. 이는 배달만을 전문으로 하는 업체의 매출이 본사의 매출발생으로만 기록되고, 실질적으로 제품이나 서비스가 제공된 점포의 매출로 기록이 되지 않는 경우가 해당될 수 있다. 이러한 경우 배송 및 배달업체에서 자료공유가 안된다면 전체 매출을 측정하기 어렵고 재화의 이동경로를 파악하기 어려울 수 있다. 이처럼 소비자가 구매한 자료만 존재할 경우 모델을 적용할 수 없다는 한계점이 있다.

본 연구에서는 중력모델을 활용하여 판매상권과 소비상권 별로 나타나는 특성을 분석하였다. 각 상권의 영향력의 추정을 위하여, 소비상권은 소비자의 연령층, 구매건수, 구매액을 활용하였으며, 판매상권은 업종, 매출액, 매출건수를 활용하였다. 향후 연구에서는 이러한 변수 이외에도 소비 및 판매상권을 구분할 수 있는 변수를 반영하여 영향력을 측정하고 비교할 수 있다면 상권별 특성을 좀 더 세밀하게 파악할 수 있을 것이다. 예를 들어 브랜드 또는 제품의 유형 등의 변수를 추가하여 모델을 확장시킨다면 상권에 따른 브랜드 및 제품의 소비 및 판매 영향력의 차이와 거리에 따른 민감도를 더욱 세밀하게 분석할 수 있는 모델이 개발될 수 있을 것이다.

논문접수일: 2021.09.01.

1차 수정본 접수일: 2021.10.12.

게재확정일: 2021.10.20.

## 참고문헌

- 권용걸, 강양석 (2002), “대형할인점 입지 결정 요인에 관한 연구,” *국토계획*, 37 (1), 207-217.
- 권태구, 성낙일 (2014), “대규모점포의 시장진입과 소매업종별 사업체 수의 변화: 실증분석,” *경제분석*, 20 (2), 56-91.
- 김주영 (1994), “상표선택 중력모델,” *소비자학연구*, 5 (2), 63-78.
- 서용구, 한경동 (2004), “대형 할인점 포화지수 산출과 결정 요인 분석: 수도권을 중심으로,” *유통연구*, 9 (4), 65-83.
- 서진형, 조춘한 (2020), “초저가 프랜차이즈 업체의 점포 출점이 상권 및 주변점포에 미치는 영향: 다이소에 대한 사례 연구,” *프랜차이즈저널*, 6 (1), 1-26.
- 이태정, 조성현 (2007), “대형할인점의 매출과 입지요인분석-상위 3 개 업체의 전국매장현황을 중심으로,” *응용경제*, 9 (1), 103-134.
- 조영재, 최동욱 (2021), “지역내 편의점 출점의 효과: 2010-2015 경제총조사 자료를 이용한 분석,” *유통연구*, 26 (1), 1-24.
- 조춘한, 정강욱, 정환, 김주영 (2020), “상권영향 평가와 대규모점포 출점규제,” *유통연구*, 25 (4), 1-29.
- 통계청 (2021a), “온라인쇼핑 동향, 통계청,” 2021년 6월호, [https://www.kostat.go.kr/portal/korea/kor\\_nw/1/12/3/index.board?bmode=read&bSeq=&aSeq=391169&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=](https://www.kostat.go.kr/portal/korea/kor_nw/1/12/3/index.board?bmode=read&bSeq=&aSeq=391169&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt=).
- 통계청 (2021b), “외식업체경영실태조사 : 배달 앱 및 배달대행 이용현황,” [https://kosis.kr/statHtml/statHtml.do?orgId=114&tblId=DT\\_114054\\_016&vw\\_cd=MT\\_ZTITLE&list\\_id=K2\\_003\\_001&scrId=](https://kosis.kr/statHtml/statHtml.do?orgId=114&tblId=DT_114054_016&vw_cd=MT_ZTITLE&list_id=K2_003_001&scrId=).
- Abdallah, Wissam, Marc Goergen, and Noel O’Sullivan (2015), “Endogeneity: How Failure to Correct for It Can Cause Wrong Inferences and Some Remedies,” *British Journal of Management*, 26 (4), 791-804.
- Anderson, James E. and Eric van Wincoop (2003), “Gravity with Gravitas: A Solution to the Border Puzzle,” *American Economic Review*, 93 (1), 170-192.
- Applebaum, William and Saul B. Cohen (1961), “The Dynamics of Store Trading Areas and Market Equilibrium,” *Annals of the Association of American Geographers*, 51 (1), 73-101.
- Arthur, David and Sergei Vassilvitskii (2007), “K-means++: The Advantages of Careful Seeding,” in *The Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, New Orleans Louisiana, Association for Computing Machinery, 1027-1035.
- Davies, Ross L. (1976), *Marketing Geography : with Special Reference to Retailing*, London: Methuen & Co Ltd.

- Daranto, Ahmad (2020), "EndoS: An SPSS Macro to Assess Endogeneity," *The Quantitative Methods for Psychology*, 16 (1), 56-70.
- DeSarbo, Wayne S., Ju Young Kim, S. Chan Choi, and Melinda Spaulding (2002), "A Gravity-Based Multidimensional Scaling Model for Deriving Spatial Structures Underlying Consumer Preference/Choice Judgments," *Journal of Consumer Research*, 29 (1), 91-100.
- Fotheringham, Alexander S. (2001), "Spatial Interaction Models," in *International Encyclopedia of the Social & Behavioral Sciences*, Neil J. Smelser and Paul B. Baltes, ed. Amsterdam: Elsevier, 14794-14800.
- Ghosh, Avijit (2001), "Retail Trade," in *International Encyclopedia of the Social & Behavioral Sciences*, Neil J. Smelser and Paul B. Baltes, ed. Amsterdam: Elsevier, 13283-13286.
- Hausman, Jerry A. (1978), "Specification Tests in Econometrics," *Econometrica: Journal of The Econometric Society*, 46 (6), 1251-1271.
- Haynes, Kingsley E. and Alexander S. Fotheringham (1984), *Gravity and Spatial Interaction Models*, Beverly Hills, CA: Sage Publications.
- Helpman, Elhanan, Marc Melitz, and Yona Rubinstein (2008), "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123 (2), 441-487.
- Huff, David L. (1964), "Defining and Estimating a Trading Area," *Journal of Marketing*, 28 (3), 34-38.
- Mayo, Edward, Lance P. Jarvis, and James A. Xander (1988), "Beyond the Gravity Model," *Journal of the Academy of Marketing Science*, 16 (3), 23-29.
- Nakanishi, Masao and Lee G. Cooper (1974), "Parameter Estimation for a Multiplicative Competitive Interaction Model-Least Squares Approach," *Journal of Marketing Research*, 11 (3), 303-311.
- Reilly, William J. (1931), *The Law of Retail Gravitation*, New York: Knickerbocker Press.
- Silva, J. Santos and Silvana Tenreyro (2006), "The Log of Gravity," *The Review of Economics and Statistics*, 88 (4), 641-658.
- Silva, J. Santos and Silvana Tenreyro (2011), "Further Simulation Evidence on the Performance of the Poisson Pseudo-maximum Likelihood Estimator," *Economics Letters*, 112 (2), 220-222.
- Stanley, Thomas J. and Murphy A. Sewall (1976), "Image Inputs to a Probabilistic Model: Predicting Retail Potential," *Journal of Marketing*, 40 (3), 48-53.
- Timmermans, Harry (2001), "Spatial Choice Models," in *International Encyclopedia of the Social & Behavioral Sciences*, Neil J. Smelser and Paul B. Baltes, ed. Amsterdam: Elsevier, 114768-14771.

## Trade Area Analysis through Gravity Model

Juyoung Kim\*, Jeongsoo Kim\*\*

### ABSTRACT

The industry is changing rapidly using big data and AI, and the pandemic caused by the corona is accelerating these changes. In the distribution industry, non-face-to-face transactions such as delivery are rapidly increasing. In this situation, the consumer trade area based on the consumer's residence may become more important than the sales trade area based on the stores. In order to accurately identify changes in the distribution industry, both sales and consumer trade area will have to be considered at the same time.

This study tried to analyze the sales and consumer trade areas in an integrated way by using the gravity model. Furthermore, the sensitivity of each industry in the consumer and sales trade area was identified. We also analyzed the influence and sensitivity changes in each industry after the large mall was entered in the region.

Consumer and sales commercial transaction history were extracted using credit card panel data while keeping the personal information regulation law. We extracted 45 consumer trade areas and 38 sales trade areas from the data by k-means methodology by Sklearn, python. Five regression equations based on the gravity model were estimated for each industry, and the estimated summary of each regression equation is shown in <Table 1> below.

<Table 1> Summary of regression equations by industry

Industry	R (test sample)	R (hold-out sample)	R-square	Adj. R-Square	Degree of Freedom	F-value	p-value	DW (test sample)	DW (hold-out sample)
large-scale stores	0.233	0.199	0.054	0.051	(36, 11445)	18.300	0.000	1.009	1.169
Convenience store	0.400	0.396	0.160	0.159	(37, 373884)	1929.286	0.000	1.556	1.548
Restaurants	0.339	0.340	0.115	0.115	(37, 697500)	2450.943	0.000	1.687	1.624
Discount stores	0.313	0.320	0.102	0.102	(37, 236257)	723.270	0.000	1.412	1.415
Fashion Stores	0.260	0.261	0.067	0.067	(37, 71123)	138.894	0.000	1.469	1.510

\* Professor of Marketing at Sogang Business School (jkimsg@sogang.ac.kr), First Author

\*\* Doctoral Student of Marketing at Graduate School Sogang University (esko445@sogang.ac.kr), Corresponding Author

<Table 2> shows the differences according to the age of the consumer, the existing or new store, and the operating periods of each store. It can be said that the estimated parameters vary by industry. The constant term shows the average size of transactions for each industry, for example, large malls and fashion stores relatively large, and restaurants and discount stores in order.

<Table 2> Model estimate by industry

	large-scale stores	Convenience store	Restaurants	Discount stores	Fashion Stores
Constant*	14.520	11.205	13.157	12.819	13.899
$\alpha_{k1}$ (Under the age of 30)	-0.060	0.174	0.114	-0.369 <sup>^</sup>	-0.179
$\alpha_{k2}$ (Under the age of 45)	0.089	-0.067 <sup>^</sup>	-0.104 <sup>^</sup>	0.029 <sup>^</sup>	0.356
$\alpha_{k3}$ (Over the age of 45)	-0.062	-0.120	-0.038	0.292 <sup>^</sup>	-0.280
$\beta_{k1}$ (Existing Store)	-0.003	-0.011 <sup>^</sup>	0.001	-0.018	0.002
$\beta_{k2}$ (New Store)	0.010	-0.009 <sup>^</sup>	-0.035 <sup>^</sup>	0.023	0.020
$\delta_k(DUR, Time)$	0.006	0.055 <sup>^</sup>	-0.007 <sup>^</sup>	-0.098 <sup>^</sup>	0.001
$\gamma_{(ij)k}^{***}$	3/30	0/31	0/31	1/31	3/31

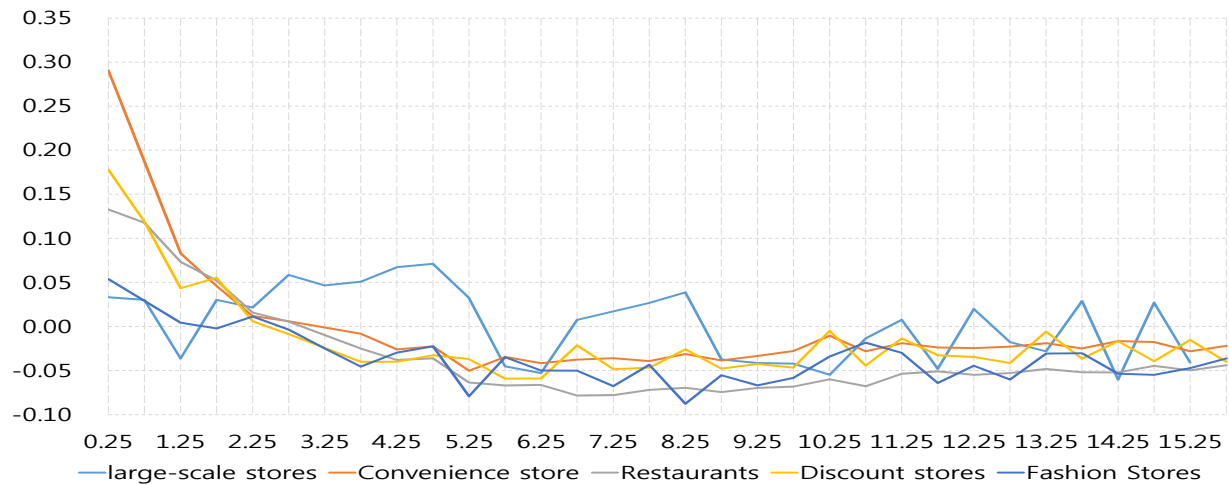
\* Only constants are non-standardized values, and all other parameters are standardized (beta values)

\*\* All parameters were statistically significant at the significance level of 0.01

\*\*\* Number of estimates of payment level 0.05 or more

<sup>^</sup> The estimated value that shows the differences in the level of 0.05 or more as a result of comparing the coefficients estimated by the holdout specimen and the model estimated sample

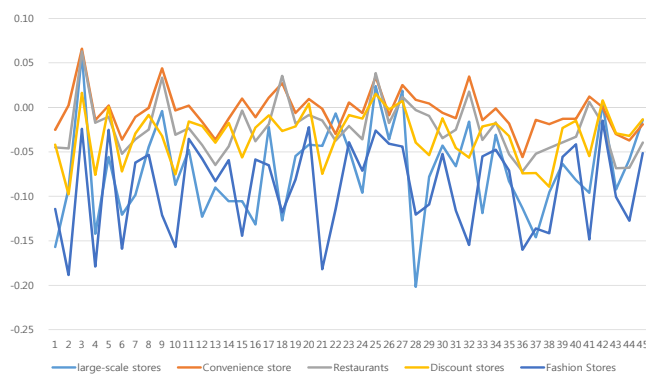
In the case of large-scale stores, the purchasing power of 30-45 year-old (0.089) is higher than that of other age groups, while at convenience stores, the younger age group, under 30 (0.174), have higher purchasing power. In restaurants, the purchasing power (0.114) of young people under the age of 30 is also high. As discount stores include supermarkets, it is found that those over 45 (0.292) have higher purchasing power than the younger generation. It can be seen that the purchasing power (0.356) of the 30-45 year old is also high in the fashion store.



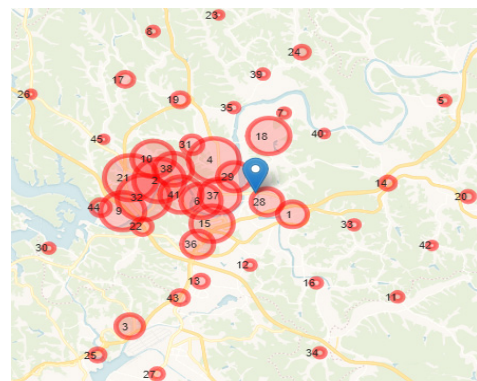
**<Figure 1> Effect of distance between sales and consumer trade area**

<Figure 1> shows that the closer they are in all industries, the more they purchase. Discount stores and restaurants are sensitive to distance. Fashion stores are less sensitive to purchases even if they moved farther away.

Below, we show the characteristics of consumer and sales trade area using the estimated parameters and the amount of effects that are calculated as parameters multiplied by size of transaction.

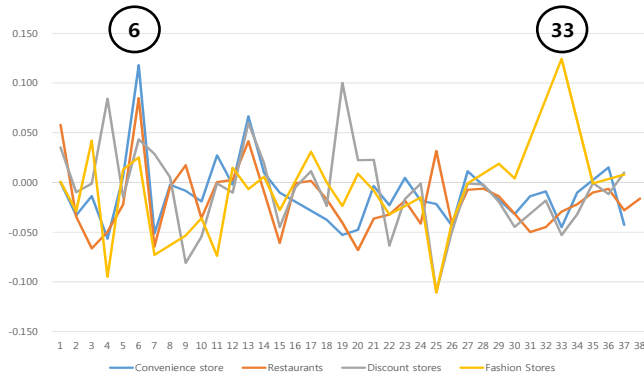


**<Figure 2> Estimated consumer trade area influence parameter**

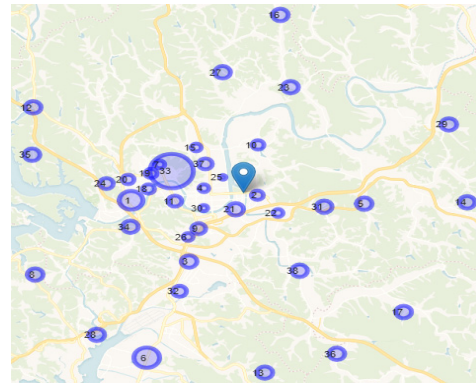


**<Figure 3> The influence of consumer markets (Parameters\*Trade Size)**

In <Figure 2>, the consumer trade area is shown on the horizontal axis and the size of the parameter by industry of the consumer trade area is on the vertical axis. In <Figure 3>, the magnitude of each consumer trade area influence is expressed as the size of the circle of the point where the consumer trade area is located.



<Figure 4> Estimated sales trade area influence parameter



<Figure 5> The influence of sales markets (Parameters\*Trade Size)

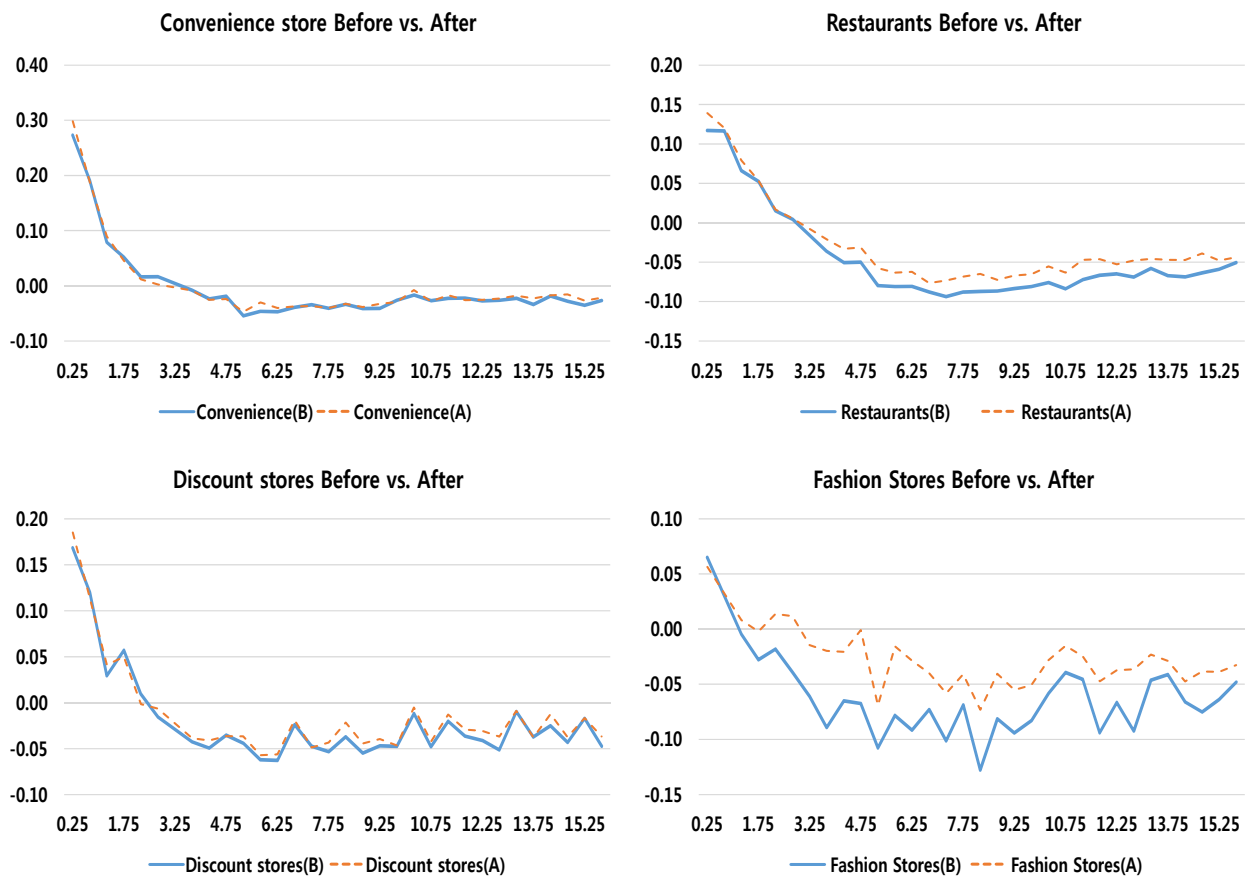
<Figure 4> and <Figure 5> are analysis results related to the sales trade area. This analysis also makes it easy to identify the differences in characteristics of sales trade area.

Under the current model structure, it is also possible to compare changes in commercial areas after large-scale stores enter. We can easily check how the influence of the consumer and sales trade area or the influence of the street changed after the large mall entered.

<Table 3> Influence parameters of commercial trade area before and after entering large-scale malls

	Before entering the large-scale mall				After entering the large-scale mall			
	Convenience store(B)	Restaurants (B)	Discount stores(B)	Fashion Stores(B)	Convenience store(A)	Restaurants (A)	Discount stores(A)	Fashion Stores(A)
$\alpha_{k1}$ (Under the age of 30)	0.213	0.136	-0.415	-0.435	0.164	0.107	-0.363	-0.003
$\alpha_{k2}$ (Under the age of 45)	-0.083	-0.130	-0.241	0.424	-0.065	-0.103	0.128	0.335
$\alpha_{k3}$ (Over the age of 45)	-0.148	-0.052	0.579	-0.125	-0.106	-0.022	0.194	-0.413
$\beta_{k1}$ (Existing Store)	0.007	0.006	-0.015	0.018	-0.049	-0.027	-0.023	-0.059
$\beta_{k2}$ (New Store)	-0.012	-0.006	0.004	-0.022	0.033	0.005	0.023	0.084
$\delta_k$ (DUR, Time)	0.085	0.014	-0.043	-0.006	0.026	0.010	-0.058	0.026

After entering a large-scale mall, the influence of new stores has increased compared to that of existing stores in sales businesses. In the case of discount stores, the influence of the 30-45-year-old consumer trade area has increased, which may be caused by the increase in the inflow population with new consumption capabilities.



<Figure 6> Effect of distance before and after entering a large-scale mall

<Figure 6> shows the impact of distance between consumer and sales trade area. After entering the large-scale mall, there is little change in convenience stores, but the sales amount increases for all stores no matter how they are away from consumer trade area in the case of fashion stores. Thus, we can identify the changes in the influence of local businesses due to the entry of large-scale malls.

Keywords: Trade Area, Sales Area, Consumer area, Trade Area Analysis, Gravity Model