

# 서울시 내 장기체류 외국인의 생활인구 밀집지역과 그 결정요인의 분석

---

- 생활인구 빅데이터를 활용하여 -

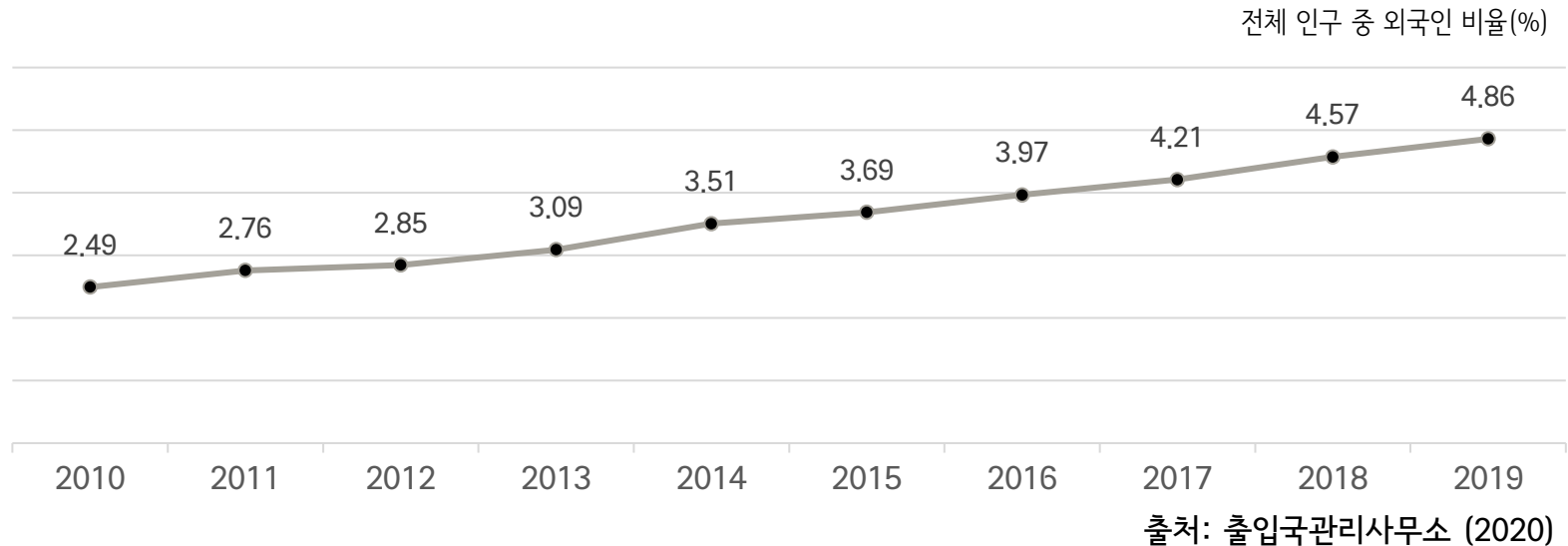
이준호(cosmicflush93@gmail.com), 홍영주(enshriner@gmail.com)

## 목차

- 1 서론
- 2 선행연구 및 기존연구와의 차별점
- 3 분석 I: 핫스팟 분석
- 4 분석 II: 로지스틱 회귀분석
- 5 분석 결과 및 제언

# 1. 서론

## ■ 대한민국의 현 위치는?



## ■ 왜 외국인 밀집지역에 대한 연구가 필요한가?

- 외국인의 정착 시에는 '밀집 지역', 혹은 분화(segregation)가 자주 일어남



출처: 박세훈 외 (2009)

# 1. 서론

- 외국인 밀집지역 관련 연구의 두 가지 도전 (Challenges)

- '외국인' 집단 내의 다양성
- 주민등록 인구 기반 연구의 한계

➡ 더 많은 국적을 고려한 연구, 생활인구를 고려한 연구

- 생활인구(de facto population)란?

$$\text{생활인구} = \text{주민등록인구(거주인구)} + (\text{유입인구} - \text{유출인구})$$

- 상주인구(de jure population)에 대치되는 개념
- A구에서 거주하는 사람의 직장이 B구라면?

## 연구 범위

### “서울시 내 장기체류 외국인”

- 상대적으로 긴 외국인 이민의 역사
- 높은 외국인 주민 비율
- 다양한 구득 가능 자료

- 90일 이상 체류 외국인 대상
- 단기체류 외국인 및 영주권자와 차이가 존재

- 시간적 범위: 2020년 1월 1일 - 2020년 5월 31일
  - 공간적 단위: 행정동 (제공된 자료와 동일)

## 2. 선행연구 - 외국인의 밀집에 대한 이론

- Robert E. Park(1950)의 동화이론 (Assimilation Theory)
  - 이주 초기에는 분화 상태에서 갈등을 일으키나 이후 주류 문화에 편입됨
- Basch et al.(1994)의 초국가적 사회공간론
  - 거주 외국인은 끊임없이 국가 경계를 넘나들며 초국가적 사회를 형성함
  - 이 과정에서 사회자본을 형성해 밀집을 이룸
- Faist(2000)의 사회자본론 (Social Capital Theory)
  - 정착 후에도 공동체 내에서 '네트워크'를 이룸
  - 이러한 사회자본은 입지특수적(location-specific)
- Sassen(2006)의 세계도시론
  - 소수민족 이주민들은 주류 사회가 기피하는 일을 부담하게 됨
  - 계층적 분화와 민족별 분리는 상호 강화되어 결국 사회 문제화됨

## 2. 선행연구 - 외국인의 밀집거주 관련 연구

### ■ 탐색적 조사

- 서울 시내 외국인 분포를 국적별로 분석  
(김혜진 외, 2018; 박세훈 외, 2010; 최재현 외, 2003)

### ■ 실증적 분석

- 비유사성 지수 이용해 군집분석 (이정아 외, 2013; 이진영 외, 2012; 하성규 외, 2011)
- 수도권 외국인 거주지 분포의 특성을 다중회귀분석을 통해 규명
  - 이진영 외(2012): 인종/국적별 연구, 외국인 지원센터, 대학교, 내국인 수 등에서 유의
  - 정지은 외(2011): 인종별 연구, 외국인 지원센터, 제조업 종사자 비율, 대중교통에서 유의

## 2. 기존 연구와의 차별점

- 생활인구 분포 기반 연구
  - 현재까지 연구는 거주인구 분포에 대한 연구만을 수행함
  - 거주인구와 생활인구 간의 간극이 가장 심한 평일 오전 9시~오후 6시의 외국인 분포에 대하여 연구
- 더 많은 국적 포함
  - 중국인, 일본인, 미국인의 분석에만 그친 연구 다수 존재
  - 대만인, 베트남인, 캐나다인을 추가하여 더욱 많은 자료 제공



### 3. 핫스팟 분석 (Hotspot Analysis)

- Incremental Spatial Autocorrelation

공간상관성이 극대화되는 지점을 선택: 6,000m 지점

	중국	미국	베트남	일본	대만	캐나다
First Peak	2947.46	5087.05	2709.73	3185.20	2079.73	5087.05
Max Peak	2947.46	6037.98	6988.91	5800.25	5324.78	6037.98

- Getis-Ord General G

공간상관성이 존재하는가? ○

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} x_i x_j}{\sum_{i=1}^n \sum_{j=1}^n x_i x_j}, \quad \forall j \neq i$$

- Getis-Ord Gi\*

공간상관성은 어디서 존재하는가?  
COVID-19에 의한 영향 고려

$$G = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{[n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2]}{n-1}}} \quad \bar{X} = \frac{\sum_{j=1}^n x_j}{n}$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2}$$

\* $x_i$ 와  $x_j$ 는  $i$ 와  $j$ 의 속성값 (Attribute Value)을,  $w_{i,j}$ 는  $i$ 와  $j$  사이의 공간가중치(Spatial Weight)를 나타냄

### 3. 핫스팟 분석 (Hotspot Analysis): 결과



미국인 핫스팟 지역



캐나다인 핫스팟 지역



일본인 핫스팟 지역

- 미국인과 캐나다인의 핫스팟 일치율이 매우 높음
  - 미국 기준 112개 중 101개(90.18%), 캐나다 기준 103개 중 101개(98.06%)
- 일본인의 경우 생활 환경이 기존 밀집지역으로 알려졌던 용산구에서 다른 구로 변해가고 있음

### 3. 핫스팟 분석 (Hotspot Analysis): 결과



중국인 핫스팟 지역



베트남인 핫스팟 지역



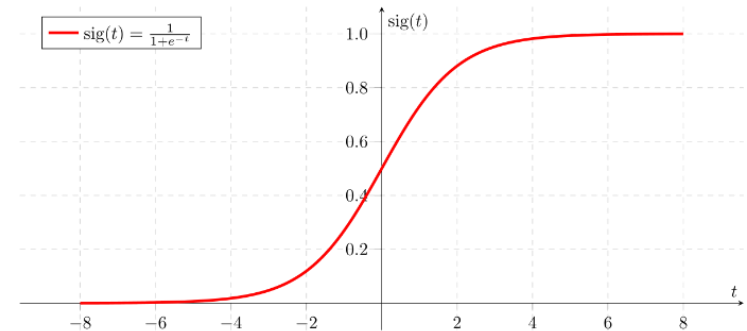
대만인 핫스팟 지역

- 중국인의 경우 원 밀집지역(영등포구, 금천구 등)을 중심으로 핫스팟이 드러남
- 베트남인에서는 고등교육기관이 밀집한 성북구, 동대문구 등과 단순 비전문인력·결혼 이민자 비율이 높은 중랑구 및 성동구에서 밀집 지역이 나타남
- 대만인은 전통적인 대만인 밀집지역 서대문구 뿐 아니라 이외 은평구, 종로구, 마포구 등에 서도 밀집지역이 드러남

# 4. 로지스틱 회귀분석

## ■ Logistic Regression

$$p = (Y = 1 | X = x_p) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}$$



## ■ 반응변수

- 행정동별 핫스팟 여부 (0과 1의 값을 가지는 이진변수)

## ■ 독립변수

- 행정동의 특성을 나타낼 수 있는 지표들

## 4. 로지스틱 회귀분석: 사용된 자료

- 서울생활인구 추정자료 (서울특별시, 2020)
  - 시공간적 자료
  - 6개국 외국인을 대상으로 평일 오전 9시~오후 6시 만을 추림
- 2019 서울서베이 (서울특별시, 2020)
  - 변수 설정의 방향성
  - 필요시 국적별 평균 이용
- 기타 행정동의 특성을 나타낼 수 있는 변수들
  - 총 17개
  - 2019 서울서베이 3번 문항(삶의 질 지표)과 선행연구 참고
  - 생활환경, 경제환경, 복지환경, 교육환경, 안전환경, 인구·사회학적 특성

## 4. 로지스틱 회귀분석: 사용된 변수

- 생활환경변수
  - 서울서베이 3번 문항의 '거주환경'을 연구에 맞게 수정
  - 대중교통 접근성(지하철 접근성과 버스 접근성), 시설 접근성(은행 접근성과 백화점 접근성)
  - 환경변수(미세먼지)
- 경제환경변수
  - '소득'과 '경제활동 기회' 관련 자료는 구득이 어려워 부득이하게 대리변수 사용
  - 공시지가 평균, 1km<sup>2</sup>당 신설법인 수, 1km<sup>2</sup>당 외국인 투자기업 수
- 복지환경변수
  - 복지/의료환경 관련 변수로 구성
  - 외국인 지원센터와 대사관 접근성, 종합병원 접근성과 1km<sup>2</sup>당 병/의원의 수
- 교육환경변수
  - 외국인학교 접근성과 대학 접근성
- 안전환경변수
  - 1km<sup>2</sup>당 교통사고 사망 건수와 화재 발생 건수
- 내국인 생활인구

\* 볼드 처리된 변수: 이전 연구에서 유의함이 증명된 변수

## 4. 로지스틱 회귀분석: 결과(1)

- 오즈비를 사용해 로지스틱 회귀분석의 결과 해석
  - 잠재변수를 이용한 회귀모형식은 방향만 알 수 있을 뿐 정도는 알기 힘들 (고길곤, 2017)
- 생활환경 변수
  - 지하철 접근성: 중국인(0.701\*\*), 대만인 · 일본인(1.227\*\*, 1.311\*\*)
  - 버스 접근성: 중국인 · 대만인(1.589\*\*, 1.399\*)
  - 은행과 백화점 접근성: 통계적 유의성 드러나지 않음
  - 미세먼지: 베트남인(0.995\*), 현재로서는 외적 타당성을 검증할 수 있는 자료 없음.
- 경제환경 변수
  - 공시지가 평균: 중국인(0.998\*), 미국인 · 캐나다인(1.008\*\*, 1.008\*\*)
  - 1km<sup>2</sup>당 신설 법인: 통계적 유의성 드러나지 않음.
  - 외국인 투자 기업 수: 중국인(0.742\*), 미국인 · 캐나다인(1.887\*\*, 1.589\*\*)

주) \*\*, \*: 각각 통계적 유의수준 1%, 5%에서 유의함.

## 4. 로지스틱 회귀분석: 결과(2)

### ■ 복지환경 변수

- 외국인 지원센터: 중국인 · 베트남인(2.341\*, 2.323\*)
- 종합병원 접근성: 통계적 유의성 드러나지 않음
- 1km<sup>2</sup>당 병/의원 수: 중국인 · 대만인(0.946\*, 0.943\*)
  - 의료 수요가 낮아서가 아니라 의료 기관 이용에 어려움을 느끼기 때문

### ■ 교육환경 변수

- 외국인학교 접근성: 미국인/캐나다인(4.549\*\*, 5.281\*\*)
- 대학 접근성: 중국인(0.250\*\*)대만인 · 일본인 · 캐나다인 · 베트남인(3.479\*\*, 3.783\*\*, 2.242\*, 7.310\*\*)
  - 중국인의 해석 상에는 주의 필요

### ■ 안전환경 변수

- 통계적 유의성 드러나지 않음

### ■ 인구/사회학적 변수

- 내국인 생활인구: 베트남인 · 일본인(0.925\*\*, 0.956\*\*)



## 5. 결론 - 정책적 함의

### ■ ‘동화이론’과 ‘세계도시론’ 관점의 해석

- 이주 초기 ‘분화 상태의 갈등’은 중국인 분석에서 여전히 드러남
- 중국인과 베트남인 일부의 경우 ‘세계도시론’을 지지하는 결과를 보임
- 미국인, 캐나다인은 고부가가치 산업에 종사하며 ‘세계도시의 다른 극단’을 보여줌

### ■ 중화권 장기체류 외국인을 대상으로 한 의료접근성 확충 필요

- 중국인과 대만인의 의료 서비스 수요가 낮아서 1km<sup>2</sup>당 병/의원 수가 핫스팟 형성에 부의 영향을 미친 것이 아님
- 의료 기관 이용에 어려움을 겪고 있음이 <2019 서울서베이>로 증명됨
- 의사소통이 원활한 병/의원의 확충이 필요

### ■ 외국인 학교의 접근성 확충

- 미국인과 캐나다인을 제외하면 외국인학교는 생활인구 핫스팟에 영향을 주지 못함
- 영어를 주 언어로 하는 외국인 학교의 비율은 57.9%
- 영미권에서 온 취학 연령의 아동을 제외한 다른 아동에 대한 교육 선택의 폭 확충이 필요

## 5. 결론 - 연구의 한계점

- 오전 9시부터 오후 6시까지의 평균만을 고려
  - 야간의 생활인구 변화는 고려하지 못함
- 주어진 자료 내에서만 분석
  - 월별 및 계절별 해석을 하지 못해 의미 확장을 하기 힘들
- 거주인구와 생활인구 간 차이 고려하지 못함, 행정동별 자료를 기반으로 함
  - 이는 거주인구 자료/집계구별 자료 구득이 어려웠기 때문
- 생활인구 자료에서 중국인의 구분이 모호함
  - 한국계 중국인과 중국계 중국인 간 구분이 되어있지 않음
- 외국인 사이의 네트워크 관련 검증이 부족함
  - 향후 질적 연구를 바탕으로 사회 자본과 외국인 간 네트워크에 대한 연구를 보충할 수 있을 것임

## Q & A

# 부록. 핫스팟 분석 결과

	중국	베트남	미국	일본	캐나다	대만
First Peak	2947.46	2709.73	5087.05	3185.20	5087.05	2079.73
Max Peak	2947.46	6988.91	6037.98	5800.25	6037.98	5324.78

## Incremental Spatial Autocorrelation 분석결과

\* 중국은 6,000m 부근에서 또 하나의 Peak를 보임

	중국	베트남	미국	일본	캐나다	대만
observed G	0.000086	0.000093	0.000094	0.000086	0.000089	0.000146
expected G	0.000057	0.000057	0.000057	0.000057	0.000057	0.000057
variance	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
z-score	12.259564	12.944031	12.536440	7.516383	11.306551	15.418297
p-value	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

## Getis-Ord General G 분석결과 (Threshold Distance = 6,000m)

# 부록. 변수 목록과 출처

생활 환경	지하철접근성	반경 350m 내의 지하철역 개수	
	버스 접근성	행정동을 통과하는 서울 시내 버스의 개수	서울특별시 도시교통실, 2020
	은행 접근성	행정동 내 무인점포를 제외한 금융기관 수	은행연합회, 2020
	백화점 접근성	반경 1km 내에 백화점이 있는 지 여부 (다미 변수)	서울특별시, 2020
	미세먼지( $\mu\text{g}/\text{m}^3$ )	2020년 1월~5월 09시~18시 행정동별 보정된 미세먼지 수치의 평균	서울연구원, 2020
경제 환경	공시지가 평균	행정동별 주택 공시지가의 평균	국토교통부, 2020
	1km <sup>2</sup> 당 산설법인수	행정동 내 1km <sup>2</sup> 당 행정동 내 산설법인수	중소기업벤처부, 2019
	1km <sup>2</sup> 당 외국인 투자 기업 수	행정동 내 1km <sup>2</sup> 당 외국인 투자 기업 수	산업통상자원부, 2019
복지 환경	외국인 자원센터 접근성	반경 1km 내에 외국인 자원센터가 있는 지 여부	서울글로벌센터, 2019
	종합병원 접근성	반경 1km 내에 종합병원이 있는 지 여부	서울특별시, 2020
	1km <sup>2</sup> 당 병의원 수	행정동 내 1km <sup>2</sup> 당 종합병원을 제외한 모든 병원 수	서울특별시, 2020
	대사관 접근성	반경 1km 내에 각국에 대응되는 대사관이 있는 지 여부 (예: 중국인에 대한 모형에서는 중국 대사관 반경 1km 여부)	

교육 환경	외국인 학교 접근성	반경 1km 내 외국인 학교의 존재 여부	서울특별시 교육청, 2016
	대학 접근성	반경 1km 내 대학 존재 여부	서울특별시, 2016
안전 환경	1km <sup>2</sup> 당 교통사고 사망 건수	행정동 내 1km <sup>2</sup> 당 사망자가 발생한 교통사고 건수	도로교통공단 통합DB처, 2019
	1km <sup>2</sup> 당 화재 발생 건수	행정동 내 1km <sup>2</sup> 당 화재 발생 건수	서울특별시 소방재난본부, 2019
인구/ 사회학적 특성	내국인 생활 인구	2020년 1~5월 09시~18시 행정동별 내국인의 생활 인구 평균	서울특별시, 2020

# 부록. 로지스틱 회귀분석: 기술통계량과 분석결과

	변수	전체행정동		중국인생활인구 하스팻		미국인생활인구 하스팻		베트남인생활인구 하스팻		일본인생활인구 하스팻		대만인생활인구하스팻		캐나다인생활인구 하스팻	
		평균	표준편 차	평균	표준편 차	평균	표준편 차	평균	표준편 차	평균	표준편차	평균	표준편차	평균	표준편차
생활 환경	지하철접근성	2.745	2.289	1.842	1.797	3.973	2.865	2.462	1.950	3.904	3.168	3.776	3.202	3.854	2.709
	버스접근성	1.970	1.460	2.305	1.268	2.116	1.654	3.444	2.851	2.610	2.023	2.400	1.885	2.052	1.601
	은행접근성	0.441	0.547	0.399	0.313	0.735	0.882	0.571	0.810	0.659	0.928	0.588	0.948	0.725	0.883
	백화점접근성	0.066	0.249	0.079	0.271	0.116	0.322	0.060	0.238	0.084	0.280	0.079	0.271	0.136	0.344
	미세먼지( $\mu\text{g}/\text{m}$ )	28.873	12.334	29.195	11.910	27.000	13.101	27.781	12.545	29.849	12.739	30.858	13.425	27.358	12.889
경제 환경	공시지가평균(백만원)	443.958	250.199	376.488	135.923	658.660	344.222	411.776	147.569	423.200	147.008	406.651	132.508	680.753	350.981
	1km당 신설법인수	2.490	4.452	2.620	6.327	4.408	5.749	2.328	3.005	2.823	2.988	2.609	3.133	4.496	5.929
	1km당 외국인투자기업	1.295	3.105	0.878	2.053	3.351	5.114	1.856	3.701	2.413	4.243	2.163	4.284	3.380	5.198
복지 환경	외국인지원센터접근성	0.134	0.342	0.184	0.390	0.223	0.418	0.188	0.392	0.205	0.406	0.224	0.419	0.233	0.425
	종합병원접근성	0.125	0.331	0.184	0.390	0.098	0.299	0.103	0.305	0.072	0.261	0.118	0.325	0.097	0.298
	1km당 병의원수	11.968	9.325	11.903	7.542	13.897	12.100	12.901	8.633	13.413	8.915	11.388	8.517	13.645	12.366
	대사관접근성	-	-	0.118	0.325	0.098	0.299	0.026	0.159	0.048	0.215	0.053	0.225	0.107	0.310
교육 환경	외국인학교접근성	0.090	0.286	0.053	0.225	0.196	0.399	0.068	0.253	0.108	0.313	0.066	0.250	0.204	0.405
	대학접근성	0.387	0.488	0.197	0.401	0.402	0.492	0.735	0.443	0.699	0.462	0.697	0.462	0.369	0.485
안전 환경	1km당 교통사고 사망건수	0.055	0.097	0.065	0.095	0.047	0.082	0.065	0.113	0.054	0.098	0.057	0.095	0.046	0.081
	1km당 화재발생건수	1.381	0.950	1.519	0.933	1.405	0.965	1.606	1.072	1.646	1.112	1.381	0.847	1.312	0.828
인구 특성	내국인생활인구(천명)	25.762	15.448	25.782	17.926	31.225	20.959	22.221	12.687	24.279	16.760	27.783	20.130	31.443	20.667
관측치수		424		76		112		117		83		76		103	

	변수	중국인생활인구하트스팟		미국인생활인구하트스팟		베트남인생활인구하트스팟		일본인생활인구하트스팟		대만인생활인구하트스팟		캐나다인생활인구하트스팟	
		Oddsratio	p-value	Oddsratio	p-value	Oddsratio	p-value	Oddsratio	p-value	Oddsratio	p-value	Oddsratio	p-value
생활 환경	지하철 접근성	0.701**	0.000	1.134	0.094	1.056	0.640	1.311**	0.001	1.227**	0.004	1.050	0.560
	버스 접근성	1.589**	0.001	0.947	0.704	1.262**	0.004	1.217	0.134	1.399*	0.011	1.027	0.850
	은행 접근성	1.066	0.904	1.181	0.761	2.469	0.092	1.647	0.313	0.963	0.921	0.740	0.591
	백화점 접근성	3.103	0.069	0.294	0.129	1.280	0.714	1.531	0.531	0.749	0.600	0.950	0.934
	미세먼지( $\mu\text{g}/\text{m}^3$ )	1.001	0.720	0.999	0.651	0.995*	0.039	1.015	0.172	1.016	0.188	0.999	0.801
경제 환경	공시지가평균(백만원)	0.998*	0.016	1.008**	0.000	1.000	0.833	0.999	0.377	0.999	0.065	1.008**	0.000
	1km당 신설법인수	1.107	0.053	0.901	0.103	0.920	0.305	0.975	0.561	0.957	0.275	0.969	0.594
	1km당 외국인투자기업	0.742*	0.017	1.887**	0.000	1.175	0.109	1.161	0.056	1.117	0.094	1.589**	0.001
복지 환경	외국인자원센터 접근성	2.431*	0.028	1.063	0.875	2.323*	0.027	1.466	0.303	1.781	0.117	1.336	0.475
	종합병원 접근성	2.049	0.059	0.490	0.204	0.738	0.469	0.383	0.056	0.623	0.274	0.488	0.188
	1km당 병/의원수	0.946*	0.022	1.013	0.561	0.973	0.267	0.969	0.168	0.943*	0.018	1.015	0.469
	대사관 접근성	1.186	0.722	0.516	0.237	0.278*	0.021	0.593	0.362	-	-	0.810	0.704
교육 환경	외국인학교 접근성	0.624	0.430	4.549**	0.000	0.667	0.446	1.646	0.290	0.670	0.441	5.281**	0.000
	대학 접근성	0.250**	0.001	2.353*	0.015	7.310**	0.000	3.783**	0.000	3.479**	0.000	2.242*	0.019
안전 환경	1km당 교통사고 사망건수	1.399	0.823	0.037	0.173	1.994	0.601	0.167	0.271	0.735	0.835	0.029	0.120
	1km당 화재발생건수	1.316	0.128	0.819	0.361	1.244	0.209	1.086	0.631	0.734	0.092	0.738	0.190
인구 특성	내국인생활인구(천명)	1.029	0.057	0.978	0.184	0.925**	0.000	0.956**	0.005	1.005	0.718	0.976	0.148
상수항		0.299*	0.030	0.006**	0.000	0.343*	0.033	0.090**	0.000	0.083**	0.000	0.006**	0.000
관측치수		423		423		423		423		422		423	
Wald Chi2(P>chi2)		57.22(0.0000)		79.39(0.0000)		86.77(0.0000)		66.31(0.0000)		70.80(0.0000)		73.03(0.0000)	
Cragg&Uhler'sR2		0.265		0.587		0.415		0.314		0.251		0.574	
McKelveyandZavoian'sR2		0.410		0.712		0.511		0.423		0.353		0.679	