

제2권 제1호

ISSN 2508-2183

연구방법논총

Journal of Research Methodology

2017년 봄호

차 례

【연구논문】

SNS활용과 정치참여: 한국 제20대 총선 분석 박영득·이재묵 / 1

국제이주 연구에 있어 정치학적 접근과 방법론적 쟁점 이병하 / 23

SNS 활용과 선거캠페인 효과 장우영·이관률 / 53

【연구경향】

An Introduction to Mixed Logit Model Byung-Jae Lee/ 79

『연구방법논총』 원고작성의 일반적 요령 107

『연구방법논총』 저술 윤리강령 116

『연구방법논총』 편집 및 심사 규정 119

『연구방법논총』 편집위원명단 124

[DOI] <http://dx.doi.org/10.21487/jrm.2017.05.2.1.1>

【연구논문】

SNS활용과 정치참여: 한국 제20대 총선 분석

박영득* · 이재묵**

논문요약

본 논문은 SNS 활용이 정치참여에 어떠한 영향을 미치는지를 한국 제20대 총선 자료를 분석하여 규명한다. 선행연구들은 인터넷, SNS와 같은 뉴미디어가 정치참여에 필요한 정치정보 취득 비용을 감소시킴으로써 정치참여를 촉진한다는 것을 보여주었으나 본 연구의 분석결과는 SNS 활용과 투표참여가 유의미한 관계에 있지 않다는 것을 보여준다. 그러나 SNS 활용은 비선거적 정치참여에 대해서는 유의미한 영향력을 미치고 있으며 이는 SNS가 제공하는 사회연결망과 사회자본에 힘입어 시민들이 스스로를 조직화 할 수 있기 때문인 것으로 보인다.

주제어: SNS, 정치참여, 투표참여, 비선거 정치참여, 미디어 효과

* 연세대학교

** 한국외국어대학교

I. 서론

정치참여에 대한 고전적 접근에 따르면 유권자들이 정치에 자신의 목소리를 내는 행위는 필연적으로 일단의 비용을 수반한다(자원론적 접근으로는 Verba et al. 1972를, 합리적 선택 접근의 입장에서는 Riker and Ordeshook 1968을 참조할 수 있다). 특히 선거를 중심으로 생각해보면 유권자들은 정당이나 후보자가 어떠한 공약을 내세우고 있는지를 알기 위해서 노력을 쏟아야 하고 투표장으로 가기 위해 시간을 들여야 하며, 자신의 정치적 의견을 형성하고 그것이 타인에게 효과적으로 알려지도록 하기 위해서 일련의 행위들을 해야만 한다. 요컨대 정보에 대한 접근과 의견의 전파에 대한 인지적, 물리적 제약들은 유권자들의 정보취득 비용을 증가시키며 결과적으로 행위의 비용 또한 동시에 증가시키는 것이다. 그러나 21세기를 살아가는 유권자들의 삶 속에서 가장 큰 변화라고 할 수 있는 정보통신기술(ICTs: Information Communication Technologies)의 발달은 이러한 정보관련 비용들을 현저히 감소시킴으로써 시민들의 정치참여에 필요한 제반 비용을 감소시킬 것이라는 전망을 가질 수 있게 했다.

소규모 지인들 간의 관계망을 형성하기 위해 시작된 소셜 네트워크 사이트(이하 SNS: Social Networking Sites)의 사용자가 급증함과 동시에 규모가 광범위해지면서 현대인의 삶에 한 부분으로 자리매김한 지금, 인터넷 연구 초기의 기대들은 시민들의 정치적 삶의 다양한 부분에 주목할 필요가 있다. 민주주의 국가에서 시민들이 정치에 참여하는 방법은 오직 투표만으로 국한되어 있지 않고 시위에 참여하는 등 집단적이고 더욱 강도 높은 행동 또한 가능하기 때문이다. 인터넷 매체가 일반적으로 정보비용의 감소를 통해 참여의 비용을 줄이고, 그에 따라 정치참여를 자극할 수 있다는 설명논리가 SNS의 활용이 정치참여에 미치는 영향을 논하면서도 동일하게 적용될 수 있는 것인지 이론적으로 숙고할 필요가 있다.

이에 본 연구는 한국의 20대 총선에서 SNS의 이용이 정치참여에 대해 어떠한 영향을 미쳤는지를 분석한다. 구체적으로 정치참여를 선거참여(electoral participation)와 비선거 참여(non-electoral participation)로 나누어 SNS가 두

정치참여의 유형에 대해 각각 다른 영향력을 미치는지를 살펴본다.

II. 이론적 논의

1. 선행연구 검토: SNS와 정치참여

매체가 개인의 정치적 행위에 어떠한 영향을 어떻게 미칠 것인지는 각 매체가 어떠한 구조를 가지고 있는지에 따라 매우 상이하다. 즉 동일한 내용의 콘텐츠도 어떠한 구조를 가진 매체에서 공유되고 소비되느냐에 따라 그 영향은 달라진다(McLuhan 1994). 예컨대 포털 사이트나 인터넷 뉴스 서비스 등 상대적으로 오래된 인터넷 매체와 SNS와 같은 상대적으로 최근의 인터넷 매체가 상이한 의사소통의 구조를 제공하고 있다면, 그 매체를 이용하는 행위의 정치적 효과 또한 동일하지는 않을 것이라고 기대할 수 있다는 것이다. 예컨대 박근영(2014)은 ‘카카오톡’과 ‘페이스북’ 이용이 지지후보 변경에 미치는 영향이 다르다는 것을 발견했으며 이는 두 매체가 서로 다른 의사소통의 구조를 가지고 있기 때문이다.

SNS가 매체로서 가지는 특성 중에서 특히 여타 인터넷 매체와 구분되는 것은 개인을 기초단위로 하는 소통구조를 가지고 있다는 것이다. 이는 말 그대로 시민들이 스스로를 네트워크로 조직화할 수 있는 가능성을 제공한다. 인터넷 커뮤니티와 같은 경우에도 기초 단위는 개인이라고 할 수 있으나 이들은 인터넷 상에서 익명화된 존재이지만 SNS에서 이루어지는 상호작용은 상당부분 개인의 프로필을 공개한 상태에서 이루어진다는 점에서 네트워크의 강도가 더욱 강하다. 요컨대 SNS 이용자는 정치정보를 더 손쉽게 취득하는 것뿐만 아니라 자신이 접한 어떤 의견이 구체적으로 누구의 의견인지, 그리고 그에 찬성 또는 반대하는 사람이 누구인지를 더욱 명확하게 알 수 있다. 그렇기 때문에 SNS 이용자들은 네트워크 안에 속해있는 타인들과 그들의 의견을 더욱 긴밀하게 지각하게 되며 이는 더 높은 수준의 상호 신뢰와 연대감을 형성하는데 도움을 준다.

SNS 매체의 특성은 특히 비제도적 정치참여, 달리 말하면 비선거적 정치참여에 대해 동원효과를 만들어낸다. 2011년 중동의 시민혁명은 SNS의 동시적인 의견교환과 그 광범위성에 의해 증폭될 수 있었고(조희정 2011), 유사한 맥락에서 케냐(Kenya)에서의 비제도적 정치참여를 원활하게 하는데 도움을 주고 있다(Ndavula and Mberia 2012). SNS의 정치적 동원효과는 권위주의 국가나 부분적 민주주의 국가(partial democracies) 뿐만 아니라 선진 민주주의 국가(advanced democracies)에서도 나타난다. 인터넷 매체 이용자와 SNS 이용자들의 정치참여를 비교하는 연구전략은 SNS의 동원효과를 더욱 뚜렷하게 드러내는데, 미국 SNS의 사용자와 비사용자의 정치참여를 비교분석한 연구는 인터넷 사용자들 중 SNS 이용자 집단이 SNS를 이용하지 않는 집단보다 비선거적 정치참여, 즉 항의형 참여에 더욱 적극적으로 나서고 있다는 것을 경험적으로 보여주었다(송경재 2014). 또한 SNS 이용자들은 일반 인터넷 이용자에 비해 정치정보를 더욱 적극적으로 습득하려함과 동시에 온라인 정치 토론 참여, 제도적, 비제도적 정치참여에 더욱 적극적으로 나선다는 사실 또한 제시되고 있다(송경재 2011).

2. 정치참여의 유형과 SNS

민주주의 국가에서 유권자가 정치과정에 영향을 미치는 가장 대표적이고 강력한 방법은 투표에 참여하는 것이다. 그러나 민주주의에서 시민들에게 허락된 정치적 권리는 단지 투표장에서만 국한되는 것이 아니다. 개별 시민들은 동료시민들을 설득하고 스스로를 조직할 수 있다. 정부는 물론, 기업 등 비정부적 조직에 대해서도 각자의 불만족에 대해 항의한다. 요컨대 정치참여에 대한 고전적 연구들이 분류하듯 정치참여는 투표참여를 그 대표로 하는 선거 참여(electoral participation)와 불매운동과 같이 사회운동에 참여하거나, 정부에 항의하는 시위에 참여할 수도 있고 정당이나 시민단체와 같은 조직에 가입함으로써 정치과정에 영향력을 미치는 비선거적 참여(non-electoral participation)로 나뉜다.

각각의 정치참여 유형에 대해 SNS는 어떠한 영향력을 가지고 있는지 이론적으로 생각해볼 필요가 있다. 왜냐하면 투표참여와 비선거적 참여는 단지 ‘무엇

을’ 하는지만 다른 것이 아니라 ‘어떻게’ 하느냐도 다르기 때문이다. 투표참여의 경우 어디까지나 개인적인 행위다. 다시 말해 개인이 투표하기 위해 반드시 여러 사람들을 조직해야 할 이유는 없다. 반면 비선거적 참여, 예컨대 시위나 사회운동의 경우 집단적인 행위로, 함께 참여할 사람들을 필요로 한다. 또한 투표참여는 비선거적 참여에 비하여 훨씬 규범성이 강하다. 민주주의 사회에서 선거의 중요성은 모든 구성원들에게 충분히 주지되어 있으며, 투표에 참여하는 것이 바람직하다는 인식은 널리 공유되고 있다. 마지막으로 비선거적 참여는 투표에 비해 훨씬 높은 비용을 필요로 하는 행위다. 예컨대 시위 등은 투표에 비해 훨씬 더 많은 시간을 할애해야 하며 시위 현장에서 발생할 수도 있는 예견치 못한 일이 발생할 수도 있다는 것 또한 감수해야 한다.

인터넷이나 SNS와 같은 뉴미디어(new media)의 활용이 정치참여에 영향을 미칠 것이라는 이론적 기대는 서론에서 언급한 것처럼 뉴미디어가 전통적 대중매체에 비해 훨씬 저렴한 정보비용으로 공공문제에 관한 정보에 접근 가능하게 해준다는 점에 기반하고 있다. 그러나 보통의 인터넷 매체가 정보에 대한 접근 및 유통에서 그 특징을 가지고 있다면 SNS는 관계적인 매체라는 특징을 갖는다는 점에서 다른 인터넷 매체들과 구분될 수 있다. 즉 SNS는 다른 인터넷 매체와 마찬가지로 정보의 공유와 확산이 가능함과 동시에, 사회적 관계를 지향하는 목적으로 이용되고 있다는 점은 특징적이다. 또한 일반적인 인터넷 매체가 비대면 관계에 있는 사람들을 중심으로 약한 연대를 형성한다면 SNS는 대면관계가 형성된 사람들 사이에서의 강한 연대를 형성할 수 있다.

SNS의 매체로서의 속성은 특히 비선거 정치참여 행위를 자극하는데 크게 기여할 수 있다. 이러한 가능성은 2011년 중동지역에서 나타난 중동 시민혁명에서 확인되었다(조희정 2011).¹⁾ SNS를 통해 시민들은 자신들의 지인이 사회적 문제에 대해 어떠한 생각을 가지고 있는지를 서로 확인할 수 있고 의견을 교환할 수 있다. 이러한 요소는 전통적 대중매체는 물론, 온라인 커뮤니티, 포털사이트 등

1) 조희정(2011)은 SNS가 중동 시민혁명을 가능케 한 ‘조건’은 아니지만 혁명을 ‘가속화’할 수 있는 매개체로서의 역할을 했다고 분석했다. 이에 따르면 SNS가 사회변동에 직접적인 효과를 가지는 것은 아니지만 시민들의 사회적 관계망이 형성·확장되고 여론을 형성하는 매개적 효과를 수행했다고 볼 수 있다.

여타 인터넷 매체와 구분되는 특성이다. SNS를 제외한 여타 인터넷 매체도 시민들 간 관계망을 형성하고 정보를 공유하며, 의견을 교환하는 장이 되기는 하지만 개별 사용자들 간의 관계의 긴밀도는 SNS가 훨씬 강하다. 개인으로서 행하는 투표와 달리 집단적으로 참여하는 시위와 같은 정치참여에서 보다 긴밀한 관계에 있는 동료 시민의 메시지의 동원효과는 더욱 두드러질 수 있다는 것이다.

송경재, 임정빈, 장우영이 강조하는 SNS에서의 커뮤니케이션의 몇 가지 특징 중 가장 눈여겨 볼만한 특징은 SNS는 정보의 수용자가 자발적으로 참여하는 개방적 커뮤니케이션이 이루어지는 공간이라는 것이다. 이들은 SNS가 개인 기반의 사회적 연결망을 제공하기 때문에 자발성이 더욱 뚜렷하게 나타난다고 본다. 또한 SNS에서 정보의 유통을 담당하는 것이 개별 이용자들이라는 사실에 주목할 필요가 있다(송경재 외 2016, 157). 즉 시민들 간 자발적인 조직화와 의제의 확산, 동의 등을 필요로 하는 비선거 정치참여는 SNS 상에서 이루어지는 커뮤니케이션의 특성이 정치에 가져다주는 효과에 의해 활성화 될 수 있다. 예컨대 시위에 참여하려는 동기를 가지고 있다 하더라도 그러한 동기를 실행으로 옮기 위해 함께 참여할 동료 시민을 찾고, 자신과 같이 참여하고자 하는 동료 시민들이 자신의 주변에도 있다는 것을 SNS를 통해 지속적으로 확인함으로써 시위참가에 관여되는 여러 심리적 장벽들을 제거할 수 있다는 것이다.

집합자료를 사용한 국가 간 비교연구에 따르면 인터넷 매체의 이용이 투표참여에는 유의미한 영향력을 미치지 않지만, 비선거적 참여(보이콧 운동 등)에는 유의미한 영향력을 표출하고 있다는 것을 보여준 바 있다(정재관 2013). 정재관에 따르면 인터넷 매체 활용이 비선거적 참여에 직접적인 영향력은 표출하고 있지 않지만, 인터넷 이용과 사회자본(social capital)의 형성과는 긴밀하게 연결되어 있으며 이러한 영향력이 간접적으로 비선거적 참여로 이어지고 있다(정재관 2013 155).

SNS가 가지는 매체로서의 특징, 즉 SNS 상에서의 커뮤니케이션의 자발성과 관계성은 시민들을 비선거적인 정치적 행동에 더욱 쉽게 참여할 수 있도록 돕는다. 공공정보를 얻는 도구로서 SNS를 많이 활용한다는 것은 정치정보를 자신의 지인들이 자발적으로 공유하는 정보를 정보와 그들의 의견을 접할 수 있다는 것

을 의미하기 때문이다. 또한 그러한 관계적인 커뮤니케이션은 다소 높은 관여도를 가지는 비선거적 참여를 더욱 쉽게 만들 것이다. 그러나 투표참여의 경우 SNS가 특별히 기여할 수 있는 부분은 크지 않다. 다시 말해 시민들이 광범위하게 자발적 조직화를 이루고 연대감을 형성하며, 그로 인해 운동에 참여하는 것은 SNS를 통해서 ‘약간의 도움’을 받는 정도가 아니라 그로 인해 ‘가능한’ 정도의 기여를 보인다고 하겠다. 그러나 SNS가 투표참여의 비용을 현저하게 감소시킨다든지 아니면 효용을 매우 크게 한다는 이론적 기대를 하기는 어렵다.

III. 분석모형 설계

1. 변인과 데이터, 방법

본 연구는 위에서 논의한 내용에 따라 SNS 이용이 정치참여, 보다 자세히는 투표참여와 비선거 정치참여에 각각 미친 영향을 추정한다. 이를 위해 지난 20대 국회의원 선거에서 1,215명의 유권자(만 19세 이상 성인유권자)를 대상으로 수행한 여론조사(조사주관: 한국정치학회) 자료를 사용하여 통계분석 한다. 이 설문조사의 표본은 지역, 성, 연령을 기준으로 하여 비례할당추출(proportional quota sampling) 방식으로 표집되었고 일대일 대면조사 방식으로 응답을 획득하였다.

종속변인

투표참여는 투표참여 여부로 측정하였다. ‘귀하는 이번 20대 국회의원 선거에서 투표하셨습니까?’라는 질문에 ‘투표했다’와 ‘사전투표했다’고 응답한 유권자를 투표자(1)로, 이외의 유효 응답을 한 유권자들은 기권자(0)로 부호화 했다.

비선거적 참여는 크게 사회운동(불매운동 등)에 참여, 시위(촛불시위 등)에 참여, 정당이나 시민단체에 가입하는 행위로 측정했다. 각각의 정치적 행동을 해본 적이 있다고 응답한 경우를 1로, 그렇지 않은 경우를 0으로 부호화했다. 세 부

류의 비선거적 참여에 대한 효과를 각각 살펴볼 수도 있지만 다소 포괄적으로 단 한 종류의 비선거적 정치참여를 해본 경험에 대한 효과를 추정해보는 것도 의미가 있을 것이다. 그러므로 사회운동, 시위, 정당 또는 사회단체 가입 중 단 하나라도 경험이 있다고 응답한 경우를 1로, 그렇지 않은 경우를 0으로 하였다. 본 연구는 사회운동, 시위, 정당 또는 사회단체 가입에 대한 SNS 이용빈도의 효과를 각각 살펴봄과 동시에 더욱 포괄적인 의미에서 비선거적 참여에 미치는 영향 또한 살펴볼 것이다.

독립변인

본 연구에서 주목하는 SNS의 활용은 공적 문제에 대한 정보를 획득하는 경로로서 SNS의 활용이다. 그러므로 ‘귀하께서는 공공문제에 대한 정보를 얻기 위해 다음과 같은 매체를 얼마나 자주 이용하십니까?’라는 질문에서 SNS(트위터, 페이스북, 인스타그램 등)에 대한 응답을 SNS 이용빈도를 측정하는데 활용했다. 구체적인 응답의 범주는 ‘이용하지 않는다(1점)’, ‘일주일에 1~2일 이용한다(2점)’, ‘일주일에 3~4일 이용한다(3점)’, ‘매일 이용한다’(4점)로 주어졌다.

통제변인

정치참여에 영향을 미치는 요인들로서 통제변인으로 투입된 것은 첫째 사회경제적 요인(socioeconomic factors)이다. 이는 정치참여에는 인지적, 물질적 자원이 소요된다는 자원이론(resource theory)의 입장(Verba and Nie 1972)에 근거하고 있으며 구체적으로 성별, 연령, 소득, 교육수준이다. 연령은 응답자의 만 나이로 측정되었고 소득은 가구소득(household income)으로 측정했다. 교육수준은 중졸이하, 고졸, 대졸, 대학원 재학(이상)으로 구분되는 4점 척도로 측정하였으며 성별은 여성을 기준범주(reference category)로 하여 측정했다(여성=0, 남성=1). 둘째로는 정당일체감을 통제했다. 미시건 학파의 연구에 따르면 특정 정당에 대해 느끼는 일체감이 시민들의 정치적 행위에 핵심적 역할을 하는 변인이다(Campbell et al. 1960). 정치참여는 개인적인 행위이기도 하지만 동시에 외부의 동원으로부터 크게 영향을 받는 행위이기도 하다(Rosenstone and Hansen

1993). 즉 정치적 동원(political mobilization)의 효과를 통제하기 위해 두 가지 측면의 동원효과를 고려했다(정당과 후보자로부터의 동원, 이웃으로부터의 동원). 정치적 동원의 효과는 각각 ‘선거운동 기간 동안 정당 또는 후보자가 직접 또는 기타 다른 방식을 통해 귀하께 접촉한 적이 있습니까?’라는 문항과 ‘선거운동 기간 동안 친구, 가족, 이웃, 직장 동료 또는 다른 지인이 특정 정당이나 후보자에게 투표하라고 귀하를 설득한 적이 있습니까?’라는 문항으로 측정했다.²⁾ 정치에서 이념이 가지는 중요성을 고려하여 정치이념을 통제했다. 이념은 자기규정(self-placement)을 통해 측정하였다. 즉 응답자 스스로 매우 진보(0점)부터 매우 보수(10점) 까지 1점 단위로 구분된 척도에서 자기 자신을 위치짓는 값을 통해 측정했다. 또한 한국정치에서 중요한 요인중 하나인 지역변수를 통제했다. 지역변수는 수도권을 기준집단으로 하여 호남, 부산/경남(PK), 대구/경북을 각각 통제하였다. 정치심리적인 요인으로는 정치지식과 정치관심도를 통제하였다. 정치관심도는 ‘귀하께서는 정치에 얼마나 관심이 있으십니까’라는 문항(4점 척도)으로 측정하였고 정치지식은 정치에 관한 질문(20대 국회의원 정수, 국회의원 임기, 20대 총선 결과 원내 제2당, 응답자 지역구 당선자 이름, 국회의장 이름)의 정답(각 1점 부여)/오답(각 0점 부여)을 측정하였다. 이는 0부터 1까지 분포하도록 표준화하여 분석에 투입했다. 투표참여의 경우 두 가지 변수를 추가적으로 통제했다. 합리적 선택의 접근법에 입각해보면 정당 간의 정책차이는 자신이 지지하는 후보(또는 정당)이 당선되거나 낙선될 때 얻게 될 기대효용의 크기를 그만큼 크게 만든다(조성대 2006). 이에 따라 정당 간 정책차이를 얼마나 크게 인식하는지를 통제하였고, 라이커와 오데훅(Riker and Ordeshook 1968)에 따르면 투표행위에 대해 느끼는 의무감은 투표참여 여부를 결정짓는 매우 핵심적인 요소이므로 이를 통제하였다.

2) 예비적 분석

2) 물론 동원효과는 비선거적 참여든 투표참여든 효과를 가질 것이라는 이론적 기대가 가능하지만, 문항의 내용이 선거나 투표를 적시하고 있어 비선거적 참여를 검증하는 모형에서는 제외하였다.

회귀분석으로 진행하기에 앞서 주요 관심변수들, 즉 SNS 이용빈도와 정치참여의 양변인간 관계를 교차분석을 통해 살펴보았다.

〈표 1〉 SNS 이용빈도와 투표참여 교차분석

	투표	기권	계
이용하지 않는다	539 (83.31)	108 (16.69)	647 (100)
일주일에 1~2일 이용한다	119 (73.46)	43 (26.54)	162 (100)
일주일에 3~4일 이용한다	80 (75.47)	26 (24.53)	106 (100)
매일 이용한다	206 (68.90)	93 (31.10)	299 (100)
계	944 (77.76)	270 (22.24)	1,214 (100)

$\chi^2=27.152$, P-value=0.000

〈표 1〉은 SNS 이용빈도에 따라 투표참여와 불참 비율이 어떻게 달라지는지를 보여준다. 카이스퀘어 검정을 통해 살펴보면 우선 두 변인 간은 통계적으로 독립적이지 않다는 것을 확인할 수 있다. 다시 말해 두 변인 간에는 유의미한 통계적 연관(statistical association)이 관측되며, 이용빈도가 높아질수록 투표자의 비율이 오히려 낮아지고 있다. 이는 동원가설(mobilization hypotheses)에서 기대하는 것과는 반대의 결과이다. 그러나 미디어 이용에는 연령 등의 다른 변수들이 깊게 관여되어있고, SNS 이용에 영향을 미치는 변인들이 동시에 투표참여에도 영향을 미치기 때문에 이 관계가 다른 요인들을 통제한 뒤에도 동일할 것 인지는 미지수다. 이는 다음 장의 회귀분석을 통해 살펴본다.

이어지는 〈표 2〉부터 〈표 4〉까지는 비선거적 정치참여 행위를 각각 분석한 것이다. 〈표 2〉를 통해 보면 SNS 이용빈도가 높아질수록 사회운동 참여 비율이 점차 높아짐을 확인할 수 있다. 이러한 경향은 시위참여에 대해서도 동일하게 나타난다. 통계적 연관성 또한 카이스퀘어 검정을 통해 통계적으로 유의미한 관계에 있다는 것을 알 수 있다. 그러나 〈표 4〉에서 나타나듯, 정당 및 사회단체

가입과는 유의미한 관계가 관측되지 않는다. 마지막 <표 5>에서는 비선거 참여를 단 하나라도 해본 경험이 있는지를 측정한 변수와 SNS 이용빈도 간의 교차분석이다. 이러한 측정전략을 통해서 보면 SNS 이용빈도는 비선거적 정치 참여를 증진시키는 유의미한 영향력을 보이고 있다는 것을 확인할 수 있다.

<표 2> SNS 이용빈도와 사회운동 참여 교차분석

	참여	불참	계
이용하지 않는다	44 (6.81)	602 (93.19)	646 (100)
일주일에 1~2일 이용한다	20 (12.35)	142 (87.65)	162 (100)
일주일에 3~4일 이용한다	13 (12.26)	93 (87.74)	106 (100)
매일 이용한다	35 (11.71)	264 (88.29)	299 (100)
계	112 (9.23)	1,101 (90.77)	1,214 (100)

$\chi^2=9.737$, P-value=0.021

<표 3> SNS 이용빈도와 시위 참여 교차분석

	참여	불참	계
이용하지 않는다	22 (3.41)	624 (96.59)	646 (100)
일주일에 1~2일 이용한다	12 (7.41)	150 (92.59)	162 (100)
일주일에 3~4일 이용한다	3 (2.83)	103 (97.17)	106 (100)
매일 이용한다	25 (8.36)	274 (91.64)	299 (100)
계	1,151 (94.89)	62 (5.11)	1,213 (100)

$\chi^2=13.285$, P-value=0.004

〈표 4〉 SNS 이용빈도와 정당 또는 시민단체 가입 교차분석

	가입	미가입	계
이용하지 않는다	20 (3.1)	626 (96.9)	646 (100)
일주일에 1~2일 이용한다	1 (0.62)	161 (99.38)	162 (100)
일주일에 3~4일 이용한다	3 (2.83)	103 (97.17)	106 (100)
매일 이용한다	12 (4.01)	287 (95.99)	299 (100)
계	36 (2.97)	1,177 (97.03)	1,213 (100)

$\chi^2=4.287$, P-value=0.232

〈표 5〉 SNS 이용빈도와 비선거 참여 교차분석

	참여	불참	계
이용하지 않는다	58 (8.96)	589 (91.04)	647 (100)
일주일에 1~2일 이용한다	26 (16.05)	136 (83.95)	162 (100)
일주일에 3~4일 이용한다	16 (15.09)	90 (84.91)	106 (100)
매일 이용한다	51 (17.06)	248 (82.94)	299 (100)
계	151 (12.44)	1,063 (87.56)	1,214 (100)

$\chi^2=15.651$, P-value=0.001

IV. SNS 이용의 정치참여에 대한 효과

이상에서 확인한 SNS 이용빈도와 정치참여의 관계를 보다 엄격하게 살펴보기 위해 통제변수를 포함하여 로지스틱 회귀분석을 시행하였다. 우선 〈표 6〉은 투

표참여를 종속변인으로 한 로지스틱 회귀분석의 결과다. 통제변인들을 먼저 살펴보면 연령이 높을수록, 소득이 높을수록 투표하는 경향이 높다. 교육수준이 높은 유권자는 상대적으로 투표할 확률이 낮은 반면, 정치지식은 통계적으로 매우 견고한 긍정적 영향력을 표출하고 있다. 정당일체감을 가진 유권자는 어떠한 정당에 대해 정당일체감을 갖든지 무당파 유권자에 비해 투표참여에 더욱 활발하다는 점을 우선 확인할 수 있다. 정당과 이웃의 동원효과는 나타나지 않았으며 정당 간 정책차이에 대한 인식과 투표 의무감은 투표참여를 자극하는 요인이었으나 확인되었다. 로지스틱 회귀분석 결과 SNS 이용빈도는 투표참여에 유의미한 영향력을 보이고 있지 않으며 계수의 크기 또한 매우 작다는 사실이 관측되고 있다.³⁾ 통제변인에 대한 분석을 통해 생각해보면, <표 1>에서 확인했던 SNS와 투표참여 간의 부정적 관계는 연령변수 등 주요 통제변수들이 누락됨에 의해 나타난 허위적 관계였을 가능성이 매우 높다고 할 수 있다.

왜 SNS 이용빈도는 투표참여에 영향을 미치지 못했는가? 간단한 분석을 통해 살펴보면 우선 SNS의 이용이 정치에 대한 정보를 제공하고, 정치지식을 증가시켜 유권자의 인지적 자원을 강화한다는 이론적 기대는 사실이 아니라는 것을 알 수 있다. 표준화된 정치지식 변수와 SNS 이용빈도의 상관관계 분석 결과를 보면 피어슨 상관계수는 $-0.021(p\text{-value}=0.474)$ 로 통계적으로 유의미한 관계를 찾아볼 수 없었다. 또, 동료 시민들과 사회적 관계망을 형성함으로써 투표참여에 대한 의무감을 느낄 수 있다는 이론적 기대 또한 충족되는 것으로 보이지 않는다. 투표의무감과 SNS 이용빈도의 상관관계는 $-0.068(p\text{-value}=0.0175)$ 로, 통계적으로 유의미한 관계이기는 하나 의무감을 오히려 낮추는 것으로 나타난다.

그렇다면 공공정보를 얻기 위해 매체를 활용한다는 것은 투표참여에서 어떤 의미인가? 즉 단지 SNS 활용이 의미있는 영향력을 미치지 않는 것이라고 결론 짓기보다는 과연 정보취득 행위 자체가 투표참여에 어떠한 의미인지를 확인해보기 위해 모든 매체(종이신문, TV, 인터넷, SNS)의 활용빈도를 합하여 매체를 막

3) 본 연구와는 구체적인 조작정의나 분석전략이 상이하기는 하지만 정재관(2013)의 최근 연구에서도 이러한 경향이 발견된 바 있다. 정재관은 집합데이터를 통한 국가간 비교연구를 통해 인터넷 이용과 투표참여 간에는 유의미한 영향력이 발견되지 않는다는 것을 보고하였다.

론하고 공공정보 취득을 위해 매체를 이용하는 정도를 측정하여 모형에 투입했다. 이 변수는 SNS 활용빈도의 대체변수로서 투입하였는데, 여기에서도 유의미한 관계는 발견되지 않았다. 이를 통해 보면 과연 정보취득을 위해 매체를 활발히 이용한다는 것이 과연 투표참여에 실질적인 의미를 가지고 있는 행위인지를 근본적으로 고찰해야 할 필요성이 있다고 할 수 있다.

물론 이러한 해석을 강하게 뒷받침할 수는 없다. 즉 현재 단계에서의 통계분석은 물론 이론적 고려 또한 지나치게 단조롭기 때문에 무리가 따르는 것은 사실이다. 그러나 SNS의 활용이 투표참여에 긍정적 영향을 미치리라 기대했던 이론적 기대를 그대로 유지하기에는 많은 의구심을 가지지 않을 수 없는 것 또한 사실이다. 이 문제를 규명하기 위해서는 정치참여에 관여되는 여러 인자들과 SNS 이용 간의 이론적, 경험적 관계를 새로운 시각에서 살펴볼 필요성이 제기될 것이다.

〈표 6〉 투표참여의 결정요인

		종속변인: 투표참여	
		(1)	(2)
연령		0.399 ^{***} (0.084)	0.396 ^{***} (0.079)
성별		-0.208 (0.183)	-0.207 (0.183)
소득		0.162 [*] (0.069)	0.163 [*] (0.070)
교육수준		-0.254 [†] (0.143)	-0.252 [†] (0.145)
이념		0.011 (0.049)	0.010 (0.049)
정치지식		1.374 ^{***} (0.367)	1.377 ^{***} (0.369)
지역	호남	0.705 [*] (0.276)	0.702 [*] (0.277)
	TK	0.102 (0.272)	0.101 (0.272)
	PK	0.040 (0.246)	0.043 (0.247)

정당일체감	새누리당	0.920 ^{***} (0.268)	0.921 ^{***} (0.268)
	더불어민주당	1.038 ^{***} (0.261)	1.041 ^{***} (0.261)
	국민의당	0.992 ^{**} (0.362)	0.989 ^{**} (0.362)
정당 접촉		0.165 (0.213)	0.165 (0.213)
이웃 접촉		0.420 (0.280)	0.420 (0.280)
정책차이 인식		0.247 [†] (0.136)	0.247 [†] (0.136)
투표 의무감		0.622 ^{***} (0.090)	0.622 ^{***} (0.090)
SNS 이용빈도		0.008 (0.077)	
공공정보 매체이용(총합)			-0.019 (0.421)
상수항		-2.656 ^{***} (0.684)	-2.629 ^{***} (0.664)
N		1,082	1,082
Pseudo R-Squared		0.223	0.223
Log Likelihood		-423.542	-423.547

† p<0.1, * p<0.05, ** p<0.01, *** p<0.001

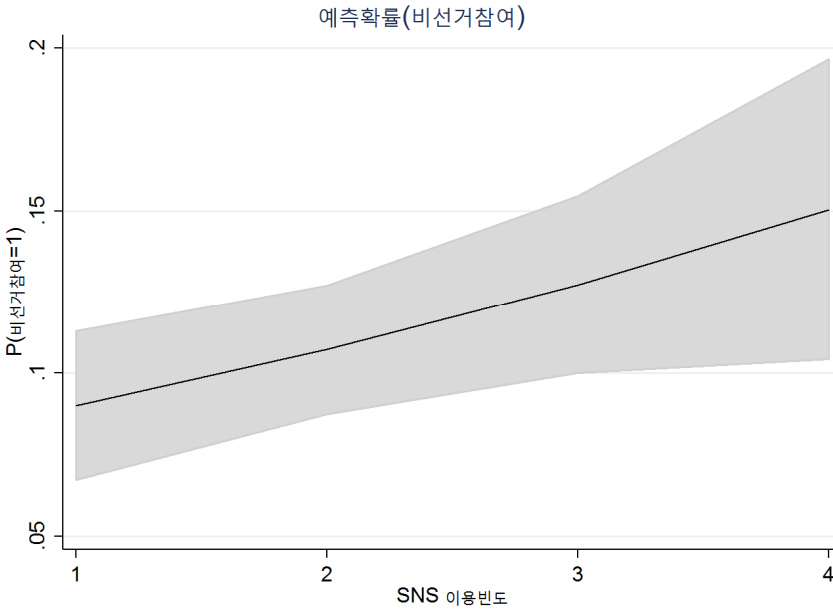
그렇다고 해서 SNS 이용이 정치참여에 어떠한 영향력도 없는 것은 아니다. 아래 <표 7>의 모형 2를 보면, 시위참여에는 유의미한 영향력을 표출하고 있다. 지난 2008년 미국산 쇠고기 수입반대 시위부터 최근의 청와대 비선실세 논란에 이은 시민들의 저항이 SNS를 통해 관련 정보가 확산되고, 시민들의 참여가 독려되고 있다는 점을 생각해보면 SNS는 비선거적 참여에서 매우 중요한 동원기능을 수행하고 있다고 할 수 있다. 다른 비선거적 참여에 대해서는 유의미한 효과가 표출되지 않았지만 기준을 조금 완화하여 분석한 모형 4, 즉 사회운동 참여, 시위 참여, 정당 단체 가입 중 하나라도 참여해본 경험이 있는지의 여부를 분석한 결과에서도 SNS 이용빈도의 영향력은 유의미하고 긍정적으로 나타나고 있다.

〈표 7〉 비선거 참여의 결정요인

		(1) 사회운동 참여	(2) 시위 참여	(3) 정당, 단체 가입	(4) 비선거적 참여
연령		-0.146 (0.106)	-0.111 (0.137)	-0.112 (0.171)	-0.109 (0.092)
성별		-0.292 (0.226)	0.310 (0.307)	0.421 (0.391)	-0.153 (0.198)
소득		-0.119 (0.084)	-0.398*** (0.112)	-0.340* (0.138)	-0.149* (0.073)
교육수준		0.530** (0.176)	0.695** (0.230)	0.178 (0.284)	0.380 (0.152)
이념		0.017 (0.059)	0.047 (0.075)	-0.073 (0.092)	0.029 (0.051)
정치지식		1.938*** (0.529)	0.182 (0.658)	0.267 (0.842)	1.271** (0.446)
지역	호남	-0.399 (0.305)	-0.012 (0.363)	-0.298 (0.480)	0.083 (0.248)
	TK	-0.969* (0.395)	-1.381* (0.624)	-0.203 (0.523)	-0.548 (0.313)
	PK	-1.035** (0.373)	-1.072* (0.499)	-2.176* (1.033)	-0.950** (0.328)
정당일체감	새누리당	-0.006 (0.320)	0.501 (0.448)	0.826 (0.562)	-0.056 (0.283)
	더불어민주당	-0.123 (0.296)	1.078** (0.361)	1.101* (0.483)	0.146 (0.250)
	국민의당	0.040 (0.373)	0.571 (0.491)	1.180* (0.572)	0.036 (0.335)
SNS 이용빈도		0.103 (0.094)	0.238* (0.123)	0.135 (0.159)	0.194* (0.083)
상수항		-3.914*** (0.776)	-4.004*** (0.982)	-2.950 (1.196)	-3.188*** (0.663)
N		1,089	1,089	1,089	1,089
Pseudo R-Squared		0.076	0.109	0.088	0.058
Log Likelihood		-308.777	-196.606	-137.983	-384.635

† p<0.1, * p<0.05, ** p<0.01, *** p<0.001

〈그림 1〉 비선거 참여 예측확률



* 실선은 예측확률(predicted probability)을 나타내며, 음영은 95% 신뢰구간임.

〈그림 1〉은 SNS 이용빈도에 따라서 비선거적 정치참여를 할 확률을 계산하여 보여주고 있다. 이 그림은 다른 변인들을 평균에 고정시킨 상태에서 SNS 이용빈도 변인의 예측확률을 계산하여 그린 것이다. 이 그림을 통해 보면 다른 모든 변인들이 평균이라고 가정할 때, SNS 이용빈도가 높아질수록 비선거적 참여를 할 확률이 점차 높아지고 있다.

여기까지의 분석결과는 SNS의 이용이 시민들의 정치적 행동에 어떠한 영향을 미치는지에 대해 다양하게 생각해볼 문제를 던져준다. 첫째로 SNS의 이용이 이론적 기대들과는 달리 투표참여에 영향을 미치지 못하는 것에 대한 이론적, 경험적 해명이 필요하다. 단지 SNS가 투표참여에 영향을 미치지 못했다는 것에서 이해를 마치기보다 그동안 믿어왔던 소셜네트워크의 정치적 효과에 대해 재숙고해야 할 필요성이 있다는 것이다. 사실 SNS를 공공정보를 획득하기 위한 창으로도 사용하는 것이 사실이지만 기본적으로는 사적교류의 목적이 추가 되고 있는 것

또한 사실이다. SNS의 이용목적이나 이용행태가 공적인 부분보다 사적인 부분에 더욱 집중되어있다는 현실을 고려했을 때 SNS의 정치적 효과를 지나치게 과장하여 생각한 것은 아닐지 고민해볼 필요가 있다. 또한 직접적 효과를 보이지 않는다 하더라도 간접적 효과 내지는 상호작용효과를 통해 시민들의 투표참여에 영향을 미치고 있지는 않은지에 대한 이론적 고찰 또한 요구된다. 둘째로, SNS의 이용이 투표참여에는 영향이 없지만 시위참여에는 유의미한 영향을 나타내고 있다. 다시 말해 SNS는 선거적 참여보다는 비선거적 참여를 동원해내는 역할을 수행하고 있다는 것이다. 투표라는 행위는 기본적으로 개인이 홀로 하는 행위이지만, 시위의 경우 많은 사람들이 동시에 같은 공간에서 모이는 행위다. 그렇게 보면 SNS가 제공하는 사회적 연결망, 사회자본 등은 시위에 참여하려는 동기를 실제 행위로 이끌어낼 수 있는 동력을 가지고 있다고 볼 수 있을 것이다.

V. 결론

본 연구는 한국 제20대 총선에서 SNS의 정치적 동원효과가 다양한 정치참여 행위에 대해 각각 어떻게 나타났는지를 살펴보았다. 개인의 자발성이 강조되고 개인간 관계성이 두드러지기 때문에 SNS는 특히 비선거적 정치참여에 유의미한 영향력을 미칠 수 있으리라고 기대했다. 20대 총선 직후 진행된 설문조사 자료를 분석한 결과 SNS의 이용은 비선거적 정치참여를 자극하는 효과를 유의미하게 표출하고 있었다. 그러나 투표참여에는 통계적으로 유의미한 영향력을 미치지 못했으며, 이는 비단 SNS의 이용뿐만 아니라 정치정보를 획득하기 위해 매체를 활용하는 것은 투표참여와는 별반 관계가 없다는 것이다. 그러나 매체를 불문하고 정보취득이 많을수록 비선거적 참여의 가능성은 높아지고 있었다.

이러한 분석을 통해 보면 본 연구가 갖는 함의는 크게 두 가지로 정리해볼 수 있다. 첫째로 정보취득 행위가 투표참여에 유의미한 기여를 할 것이라는 기대에 대한 이론적인 재고가 필요하다는 것이다. 합리적 투표에서 상정하는 도구적 투표(instrumental voting)라고 상정하면 사실 정보취득 행위 자체가 투표참여의

산술에 영향을 미칠만한 요소는 크지 않아 보인다. 이론적으로 생각해보면 투표 참여는 기본적으로 매우 적은 비용이 필요한 행위이고 그 비용 중 정보취득이 차지하는 비중이 거의 없을지도 모른다. 둘째로 자신이 구축한 관계망 속에서 타인의 의견을 접하고 지속적으로 정치적인 의견을 교환하는 SNS 의사소통의 구조는 비선거 참여에는 유의미한 영향력을 보이고 있었다. 요컨대 SNS의 정치적 동원효과는 ‘어떠한 행위로 동원해내느냐’에 따라서 달리 나타나는 것이다. SNS를 활용할수록 정보획득량이 증가하여 이것이 단순히 참여를 증진시킬 것이라고 해석하는 것은 지나치게 단편적인 분석일 것이다. 개별 매체들이 갖추고 있는 커뮤니케이션의 구조를 파악하고 그러한 구조가 시민들의 다양한 정치적 행위와 어떻게 연결될 수 있는지를 논의해야 현재 까지 존재해왔던 다양한 매체들은 물론, 미래에 새로 등장하게 될 매체들에 대한 분석이 이론적, 경험적으로 견고한 분석이 될 것이다.

참 고 문 헌

- 박근영. 2014. “뉴미디어의 소비가 선거에서 지지후보 변경에 미치는 영향: 2014년 서울시장 선거의 예.” 『사회과학연구』 제31권 제1호: 29-55.
- 송경재. 2011. “소셜 네트워크 세대의 정치참여.” 『한국과 국제정치』 제27권 제2호: 57-88.
- 송경재·임정빈·장우영. 2016. “SNS는 정치를 어떻게 변화시키는가?: 정치정보 신뢰, 지지의 전환 및 참여 효능감을 중심으로.” 『한국콘텐츠학회논문지』 제16권 제7호: 154-167.
- 정재관. 2013. “정보통신기술 혁명은 위기의 대의민주주의를 구할 것인가?: 인터넷 이용의 정치참여 효과에 대한 국제비교.” 『국제관계연구』 제18권 제2호: 137-164.
- 조성대. 2006. “투표참여와 기권의 정치학: 합리적 선택이론의 수리모형과 17대 총선.” 『한국정치학회보』 제40권 제2호: 51-74.
- 조희정. 2011. “2011년 중동의 시민혁명과 SNS의 정치적 매개역할.” 『한국정치연구』 제20권 제2호: 309-338.
- Campbell, A., P. E. Converse, W. E. Miller and D. E. Stokes. 1960. *The American Voter*. New York: John Wiley & Sons.
- McLuhan, Marshall. 1994. *Understanding Media: The Extension of Man*. Cambridge: MIT Press.
- Ndavula, John and Hellen Mberia. 2012. “Social Networking Sites in Kenya: Trigger for Non-Institutionalized Democratic Participation.” *International Journal of Business and Social Science* 3(13): 300-306.
- Verba, Sidney and Norman Nie. 1972. *Participation in America: Political Democracy and Social Equality*. New York: Harper & Row.
- Riker, William and Peter Ordeshook. 1968. “A Theory of the Calculus of Voting.” *American Political Science Review* 62(1): 25-42.

Rosenstone, S. J. and Hansen J. M. 1993. *Mobilization, Participation, and Democracy in America*. New York: Macmillian.

Social Media Use and Political Engagement in Korea: The Case of 20th General Election in 2016

Youngdeuk Park*·Jaemook Lee**

ABSTRACT

This paper examines the effects of SNS(Social Networking Sites) utilization on political participation by analyzing the 20th general election in South Korea. Previous research has shown that new media such as the Internet and SNS promote political participation by reducing the cost of political information acquisition, but the results of this study show that SNS utilization and voting participation are not significant. However, SNS utilization has a significant influence on non-electoral political participation. This seems to be because citizens can organize themselves by the social network and social capital provided by SNS.

Keywords: Social Networking Sites, Political Participation, Turnout, Non-Electoral Participation, Media Effects

투고일: 2017.03.28. 심사일: 2017.04.03. 게재확정일: 2017.04.17.

* Yonsei University

** Hankuk University of Foreign Studies

[DOI] <http://dx.doi.org/10.21487/jrm.2017.05.2.1.23>

【연구논문】

국제이주 연구에 있어 정치학적 접근과 방법론적 쟁점

이 병 하*

논문요약

본 논문은 국제이주 연구에 있어 여타 사회과학에 비해 저발전된 정치학적 접근을 모색해보고자 한다. 이를 위해 우선 경제학적, 사회학적 접근법을 살펴보고, 정치학적 접근은 이익, 권리, 제도라는 세 가지 키워드에 기초하여 발전해왔음을 주장한다. 각 키워드에 기초하여 정치학적 접근은 정치경제학적 이론, 국제규범 혹은 사법권에 기초한 모델, 그리고 신제도주의적 이론으로 특화될 수 있음을 강조한다. 또한 방법론적 쟁점에 있어 국민국가를 분석단위로 삼는 방법론적 민족주의를 넘어 도시와 같은 대안적인 분석단위들이 모색되어야 함을 주장한다. 마지막으로 이민정책 평가에 있어 한 국가의 이민정책을 이념형을 통해 진단하는 것을 넘어 최근에 개발되고 있는 이민정책 지표들을 활용하여 지표 중심 접근법을 새로운 분석틀로 삼는 것을 제안하고자 한다.

키워드: 국제이주 연구, 정치학적 접근, 정치경제학, 신제도주의, 국제규범, 방법론적 민족주의

* 서울시립대학교. 이 논문은 2016년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2016S1A3A2923970).

I. 서론

흔히 우리는 세계화의 시대에 살고 있다고 말한다. 여기서 세계화란 우리와 멀리 떨어져 있는 지리적 공간과 삶의 영역들 간에 상호의존성이 증대하고 있다는 것을 의미한다. 세계화에서 강조되는 현상으로는 상품, 자본, 정보의 활발한 교환과 이로 인한 시공간의 압축을 들 수 있다. 그러나 세계화의 요소 중 간과되었던 것이 국경을 넘나드는 사람들 간의 이동이다. 국제이주기구(IOM)의 통계에 따르면 전 세계 이주민은 약 2억 1,400만 명으로 전 세계 인구의 약 3.1%를 차지한다(한건수 2012). 전 세계 인구에서 이주민이 차지하는 비율이 높다고 할 수는 없지만 과거 국제이주의 영향을 받는 지역이 북미와 유럽에 한정되어 있었던 반면, 현재는 전 세계 지역이 국제이주의 영향권에 있을 정도로 국제이주는 점점 중요한 사회현상이 되어가고 있다.

또한 최근 국제정치질서를 흔든 일련의 사건들을 보더라도 국제이주는 더 이상 국내외 정치의 변동을 이해하는데 있어 주변부 이슈가 아니다. 시리아 난민 사태가 전 세계 여론에 미친 영향, 브렉시트로 인한 유럽연합 정치적 질서에 대한 근본적인 도전, 도널드 트럼프의 당선과 취임 한 달 내에 벌어진 이민 관련 행정명령으로 인한 논란 등은 포퓰리즘(populism)이 전 세계적으로 확산되고 있음을 의미한다. 포퓰리즘의 근저에는 ‘순수한 시민들 대 타락한 엘리트’라는 이분법적 구도가 자리 잡고 있으며, 국제이주 문제도 이러한 이분법적 도식을 강화시키고 있다. 높은 반이민 여론에도 불구하고 이민정책이 계속 확대되는데 대한 대중들의 불만이 포퓰리즘의 한 원인이라고 할 수 있다.

또한 국내로 시선을 돌려보면, 우리에게 국제이주, 난민, 다문화라는 말은 더 이상 낯선 용어가 아니다. 시리아 난민 사태로 인해 대규모의 인구가 유럽으로 이동하는 장면은 우리에게 난민에 대한 인도주의적 관심을 새롭게 환기시키고 있으며, 2016년 12월 말 기준으로 약 200만 명이 체류하고 있으며, 이는 한국 전체 인구의 약 4%를 차지한다는 통계가 말해주듯이 이주민들은 어느덧 우리의 일상생활에 다가와 있다. 결혼이주여성, 다문화 가정 자녀들, 외국인 노동자, 재

중동포 등이 미디어의 단골 소재가 된지 오래이며, 이들은 각종 TV 프로그램의 주요 출연자이기도 하다. 이처럼 한국 사회는 국제이주의 초기 단계를 넘어 본격적으로 다인종, 다문화 사회로 진입하고 있다.

국제이주는 일반적으로 국민국가에 어떻게 국경을 통제, 관리해야 하는가하는 문제를 던져주기도 하고, 어떻게 이미 유입된 외국인, 이민자들을 주류 사회에 통합시켜야 하는지 고민하게 만든다. 더 나아가 한 사회의 멤버십과 경계에 대한 근본적인 물음을 던진다. 국경통제에 실패하면, 미등록 이주민이 증가하여 국가의 행정적 부담과 안보 문제를 야기한다. 이질적 문화와 정체성을 가진 외국인, 이주민 집단이 이민 수용국에 잘 통합되지 못하면, 교육 시스템 상에서 낙후되는 것은 물론 노동시장에서 낙오되어 사회의 부담으로 남고, 이는 선주민과 이주민 간의 갈등으로 비화된다.

자본의 이동, 상품의 이동과 더불어 세계화의 중요한 요소인 국경을 넘나드는 인적 이동은 국제적인 수준에서 국민국가는 물론 다양한 비국가 행위자들이 참여하는 정책적 대응을 중요하게 만들었다. 일례로 시리아 난민 사태로 인해 유럽연합의 근간이 도전받고 있으며, 난민의 급증으로 인해 경제적 이주민과의 경계가 모호해지면서 국민국가들은 난민에 대한 인도주의적 접근으로부터 후퇴하고 있다. 또한 국제이주와 테러리즘이 결합하는 현상이 늘어나면서 이민정책이 안보화(securitization)하고 있으며, 이주민에 의한 송금(remittance)이 증가하면서 국제이주와 국제개발협력 의제가 서로 결합하는 복합적인 현상이 생겨나고 있다. 국제이주가 기타 다른 국제적인 현상과 맞물리면서 국제사회는 협력을 통한 효율적 정책적 대응을 모색해야 할 필요성이 증가하고 있다.

국내적으로 보더라도 1980년대 후반 이후 진행된 이주노동자의 유입과 재중동포의 국내 노동시장으로의 편입, 그 이후 2000년대 들어 급증한 결혼이민 등 다양한 이주패턴으로 인해 한국사회는 이 전에 겪어 보지 못한 다문화적 현실이라는 새로운 문제를 고민하게 되었다. 이는 더 이상 한국사회가 글로벌한 현상인 국제이주의 압력으로부터 무관한 곳이 아님을 보여주는 것이며, 국제이주가 동반하는 다문화적 현실에 효율적인 정책적 대응이 필요함을 말해준다.

국내외적으로 국제이주에 대한 효율적 정책적 대응이 어느 때보다 필요한 지

금, 중요한 것은 이러한 정책적 대응이 국제이주에 대한 연구 결과물에 기반해야 한다는 점이다. 그리고 복합적인 국제이주 현상을 분석하기 위해서는 사회과학의 제반 영역에서 독자적인 접근법을 발전시키고 엄밀한 방법론을 모색해야 한다. 각 학문 영역의 발전은 더 나아가 국제이주의 복합성과 다양성을 종합적으로 조망할 수 있는 학제 간 연구로 발전되어야 한다(신지원 2011).

그러나 통합적인 학제 간 연구가 이루어지기에는 아직까지 사회과학 제반 분과 간 불균형이 존재한다. 경제학, 사회학, 인류학 등에서 국제이주의 동기와 원인, 이민자 네트워크, 이민자 집단 정체성 등의 독자적인 학문적 영역을 발전시켜온 반면, 정치학은 이에 비해 아직 국제이주 연구가 저발전된 상태이다. 따라서 본 논문은 국제이주 연구에 있어 정치학적 접근을 모색해보고 연구방법 상에 있어서 국민국가 모델과 관련된 두 가지 쟁점 즉 방법론적 민족주의(methodological nationalism)와 이민자 통합정책 모델을 토론하고자 한다. 그리고 이에 대한 대안으로 도시를 분석단위로 하는 연구의 필요성과 지표 기반 접근법(index-based approach)을 살펴보고자 한다.

II. 국제이주 연구에 있어 사회과학 분과의 접근법

국제이주 연구는 궁극적으로 학제 간 연구를 지향한다. 왜냐하면 국제이주의 결정요인, 과정과 유형, 그리고 국제이주 후의 통합 과정에는 각 사회과학 분과가 연구하는 다양한 원인들이 존재하기 때문이다(Castles and Miller 2009). 경제학, 사회학, 인류학, 정치학, 법학, 지리학, 인구학 등 다양한 사회과학 분과들이 국경을 넘나드는 인구 이동의 과정들을 개별 분과학문 내의 독창적인 접근법과 이론들을 바탕으로 접근하고 있다. 국제이주 연구가 학제 간 연구를 추구하고, 국제이주 과정을 연구함에 있어 다양한 접근법을 통해 이주 전체를 종합적으로 이해하려고 하지만, 학제 간 연구를 지향하기에 앞서 각 사회과학 분과들이 어떤 접근법을 취하고 있으며, 어떤 상이한 연구 질문을 던지고 있는

지 파악할 필요가 있다.

〈표 1〉 국제이주 연구에 있어 사회과학 분과의 접근법

사회과학 분과	연구 질문	분석수준/ 단위	주요 이론	가설의 예
인류학	이주는 어떻게 문화적 변동에 영향을 미치고, 종족적 정체성에 영향을 미치는가?	보다 미시적 수준 개인, 가정, 그룹	관계론적 이론 혹은 구조주의, 초국가주의	사회적 연결망은 문화적 차이를 유지하는데 기여한다
경제학	무엇이 이주의 원인과 영향을 설명할 수 있는가?	보다 미시적 수준 개인, 가정, 노동시장	비용-편익 흡입-배출 이론	통합은 이민자의 인적 자본에 따라 달려 있다
사회학	무엇이 이민자 통합을 설명할 수 있는가?	보다 거시적 수준 종족적 그룹, 사회적 계급	구조주의 기능주의	이민자 통합은 사회적 자본에 달려 있다
정치학	왜 국가는 이민을 통제하는데 어려움을 겪는가?	보다 거시적 수준 정치적, 국제적 시스템	합리주의적 이론 제도주의	국가는 이민확대를 원하는 이익집단의 이익에 사로잡혀 있다

출처: Brettell and Hollifield, 2000, 3. 부분 수정.

경제학은 ‘무엇이 국제이주의 원인이며, 무엇이 국제이주의 경향과 패턴을 설명할 수 있는가’하는 질문을 던진다. 경제학이 주로 사회행동을 선택의 문제로 치환한다고 보면, 국제이주 역시 개인 혹은 가족의 선택이라고 할 수 있다. 따라서 국제이주에 대한 경제학적 접근에서 주류 이론은 신고전파 경제학이다. 신고전파 경제학은 이민 송출국에 머물 때 그리고 이민 수용국으로 이주할 때의 비용과 편익을 합리적으로 계산하는 개인에 초점을 두고 국제이주를 설명한다. 개인들은 자신의 효용을 극대화한다는 가정 하에 자신들의 복지를 극대화할 수 있

는 거주 국가를 탐색하고, 이러한 선택을 위한 정보는 ‘이민 시장’에 모이게 된다. 잠재적 이민자는 이민 시장으로부터 자신의 이익을 극대화하기 위한 정보 즉 임금 수준, 고용 기회에 대한 완전한 정보를 획득한다고 가정한다. 따라서 국제이주에서 인구이동은 일반적으로 보았을 때 소득이 낮은 곳에서 높은 곳으로 혹은 인구밀도가 높은 곳에서 낮은 곳으로 향한다. 대표적인 이론이 ‘배출-흡입(push-pull)’이론으로 가난, 낮은 삶의 기준, 높은 실업률 등 개인이 출생지를 떠나게 만드는 배출 요인과 높은 임금, 높은 삶의 기준, 보다 나은 경제적 기회 등 개인이 수용국으로 향하는 흡입 요인이 만나 국제이주가 발생한다고 주장한다(Masse et al. 1993).

또 다른 경제학적 접근으로 ‘이민의 신 경제학(new economics of migration)’이 있다. 이민의 신 경제학은 국제이주를 결정하는 단위가 개인이 아니라 가족이라고 본다. 이민을 더 높은 임금을 받기 위한 개인의 결정이라고만 보기에는 부족하며, 이민을 결정하는 것은 소득의 원천을 다각화하기 위한 가족의 결정이라는 것이다. 예를 들어 농업에 종사하는 가족이 있다고 할 때, 농산물 가격의 불안정성으로 인해 모든 가족이 농업 분야에 일하는 것은 위험이 크다. 따라서 소득 다각화와 분산투자의 한 형태로 가족 구성원 중 일부가 이민을 가서 모국으로 송금을 보내는 결정을 내린다는 것이다(Castles and Miller 2009).

위의 두 이론은 분석단위로 개인과 가족을 설정한다. 반면 ‘이중노동시장 이론’은 국제이주의 원인을 더 거시적인 차원 즉 이민 수용국의 노동시장에서 찾는다. 이중노동시장 이론은 이민 수용국의 흡입 요인에 초점을 맞추어 선진국의 분절화된 이중노동시장 중 노동집중적인 산업의 노동시장에서 발생하는 노동력 부족이 국제이주의 원인이라고 주장한다. 이 이론에 따르면 선진국의 노동시장은 크게 자본집중적인 산업과 노동집중적인 산업으로 나뉜다. 자본집중적인 산업의 노동자는 고숙련 기술을 보유하고 있으며 고용주도 이러한 고급 인력의 안정성을 보장하기 위해 각종 사회복지 혜택을 제공한다. 그러나 노동집중적인 산업의 노동자는 직업이 불안정하며 저숙련 기술을 보유하고 있다. 따라서 자국의 노동자를 이러한 산업 분야로 유도하는 것이 점점 어려워진다. 노동집중적인 산업 분야의 노동력 부족 현상을 해소하기 위해 고용주는 외국인 노동자에 눈을 돌리

고, 노동집중적인 산업 분야는 점차 구조적으로 국제이주에 의존하게 된다는 것이다(Castles and Miller 2009).

사회학과 인류학은 각각 ‘무엇이 이민자 통합을 설명할 수 있는가,’ 그리고 ‘이민은 어떻게 문화적 변화와 집단 정체성에 영향을 주는가’ 같은 연구 질문을 던진다. 연구의 초점은 조금씩 다르지만 공통적으로 두 학문 분과는 인간의 사회적 관계망에 초점을 맞추고 있다. 이민자, 과거 이민자, 그리고 잠재적 이민자들이 친척, 가족관계, 공통적 공동체적 요소를 통해 연결된 대입관계의 집합으로 정의되는 이민자 네트워크를 통해 국제이주를 접근한다. 이민자, 과거 이민자, 그리고 잠재적 이민자들은 이민자 네트워크를 통해 시간과 공간을 가로질러 연결된다. 국제이주의 과정에서 이민자 네트워크는 중요한 역할을 수행한다. 이민 수용국과 송출국 사이에 친구, 친척 등의 관계망을 통해 서로 정보를 교환하고, 도움과 의무의 관계망이 발생하여 이러한 네트워크는 국제이주가 낳을 수 있는 비용과 위험을 감소시키고, 이익을 증가시켜 궁극적으로는 국제이주의 가능성을 높인다. 또한 이민자 네트워크는 이민 수용국에 정착, 통합되는 과정에서 일종의 사회적 자본의 역할을 하여 개인과 집단이 문화적 차이를 유지하고, 생계를 위한 자원과 안전망을 제공한다. 더 나아가 이민자 네트워크는 사회학이 주목하는 이민자들만의 거주 지역과 공동체 형성에 기여한다. 이민자들은 특정 거주 지역에 밀집하여 자신들만의 종교시설, 상점, 결사체 등을 발전시키고, 가족재결합을 통해 공동체의 규모를 늘려간다. 이 과정에서 이민자 네트워크는 이민자들 고유의 문화 정체성을 유지하는데 큰 도움이 되며, 더 나아가 이민자들 고유의 정체성과 이민 수용국의 정체성이 서로 융합하는 이민자 2세대 형성에도 영향을 미친다(Brettell and Hollifield 2000).

III. 국제이주 연구에 있어 정치학적 접근

위에서 살펴본 것처럼 사회과학의 각 분과는 국제이주 연구에 있어 각자 나름

의 연구질문, 주요 이론, 분석단위 등을 발전시켜왔다. 그러나 국제이주 연구를 발전시켜온 역사를 보면 정치학은 상대적으로 저발전된 상태라고 평가할 수 있다. 1990년대 이전만 하더라도 정치학 내에서 국제이주를 주된 연구 주제로 삼았던 경우는 극히 드물며, 미국정치학회 내에서 국제이주만을 다루는 분과도 2012년에 되어서야 설립되었다. 그러나 1990년대 이후 국제이주도 정치학의 주요 연구 주제로 부상하고 있으며 기존의 정치학적 이론을 활용하여 독자적인 분석틀을 모색하고 있다.

정치학에서 국제이주를 연구하는 주요 분석틀은 크게 이익(interests), 권리(rights), 제도(institutions)라는 세 가지 키워드를 중심으로 발전되어왔다. 국제이주 연구에 있어 정치학적 접근은 “서구국가에서 반(反)이민정책의 확산에도 불구하고 왜 이민정책은 확대의 방향으로 가는가”하는 연구 질문을 던진다. 이를 ‘간격 가설(gap hypothesis)’라고 하는데 주요 이민 수용국에서 관찰되는 높은 수준의 반 이민 여론이 왜 자유민주주의 체제 하에서 선거 과정을 통해 이민 확대를 저지하는 정책으로 표출되지 못하고, 정책의 결과물은 이민 확대의 경향을 가지는가를 설명하려고 노력한다. 이익에 기초한 정치경제학적 이론은 국가가 친 이민정책을 통해 이익을 얻는 이익단체의 논리에 포획되어 있다는 점을 강조하고, 권리에 기반한 이론은 국제인권규범 혹은 국내적으로 사법부의 판결 등에 의해 이민을 확대하려는 국가의 정책적 노력이 제한을 받는다는 점을 강조한다. 또한 신제도주의 이론은 단일한 국가라는 가정을 넘어 다양한 이익과 정책 목표를 가진 정부 부처 간의 경쟁에 주목함으로써 부처 간 경쟁을 통해 이민정책이 확대의 경향을 갖는 것을 설명한다. 즉, 국제이주 연구에 있어 정치학적인 접근은 국제 이주를 둘러싼 이익(interests), 권리(rights), 제도(institutions)라는 세 가지 키워드에 기초해 이루어져 왔으며(Freeman 2005), 각 이론적 모델은 ‘왜 국가가 이민 통제에 어려움을 겪는가’하는 질문에 답을 구하려고 함으로써 정치학적 접근법을 발전시키고 있다.

이익이라는 키워드에 기초한 정치경제학적 이론은 고객정치(client politics)라는 개념을 사용한다. 이에 의하면 이민의 확대에 인한 이익이 집중되는 집단들(사용자 집단 또는 이민자 그룹)의 로비를 통해 이민정책이 주로 형성된다고 본다.

정치경제학적 이론은 프리만(Gary Freeman)에 의해서 선도되었다. 그는 윌슨(James Q. Wilson)의 4가지 정책유형을 차용하여 그 중 하나인 고객정치(client politics)라는 개념을 통해 이민정책을 분석하려고 하였다. 윌슨의 정책유형에 따르면 어떤 정책으로 인한 이익이 특정 집단에 집중되고, 비용이 분산된다면, 해당 정책을 둘러싼 정치모드는 고객정치가 된다. 프리만은 이민정책이야말로 전형적인 고객정치 유형에 속한다. 왜냐하면 이민의 확대로 인해 이익을 얻는 사용자 집단이나 이민자 그룹에 이익이 집중되는 반면, 이민의 확대로 인해 늘어나는 비용은 사회 전반에 걸쳐 광범위하게 일반 대중에게 분산되기 때문이다. 이민이 확대된다면 사용자 집단은 보다 적은 비용으로 활용할 수 있는 노동력이 늘어나 이익이 집중되는 반면, 이민의 확대로 인한 사회복지비용 증가 등은 일반 대중에게 잘 인식되지 않는다. 따라서 이민정책으로 이익을 보는 고용주와 이민자 집단이 훨씬 더 잘 동원되고 조직되어 정책결정자에게 로비를 시도할 유인동기가 더 많은 반면, 이민 확대로 인한 비용을 지불하게 되는 대중들은 조직화되기 어렵다. 프리만에 의하면 이민정책은 잘 조직되지 못한 여론과는 유리된 채 정책결정자들과 조직된 집단들 사이의 고객 관계를 통해 확대의 길을 걷게 된다는 것이다(Freeman 1995).

정치경제학적 이론이 간접 가설에 대해 고객정치 유형을 통해 친 이민적인 이익집단의 영향력을 강조한다면, 권리에 기반 한 이론들은 국민국가 외부 혹은 내부에서 오는 권리의 영향력에 집중한다. 권리에 기반한 접근법은 크게 두 가지로 나눌 수 있는데 이민정책을 형성하는 국민국가의 능력을 무엇이 제한하는가에 따라 구분된다. 국민국가의 능력을 제한하고 이주민들의 권리를 향상시키는 근거를 국내에서 찾지 않고, 국제규범과 같은 국민국가 외부에서 미치는 영향에 주목하는 학자들이 있다. 특히 사센(Saskia Sassen), 제이콥슨(David Jacobson), 소이살(Yasemin Soysal) 과 같은 소위 ‘글로벌리스트(globalist)’들은 국제이주와 이민정책을 보는데 있어서 국제규범과 같은 초국가적 힘이 이민정책을 규정하는 중요한 동인이라고 주장한다. 반면 옉케(Christian Joppke)와 같은 학자들은 국제규범의 영향력은 과장된 것이며 한 국가의 이민정책에서 국민국가의 권한을 제한하고, 이민자들의 권리를 향상시켜 온 것은 국내 정치과정에 내재된 자유주

의 규범 특히 자유주의 규범을 발전시켜 온 사법부의 결정이라고 주장한다. 학계에서는 이들을 ‘제한적 주권론자’라고 부른다.

글로벌리스트들 중에서 야세민 소이살(Soyсал 1994)은 유럽에서 이주민의 시민권이 발전되고 향상되게 된 원인을 국제규범, 국제기구의 영향력 그리고 초국가적 시민권에서 찾음으로써 국민국가의 영향력이 감소하고 있다고 주장하고 있는 대표적인 학자이다. 소이살은 국민국가 외부에서 오는 국제인권규범의 영향력을 강조한다. 소이살은 초국가적 시민권과 같은 개념을 통해 정치공동체 내에서 권리를 부여하는 기준이 국적에서 보편적인 사람됨으로 변했다고 본다. 시민권을 보편적인 사람됨에 따라 부여함으로써 이주민과 외국인들의 권리를 정당화될 수 있으며, 실제로 이들의 권리는 국제인권규범, 국제법 등에 의해 향상되어 왔다고 주장한다. 야세민 소이살이 국제인권규범이 국민국가에 직접적으로 영향을 미치는데 주목한다면, 에이미 구로위츠(Gurowitz 1999)는 국제인권규범의 국내적 전용 과정에 주목한다. 국제규범이 일본 내 외국인 노동자 단체 등에 의해 어떻게 정부 정책에 도전하는 수단으로 동원되고 있는가를 주목함으로써 국제규범이 국내 정책 특히 이주민 관련 정책에 어떻게 영향을 미치는가를 비서구의 사례를 통해 연구하였다. 소이살과 구로위츠와 같은 글로벌리스트들(globalists)은 국제인권규범의 직접적인 힘 그리고 간접적 전용과정에 각각 주목하고 있지만 공통적으로는 국내적인 요인 보다는 국제규범의 영향력을 강조하는 입장을 가지고 있다.

반면, 크리스찬 욥케(Joppke 1998)는 ‘왜 자유민주주의 국가들은 원하지 않는 이민을 받아들이는가’라는 질문을 던지면서 국제인권규범과 같은 국민국가 외부의 힘 보다는 국내 정치과정에 의해 이민정책이 변화된다고 주장한다. 욥케는 국제인권규범보다는 자기제한적 주권(self-limited sovereignty)이라고 개념화된 국민국가 내 이익단체들의 영향력, 사법부의 영향력과 같은 국내적인 영향력이 실질적으로 이민정책을 변화시키고 있다고 주장하였다. 즉 서구에서 이민정책이 확대될 수밖에 없고, 이민통제를 강화하기 어려운 이유로 2차 대전 이후 인권과 같은 권리의 중요성을 내재화한 사법부의 판결을 들고 있다. 이민수용국 정부가 이민정책을 통제중심으로 전환하고 싶어도 이주민의 인권에 손을 들어주는 사법부에 의해서 제동이 걸린다는 것이다. 이 점에서 자기제한적 주권론자들은 글로

별리스트들처럼 ‘권리’라는 키워드에 기반 한 주장을 펼치지만 국제인권규범과 같은 국민국가 외부의 힘보다는 국민국가 내부의 힘에 주목하면서 국민국가 스스로 주권을 제한하는 측면을 강조하고 있다.

권리에 기반 한 접근법은 국가의 능력을 제한하는 근거가 국민국가 외부에서 오는지 아니면 내부에 내재하고 있는지에 따라 글로벌리스트와 자기제한적 주권론자로 나뉜다. 다시 말해 국민국가 외부로부터 개입하는 국제인권규범이 국내 이민정책 및 이주민 인권에 어떠한 영향을 미치는가, 그리고 그 속에서 국제인권규범 그 자체나 아니면 국내적인 요인 즉 NGOs, 정부부처, 사법부 등의 역할이 더 중요한가 하는 논쟁이 진행되고 있다. 또한 국제인권규범이 더 중요한가 아니면 국민국가 내부의 힘이 더 중요한가 하는 논쟁과 병행하여 국제인권규범이 국내 이민정책에 영향을 미치는 구체적인 메커니즘을 찾아내는 방향으로 연구들이 진행되고 있다.

정치경제학적 이론은 정책의 편익비용분석에 기초한 단순명료한 모델을 제시한다는 점에서 장점을 지니나 다원주의 국가모델을 가정함으로써 단일한 행위자로서의 국가, 이익집단 간의 갈등, 경쟁을 중재하는 심판으로서의 국가를 상정하고 있다는 비판을 받아왔다. 국가 내에는 특정 공공정책을 둘러싼 다양한 이해관계와 상이한 목적을 지닌 부처들이 있고 더 나아가 행정부, 사법부, 입법부 간의 갈등 또한 중앙정부와 지방정부 간의 경쟁이 있음에도 불구하고 국가를 하나의 수동적인 단일 행위자로 보는 것은 이민정책의 결정과정을 분석하는데 있어서 국가의 주체적 측면을 간과했다는 지적이 있었다. 이러한 비판은 대체로 신제도주의적 모델을 지지하는 학자들에 의해 제기되어 왔고 신제도주의적 모델은 위와 같은 정치경제학적 이론의 단일한 행위자로서의 국가에 대한 비판에서 시작한다.

신제도주의적 모델은 홀리필드(James Hollifield), 티치너(Daniel Tichenor), 로젠헝(Zeev Rosenhek) 등에 의해 발전되어 왔다. 신제도주의적 접근은 기본적으로 고객정치 모델이 상정하는 단일한 행위자로서의 국가, 이익집단 간의 이익 추구를 중재하는 심판으로서의 국가 개념을 비판한다. 국가는 하나의 통일체(monolithic entity)가 아니라 다양한 제도로 이루어졌다고 가정한다. 따라서 제

도간의 각축, 시민사회 내 정치적 연합의 형성, 전문가집단의 담론, 국제적 위기 상황(일례로 전쟁으로 인한 난민 급증) 등이 서로 상호작용하여 이민정책에 영향을 미친다고 주장한다(Tichenor 2002).

신제도주의 모델은 이민정책에 있어서 단일한 행위자로서의 국가가 아닌, 국가를 서로 다른 의견, 규범, 정책 목표를 가진 다양한 행위자로 나누어 파악함으로써 국가의 역할을 재조명했다는 평가를 받고 있다. 신제도주의 모델은 행정부 내 다양한 부처가 특정한 정책을 입안하고 추진하는데 있어서 각자의 역할을 확대하기 위해 노력하고 때론 경쟁한다고 가정한다. 부처 간 경쟁이 발생하는 이유는 이익을 최대화하려는 합리적 행위자로서 관료들 상정하고, 이러한 관료들의 집합체로서 부처를 가정한다면 당연한 결과이다. 또한 제도적으로 파악하더라도 부처 간 관할영역이 중복될 가능성이 높기 때문에 중복되는 관할영역을 재조정하는 과정에서 각 부처가 어떠한 형식으로 각자의 고유한 역할을 보장받으면서 동시에 이를 확대하려는 경향 때문에 부처 간 경쟁이 발생할 수 있다. 부처 간 경쟁과정에서 각 부처는 전문성 재고를 통해 관할영역을 자신의 것으로 확보하려고 노력하며 이 과정에서 종종 시민사회의 행위자 또는 전문가들과 연합하기도 한다. 이 점에서 부처 간 경쟁과정은 시민사회 행위자에게 일종의 '정치적 기회구조'가 되기도 하고 이러한 국가와 시민사회의 정치적 연합을 통해 행정부 내의 역할 분담에서 과소평가되어 있는 부처가 자신의 역할을 확대시키려 하기도 한다.

또한 보스웰(Christina Boswell)은 이민정책에 반영되는 국가기능에 초점을 맞추어 신제도주의적 모델을 발전시키려고 하였다(Boswell 2007). 그녀는 우선 국가는 이민정책을 형성하는데 있어 정당성과 통치력을 유지해야 한다고 전제한다. 이런 맥락에서 정당성은 대중의 기대와 가치에 부합하고 조용하는 정치적 행위와 실천이다. 보스웰은 이민정책을 통해 표출되는 국가 정당성을 측정하기 위한 네 가지 기능 즉 안보(security), 축적(accumulation), 공정(fairness), 제도적 정당성(institutional legitimacy)에 주목한다. 국가가 네 가지 기능을 모두 이민정책에 반영하여 추진할 수 있다면 이상적이지만 현실에서는 네 가지 기능을 한꺼번에 추진할 수 없고 각 기능들 사이에 긴장관계가 형성된다. 예를 들면, 국가

안보를 이유로 이민의 문턱을 높일 수도 있고, 축적 을 이유로 이주노동자의 인권을 무시할 수도 있다. 따라서 국가는 위와 같은 긴장관계를 해결하기 위해 네 가지 기능 중 특정 기능을 강조한다는 것이다. 보스웰은 국가가 이민정책을 수립할 때 정당성을 유지하기 위해서는 다양한 사회 집단이나 제도의 이해를 선별적으로 수용하며 위 집단이나 제도의 영향력은 정부의 주관적 관점과 공명할 때만 영향을 미치게 된다고 주장한다.

국제이주를 설명하는 세 가지 방법, 정치경제학적 모델, 국제규범모델, 신제도주의적 모델은 각자 나름의 장점을 지니고 있다. 정치경제학적 모델은 공공정책의 편익비용 분석에 기초한 단순명료한 이론들을 제시하여 검증 가능한 가설들을 만들어내지만, 각 국민국가가 처한 역사적 맥락, 독자적인 역사적 발전 경로를 간과할 수 있다. 신제도주의적 모델은 정책결정을 둘러싼 국가, 시민사회 내 다양한 행위자들의 상호작용과 제도들 간의 각축을 종합적으로 판단할 수 있는 이론들을 제공한다. 하지만 정치경제학적 이론과는 달리 검증 가능한 가설들을 제시하고 이를 테스트하는 데는 한계가 있다고 평가할 수 있다. 권리 기반 접근법 특히 국제규범모델은 이민정책 형성과정을 단순히 국내적인 요인의 역학만으로는 이해하기 어렵다는 점에서 국제적인 변수를 분석할 수 있는 이론을 제공할 수 있다. 하지만 국제규범의 전파경로를 분석하는데 있어서 국내적 수준에서 형성된 이민정책을 분석함에 있어 국제규범의 영향력만을 독립적으로 파악하기 어렵다는 단점이 있다.

IV. 비교연구의 분석단위: 국민국가 모델을 넘어서

이익, 제도, 권리에 기반 한 정치학적인 이론들은 국제이주 연구에 있어서 국제이주가 어떻게 정치적 변동의 원인으로 혹은 결과로 작동하는지를 설명하는 분석적 도구를 제공한다. 국제이주 연구의 정치학적 접근법에서 추론되는 가설, 변수 등은 적합한 연구방법에 의해 검증되어야 한다. 아직까지 국제이주 관련

데이터들이 많이 축적되지 않은 상황에서 통계기법을 활용한 연구들은 저발전된 상황이다. 국제이주에서 정치적 접근법을 취하는 많은 연구들은 단일사례 연구방법에 의존하고 있다. 국내의 선행연구들을 검토해보아도 한국, 일본, 미국, 독일 등 단일사례를 통해 이민정책의 동학을 분석하고 있다. 비교연구방법을 선택하고 있는 연구들도 대부분은 한국과 일본 간 비교연구라고 할 수 있다. 해외의 경우도 정치학 내에서 국제이주가 주요 연구주제로 등장한 시기는 1990년대 이후로 대부분 북미 혹은 유럽의 사례를 기반으로 한 것이다. 비교연구 역시 드문 편이며 미국과 캐나다를 비교하거나 유럽국가들 간의 비교연구가 대부분을 차지한다 (Schain 2012). 그러나 국제이주 연구에 있어 정치학적 접근법과 거기로부터 도출된 가설과 변수들 간의 관계를 검증하기 위해서는 ‘왜 특정 국가에서 발생한 현상이 왜 다른 국가에서는 벌어지지 않는가’하는 연구질문에 답을 제공해야 하며, 이 점에서 국제이주와 관련된 비교연구가 보다 더 진전되어야 한다. 이를 위해 본 장에서는 국제이주 비교연구가 주로 국가 간 비교연구에 기초하고 있음을 비판하는 ‘방법론적 민족주의(methodological nationalism)’논쟁을 소개하고, 국가 간 비교연구에 대한 대안으로 국가 모델을 넘어서는 도시 간 비교연구를 제시하고자 한다.

국제이주 연구에서 정치학적 접근이 비교연구방법을 채택할 때 분석단위는 주로 국민국가이다. 나라마다 상이한 이민정책의 레짐 유형을 크게 전통적 이민국가, 유럽의 이민국가, 그리고 이민후발국가로 나누고 각 유형 내에서 비교대상을 선택하여 비교연구를 진행해왔다. 전통적 이민국가는 이민이 해당 국가의 국가형성과 밀접한 연관성이 있는 국가들로 미국, 캐나다, 호주 등이 이에 해당한다. 이러한 국가들은 이민의 역사가 오래되었으며, 이민이 민족 정체성 형성과 국가형성에 핵심적인 역할을 수행하였다. 유럽의 이민국가는 지역을 기준으로 한 유형에 속한다기 보다 역사적으로 보았을 때 필요에 의해서 이민을 받아들였지만, 이민정책의 의도와는 달리 한시적 이주노동자들의 정주화로 인해 이민 문제가 국가적 차원의 의제가 된 국가들을 지칭한다. 이민후발국가는 앞선 두 유형의 국가들에 비해 뒤늦게 이민 수용국의 대열에 합류한 국가들로 아직까지 이민의 문호를 전면적으로 개방한 이민국가라기 보다는 필요한 산업 분야에 한해 제한

적으로 이민을 수용하는 유형을 의미한다(Hollifield et al. 2004). 이러한 이민정책의 레짐 유형을 기반으로 국제이주를 연구하는 정치학자들은 각 유형 내에서 비교 대상 국가를 선정, 특정 가설을 검증하는 비교연구를 진행해왔다.

분석단위를 국민국가로 설정하여 비교연구를 진행하는 것이 주된 연구방법의 흐름이 된 데는 브루베이커의 연구가 큰 영향을 미쳤다. 브루베이커(Rogers Brubaker)는 *Citizenship and Nationhood in France and Germany* 라는 책에서 독일과 프랑스의 이민정책 차이를 설명하고 있다. 브루베이커에 따르면 독일과 프랑스 이민정책의 차이는 각 국에서 역사적으로 상이하게 발전된 시민권 모델로 설명할 수 있다. 독일과 프랑스는 각각 혈통중심의 시민권 모델과 공화주의적 모델을 발전시켜 왔는데 이 같은 시민권 모델이 두 나라 간의 이민정책 차이를 낳고 있다고 주장한다(Brubaker 1992).

국제이주 연구에서 국민국가를 비교연구의 분석단위로 삼는 연구경향에 대해 일군의 학자들은 비판적인 견해를 제시하고 있다. 국제이주는 국경을 넘나드는 인구의 이동으로 그 속성 상 초국가적인 관계를 내포하고 있는데 국민국가를 분석단위로 삼는 것은 이러한 국제이주의 속성을 정확히 설명하기에 적합하지 않다는 것이다. 이들은 국가를 비교연구의 분석단위로 채택하는 것을 ‘방법론적 민족주의(methodological nationalism)’로 정의하고 이에 대한 비판을 제기하고 있다(Wimmer and Glick Schiller 2002).

위머와 글릭 실러(Andreas Wimmer and Nina Glick Schiller)는 연구의 분석단위를 국민국가로 설정하는 것이 국제이주 연구를 포함한 사회과학 전반에 만연해있다고 주장한다. 방법론적 민족주의는 “민족/국가/사회가 근대 세계의 자연스러운 사회적, 정치적 형식”이라는 가정을 취하고 있다는 것이다(Wimmer and Glick Schiller 2002, 302). 이는 역사적으로 보았을 때 사회과학이 국민국가 형성 프로젝트에 큰 영향을 받았기 때문이며, 그 발전과정 속에서 자연스럽게 정치적, 사회적 현상을 국민국가 단위로 파악한다는 것이다. 위머와 글릭 실러는 방법론적 민족주의의 세 가지 형태를 서술하면서 그 중 하나로 “사회과학적 상상의 영토화와 분석의 초점을 국민국가의 경계로 환원하는 것”을 들고 있다(Wimmer and Glick Schiller 2002, 307). 사회과학은 국민국가의 경계 안에

서 벌어지는 일을 국민국가의 경계 밖과 구분 짓고, 경계 안에서 벌어지는 현상을 동질화시켜 파악한다고 이들은 주장한다. 이것은 일종의 ‘컨테이너 이론’으로 국민국가를 분석단위로 삼으면서 문화, 정체(polity), 경제, 사회집단 등을 국민국가라는 컨테이너 안에 넣고 사고하는 것이다. 이러한 연구 경향은 초국가적 흐름과 세계화로 인한 상호의존성 증대를 속성으로 하는 국제이주 현상과 배치되며, 국제이주 현상을 정확히 설명하기 위한 적합한 분석단위가 아니라는 것이 위머와 글릭 실러의 주장이다.

국제이주의 정치학적 접근법에 있어 비교연구를 진행할 때 국민국가를 분석단위로 하는 것은 여전히 중요한 연구방법이다. 이민정책은 일반적인 수준에서 보았을 때 국민국가에 의해서 수립되며 이민의 통제, 이민자의 통합 역시 여전히 국민국가의 핵심적인 주권 영역에 속하기 때문이다. 그러나 ‘방법론적 민족주의’라는 비판을 일정 부분 수용한다면, 초국가적인 인구 이동의 속성을 보다 엄밀하게 분석하기 위해서는 방법론적 민족주의 외에 대안적인 분석단위들이 모색되어야 한다. 이 점에서 도시를 분석단위로 채택하여 비교연구에서 ‘국가적 모델’을 넘어서는 것은 의미 있는 시도가 될 수 있다.

이민정책은 일반적으로 국가 수준에서 수립된다. 물론 경제적인 요건, 인구학적인 상황에 따라 도시 간, 지역 간 차이가 있을 수 있지만, 중앙정부가 어떤 이민자를 받아들일 것인지, 이민자들의 체류조건, 시민권 요건 등을 결정함에 있어 일차적인 책임을 진다. 그러나 이민자들이 생활을 영위하는데 있어 부딪히는 실질적인 문제들 특히 주거, 보건, 교육 등의 문제는 그들이 구체적으로 거주하는 공간인 지역에서 발생한다. 그래서 많은 경우 국가 수준의 이민정책 목표와 지역에서 요구하는 구체적 정책 사이에 간격이 생겨나며 이러한 이유로 국가 수준의 이민정책과 지역 수준의 이민정책 사이의 조율이 필요하다. 국가 수준의 통일성과 지역 수준의 다양성 간의 조화를 위해서는 이민자들 다수가 거주하고 있는 도시적 맥락의 이민정책을 이론적 틀을 정립할 필요가 있다.

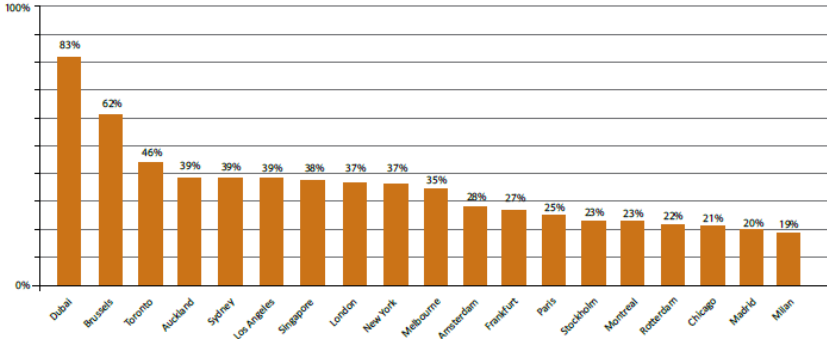
도시적 맥락의 이민정책을 탐색하는 것은 학술적으로도 중요한 의미를 가진다. 앞서 서술했지만 이민연구 특히 비교이민연구에 있어 주된 분석단위는 주로 국가였으며, 브루베이커(Rogers Brubaker)의 연구가 대표적으로(Brubaker 1992)

독일과 프랑스의 이민정책 차이를 각 국에서 역사적으로 상이하게 발전된 시민권 모델에서 찾고 있다. 그러나 이와 같은 방법론적 민족주의 경향에서 탈피하여 분석대상을 분권화하자는 일련의 연구들이 나오기 시작했다. 마르티넬로 등은 여러 유럽 도시들을 분석 대상으로 하여 이민자들의 배제와 통합의 동학을 연구하고자 하였으며, 니콜스(Walter Nicholls)와 위터마크(Justus Uitermark)는 암스테르담과 로스엔젤리스의 인종적 소수자 정치를 비교함으로써 다문화주의 정책의 도시 간 차이를 규명하고자 하였다(Nicholls and Uitermark, 2013). 이민 연구에 있어 도시 간 비교는 특히 이민정책에 있어 주거, 정치 참여, 통합, 이민자단체 등 보다 구체적인 이슈들을 규명하는데 기여할 수 있다.

현재 우리는 많은 사람들이 도시에 거주하는 도시화의 시대에 살고 있다. 예를 들어 2014년에 전 세계 인구의 54%가 도시에 거주하고 있으며, 현재 약 39억 명의 도시거주인구는 2050년에 64억 명에 달할 것으로 예측된다(UN DESA, 2014). 국제이주는 이러한 도시화 현상과 맞물려 더욱 가속되고 있으며, 도시 내에서 문화적, 인종적 다양성을 심화시키고 있다. 이러한 점을 고려하면 초국가적 인구의 흐름인 국제이주의 속성이 더 분명히 드러나는 공간적 단위는 국민국가가 아니라 도시라고 할 수 있다.

〈그림 1〉은 전 세계 주요 도시의 이민자 비율을 보여준다. 두바이, 브뤼셀 등은 예외로 하더라도 런던, 로스엔젤리스, 싱가포르, 시드니 등의 도시는 30% 후반에 해당하는 이민자 비율을 가지고 있다. 주요 도시의 이민자 비율은 이들 도시가 위치한 국가의 평균 이민자 비율을 훨씬 상회하며, 많은 이민자들이 주요 도시를 거점으로 거주하고 있음을 보여준다. 캐나다 이민자의 46%가 토론토에 거주하고 있고, 미국 이민자의 40%가 뉴욕, 로스엔젤리스, 마이애미, 시카고, 샌프란시스코에 살고 있으며, 호주 이민자의 28%가 시드니와 멜버른에 거주하고 있다(IOM, 2015).

〈그림 1〉 주요 도시의 이민자 비율



출처: IOM, 2015

전통적 이민국가의 도시들에서 후발 이민국가의 도시들 특히 아시아 지역의 도시들로 시선을 돌리면 각 도시들 간 이민자 비율은 큰 격차를 보인다. 〈표 2〉는 아시아 지역의 국가와 주요 도시의 이민자 비율을 보여준다.

〈표 2〉 아시아 국가와 주요 도시의 이민자 비율

City	Foreign-born (%)	Country	Foreign-born (%)
Beijing	0.5	China	0.05
Kuala Lumpur	9.0	Malaysia	8.40
Mumbai	1.4	India	0.50
Seoul	3.7	Republic of Korea	2.80
Singapore	38.0	Singapore	38.00
Sydney	39.0	Australia	28.00
Tokyo	3.0	Japan	1.60

출처: IOM, 2015

싱가포르가 38%로 가장 높은 비율을 나타낸 반면, 서울, 도쿄, 쿠알라 룸프르 등은 상대적으로 낮은 비율을 나타낸다. 하지만 서울의 경우, 2004년 약 11만 명의 이민자수가 2014년 27만 명으로 최근 10년 사이 두 배로 증가하였으며 베이징의 경우도 약 50%의 증가세를 보이고 있다. 전통적 이민국가의 도시들에

비해 낮은 이민자 비율을 보이고 있지만, 증가 속도가 빠르다는 것을 고려하면 향후 아시아 지역의 도시들에서 이민자 통합과 관련된 이슈들이 중요하게 대두될 수 있음을 예측할 수 있다. 특히 한국과 일본의 경우 전형적인 후발 이민국가로서 상대적으로 제한적인 이민정책을 실행하고 있기 때문에 이민자 통합이라는 부담이 실제 이민자들이 거주하는 도시에 부과될 가능성이 높다. 따라서 국가 수준의 제한적 이민정책과 도시 수준의 다문화적 현실 사이의 딜레마를 해결하기 위해서라도 국제이주에 대응하는 도시적 맥락의 이민정책을 탐색할 필요가 있고, 이를 위해서는 도시를 분석단위로 하는 비교연구가 보다 많이 진행되어야 한다.

또한 도시를 분석단위로 설정한다면 국제이주의 비교연구 대상을 늘릴 수 있다. 앞서 서술했듯이 국제이주 연구에서 정치학적 접근법은 전통적 이민국가, 유럽의 이민국가, 이민후발국가라는 세 가지 이민정책 레짐 유형 속에서 비교대상을 주로 선택해왔다. 하지만 도시를 분석단위로 채택한다면 레짐 유형을 넘나드는 비교연구가 가능하다. 예를 들어 전통적 이민국가에 속하는 캐나다와 유럽의 이민국가에 속하는 독일을 비교할 때 국가를 분석단위로 설정하면 많은 난관에 봉착하게 된다. 캐나다는 대표적으로 영주이민을 주로 받아 이민자의 정착을 유도하는 국가인 반면, 독일은 최근의 국적법 개정에도 불구하고 오랫동안 이민국가임을 부정해온 이민을 꺼리는 국가였기 때문이다(Schmidtke 2014). 그러나 도시를 분석단위로 이민정책의 다층적 거버넌스의 한 축으로 설정하고 이를 분석한다면, 레짐 유형을 넘어서는 비교대상을 찾을 수 있다. 또한 최근 들어 이민정책이 중앙정부에서 지방정부로 분권화하는 추세 속에서 도시가 중앙정부와는 다른 정책방향을 가진 이민정책을 실험하는 주체임을 고려한다면, 도시를 분석단위로 설정하는 것은 국제이주의 비교연구의 지평을 넓힐 수 있다.

비교연구에서 분석단위로서 도시를 선택하는 것은 국민국가 중심의 비교연구와 반드시 배치되지 않는다. 국민국가를 단위로 하는 연구와 병행하여 진행한다면 보다 정확한 연구결과를 도출할 수 있다. 예를 들어 아이린 블룸라드(Irene Bloemraad)는 *Becoming a Citizen*이라는 책에서 국가 간 비교와 도시 간 비교를 함께 사용하고 있다. 블룸라드는 '왜 미국과 캐나다는 많은 공통점에도 불구

하고 서로 다른 귀화율을 보이는가'하는 연구질문에 답하기 위해 국민국가 수준에서 미국과 캐나다를 비교한 후, 거기서 도출된 해답을 보스톤과 토론토를 분석단위로 다시 검증하는 연구방법을 채택한 바 있다. 이처럼 도시를 분석단위로 하는 연구는 국제이주 연구에 있어 비교대상을 확장시킬 수 있으며, 연구방법을 보다 풍성하게 할 수 있다.

V. 이민자 통합정책 모델에서 지표 중심 접근법으로

국제이주 연구의 정치학적 접근에서 한 국가 혹은 여러 국가를 비교하여 이민자 통합정책을 평가할 때 많은 연구자들은 이민자 통합정책 모델을 상정하고 이를 기반으로 연구를 진행해왔다. 그러나 이 같은 모델은 일종의 이념형으로 보다 구체적으로 한 국가의 이민정책을 평가하기에는 부족한 점이 많다. 예를 들어 프랑스는 대표적으로 동화주의 모형에 속하는 국가이지만, 모든 정책 분야에서 동화주의 모형에서 도출된 특성들이 발견되는 것은 아니다. 교육 분야만 보더라도 전반적으로 인종, 종교 등을 나타내는 상징들을 공적 장소에서 드러내지 못하는 동화주의적 특성이 강하게 나타나지만, 한편으로는 교사들을 대상으로 다양한 다문화교육 프로그램을 실시하는 등 여러 모델이 혼재하고 있다(한승준 2008). 이처럼 이념형으로서 이민자 통합정책 모델에 기반 한 단일사례연구, 비교연구는 이민자 통합정책 평가와 정책의 변화를 설명하는데 있는 한계가 있다. 이를 극복하기 위해 학계는 이민정책을 평가하는 다양한 지표들을 개발하고 지표 중심 접근법을 새로운 분석틀로 삼아 이민자 통합정책 모델에 의한 평가로부터 전환을 모색하고 있다.

이민자 통합정책을 유형화하자면, 흔히 '구분배제모형(differentiated exclusionary model),' '동화모형(assmilationist model),' '다문화주의 모형(multicultural model)'으로 나눌 수 있다. 구분배제모형은 외국인 노동자를 노동력 부족현상이 있는 특정 산업분야에 한해서만 수용하며, 이들에게 사회적, 정

치적 권리를 주지 않는 것을 원칙으로 한다. 노동력 유입국 사회가 이주노동자나 이민자를 3-D직종의 노동시장과 같은 특정 경제적 영역에만 받아들이고, 복지 혜택, 국적(또는 시민권), 선거권·피선거권 부여 또는 국방의 의무 부여와 같은 사회적·정치적 영역에는 절대 받아들이지 않는 것을 말한다. 우리나라의 저숙련 외국인 노동자 정책이 대표적이다. 동화모형은 이민자가 출신국의 문화적인 특성을 포기하고 주류 사회에 완전히 동화되는 것을 목적으로 한다. 프랑스의 이민자 통합정책이 대표적인 것으로 프랑스의 공화주의적 가치를 받아들인다면 주류 성원들과의 차별을 두지 않는 것을 의미한다. 다문화주의 모형은 이민자들이 출신국의 문화와 언어를 ‘집단적으로’ 유지하는 것을 장려하여 다양한 이민자 그룹들 간의 공존을 꾀하는 것이다. 다원적인 사회에서 다양한 하위문화들이 평등하게 타당성을 지니는 것으로 인정된다. 진정으로 다원적인 사회의 발전을 촉진시키는 것이 가장 적합한 통합 방식이라는 것이다. 다문화주의 모형은 정책을 시행할 때 이민자를 한 개인으로 접근하지 않고 특정 그룹의 일원으로 간주하며, 캐나다의 다문화주의 정책이 대표적이다(설동훈 외 2013).

국제이주 연구를 하는 정치학자들은 이민자 통합을 연구할 때 개인 혹은 집단 수준의 통합 정도를 이민 수용국의 제도적 영향력에 초점을 맞추어 설명한다(Triadafilopoulos 2012). 이러한 제도적 영향력을 설명하기 위한 분석틀 중 많이 활용된 것이 위에서 설명한 이민자 통합정책 모델이었다. 그러나 이 같은 이념형에 기반 한 설명은 고용, 교육, 반차별, 국적 취득 등 이민정책의 세부 분야를 자세하게 파악하기 어렵다는 단점이 있다. 따라서 최근 들어서는 이민정책을 지수화하여 이를 기반으로 연구를 진행하자는 움직임이 있다(Triadafilopoulos 2012). 대표적인 것이 이민자 통합 정책 지수인 MIPEX이다.

‘이민자 통합 정책 지수(Migrant Integration Policy Index: MIPEX)’는 여러 나라의 이민자 통합정책을 체계적으로 평가하기 위한 노력의 일환으로, 이민자 통합에 관한 일곱 가지 정책 영역에 걸쳐 총 148개 항목의 세부 지표들을 통해 각국의 이민자 통합정책을 평가하고 있다. MIPEX의 일곱 가지 정책 영역은 ①이민자의 노동시장 이동(labour market mobility), ②이민자의 가족 재결합(family reunion), ③교육(education), ④이민자의 정치참여(political participation), ⑤이

민자의 장기거주(long-term residence), ⑥이민자의 국적 취득 가능성(access to nationality), ⑦이민자에 대한 반차별 조치(anti-discrimination) 등으로 현재 유럽, 북미 국가 그리고 호주, 일본 등 총 33개국이 참여하고 있다.

MIPEX는 각국의 이민정책을 다문화주의모형, 동화모형, 구분배제모형 등 이념 형에 기반 한 유형화에 의해 평가하던 기존의 방식과 달리, 세부 정책 별로 비교 가능한 지표체계를 사용하여 평가한다는 점에서 그 의의가 있다고 하겠다. 특정 국가 이민자 통합정책의 수준을 비교가능하게 하여 정책의 도입, 평가, 그리고 개선을 유도할 수 있다는 장점이 있다. 또한 상세한 지표를 제공함으로써 이민자 통합정책 결정과정에 참여하는 정책결정자, 시민사회, 미디어 등에 이민자 통합정책에 대한 기준을 제시하고, 국가 간 비교에 의한 정책 학습기회를 제공한다.

학술적으로도 이민자 통합 정책 지수와 같은 지표 기반 접근법은 추상적인 논쟁의 반복으로 평행선을 달리는 논의를 경험적 연구 결과물을 통해 넘어설 수 있게 한다. 예를 들어 국제이주 연구에서 큰 논쟁 중 하나는 과연 주요 이민 수용국에서 다문화주의가 쇠퇴하고 있는가 하는 것이다. 크리스찬 욥케(Christian Joppke)는 네덜란드를 시작으로 주요 이민 수용국의 이민자 통합 정책이 다문화주의에서 시민통합(civic integration)으로 전환되고 있다고 주장했다(Joppke 2007). 시민 통합에 기반 한 이민자 통합정책은 다문화주의 모델이 이민자들의 문화적 권리를 “집단”의 수준에 기초한 것에 반해 이민자가 수용국 사회에 적응할 책임을 “개인”에 지우고, 이민자 개개인이 수용국 사회의 언어, 역사, 제도에 대한 기본적인 지식을 습득하는 것이 사회통합에 필수적인 요소라고 보고 있다. 이에 따라 이민자가 수용국에 입국하기 전부터 수용국 언어에 대한 테스트가 이루어지는 경우도 있고, 입국 후에도 계속 일정 시간 언어교육, 사회교육, 직업교육을 이수해야 한다. 이수과정을 불이행할 경우, 벌금이 부과되거나 혹은 영주권을 거부당할 수도 있다. 이러한 교육과정을 통해 수용국의 노동시장에 통합되어야 함을 목적으로 하고 있다. 이러한 시민 통합 정책은 1990년대 후반 네덜란드에서 시작되어 그 이후 핀란드, 덴마크, 오스트리아, 독일, 프랑스 등 유럽의 각 지역으로 퍼져나갔다(Joppke, 2007). 이에 대해 대표적인 다문화주의 정치철학

자인 킴리카(Will Kymlika)는 다문화주의가 쇠퇴하고 있지 않다고 주장했다. 이러한 다문화주의 쇠퇴 논쟁이 이전의 논쟁과 달랐던 점은 그 근거를 지표 기반 접근법에서 도출했다는 것이다. 반팅과 킴리카는 ‘다문화주의 정책 지수(Multiculturalism Policy Index)’에 기반하여 다문화주의가 쇠퇴하고 있다고 유펜케가 주장하는 국가들의 정책을 분석, 그러한 국가들에서 다문화주의는 쇠퇴하고 있지 않으며, 높은 점수를 받은 다문화주의 정책이 더 나은 통합의 결과를 만들고 있다고 주장했다(Banting and Kymlika 2012).

‘이민자 통합 정책 지수’와 ‘다문화주의 정책 지수’외에도 ‘시민 통합 지수(Civic Integration Index),’ ‘시민권 정책 지수(Citizenship Policy Index)’ 등이 개발되어 연구의 영역을 이민자 통합 정책 평가에서 시민권, 복지정책과의 연관성 등으로 확장시키고 있다. 지표 기반 접근법은 개념의 명확성, 측정 방법 등에서 향후 개선되어야 할 점이 많지만(Chun and Yoon 2013), 데이터의 축적을 통해 이민자 태도나 여론 등에 국한되어 있었던 양적 연구 방법을 정책 분야로 확산시킬 것으로 기대된다. 그러나 지표 기반 접근법은 측정의 객관성을 위해 법, 정책 등을 기반으로 하여 정책이 영향을 미치는 이민자들의 의견이나 입장이 반영되지 못했다는 비판이 있다(Triadafilopoulos 2012). 따라서 지표 기반 접근법은 이민자들의 의견과 입장을 분석하기 위한 심층 면접, 인터뷰 기법 등과 함께 혼용되어 쓰일 필요가 있으며, 이 같은 혼합 연구 방법(mixed method)은 향후 국제이주 연구의 방법론을 다양화하는데 기여할 것이다.

VI. 결론

본 논문은 국제적으로 시리아 난민 사태와 포퓰리즘의 확산, 그리고 국내적으로 다문화 사회의 도래 등 국제이주가 더 이상 정치학의 주변부 이슈가 아니라는 판단 하에 국제이주 연구에 있어 정치학적 접근법을 모색하고 정치학적 접근법에서 도출된 가설과 변수들 간의 관계를 파악하기 위한 연구방법에 있어 분석

단위의 문제와 이념형 모델에 의한 정책의 평가 등을 토론하고자 하였다.

국제이주 연구에 있어 정치학은 상대적으로 후발주자이다(Triadafilopoulos 2012). 경제학이 국제이주의 원인과 동기를 분석하는데 중점을 두고, 사회학과 인류학이 이민자 통합을 설명하고 이민이 집단 정체성에 미치는 영향을 연구하는데 오랫동안 심혈을 기울여온 데 비해 정치학은 1990년대 이후 본격적으로 독자적인 접근법을 모색하기 시작했다. 본 논문에서는 국제이주 연구에 있어 정치학적 접근은 이익, 권리, 제도라는 세 가지 키워드를 중심으로 각각 정치경제학적 모델, 권리 기반 모델, 신제도주의 모델로 발전해왔으며, 각 모델은 ‘왜 주요 이민 수용국에서 반 이민 여론에도 불구하고 이민정책은 확대의 방향으로 나아가는가’를 설명하고 노력한다고 보았다.

국제이주 연구에 있어 정치학적 접근법을 검증하기 위해 여러 방법론적 쟁점이 있을 수 있지만 본 논문에서는 비교연구에 있어 국가를 핵심 분석단위로 설정하는 경향과 이러한 경향을 ‘방법론적 민족주의’로 정의하고 비판하는 관점을 소개하였다. 국민국가를 분석단위로 설정하여 국가를 일종의 ‘컨테이너’로 보고 그 속에서 민족/사회/이민자 집단 등을 동질화시키는 접근법은 초국가적 흐름을 특징으로 하는 국제이주 현상과 배치된다는 것이다. 이에 대해 본 논문은 국민국가가 여전히 중요한 분석단위임에도 불구하고, 다른 대안적인 분석단위를 모색해야 한다고 주장한다. 특히 국민국가 모델을 넘어서 도시를 분석단위로 설정한다면 국민국가 하부단위(subnational unit)의 이민정책 연구를 통해 이민정책의 다층성을 파악할 수 있으며, 비교연구 대상을 확장할 수 있다고 보았다. 또한 한 국가의 이민자 통합정책을 평가함에 있어 구분배제모형, 동화모형, 다문화주의 모형 등 일종의 이념형을 분석틀로 활용하는 것을 비판하고 지표 기반 접근법을 통해 이민정책의 세부 분야를 포괄적으로 분석해야 함과 연구방법의 다양화를 모색해야 함을 제안하였다.

본 논문은 국제이주 연구에 있어 정치학적 접근법과 방법론 상 쟁점을 몇 가지에 국한하여 모색한 시론적 연구에 해당한다. 따라서 정치학적 접근법에 있어 기타 연구영역 즉 시민권, 이민자의 투표행태, 이민정책에 있어 정당의 역할 등을 토론하지 못한 한계가 존재한다. 또한 방법론 상 쟁점에 있어서도 보다 구체

적으로 이민정책 관련 지표들을 소개하고, 그 한계를 논하지 못했다. 본 논문에서 다루지 못한 주제에 대해서는 후속 연구를 통해 보완하도록 노력하겠다.

참고문헌

- 신지원. 2011. 『이민정책의 이론적 기초와 방향성 연구』 일산: IOM 이민정책 연구원.
- 이혜경 외. 2016. 『이민정책론』. 서울: 박영사.
- 한건수. 2012. “한국사회의 다문화주의 혐오증과 실패론: 어떤 다문화주의인가?” 『다문화와 인간』 제1권 제1호: 113-143.
- 한승준. 2008. “프랑스 동화주의 다문화정책의 위기와 재편에 관한 연구.” 『한국 행정학보』 제42권 제3호: 463-486.
- Bloemraad, Irene. 2006. *Becoming A Citizen: Incorporating Immigrants and Refugees in the United States and Canada*. Berkeley: University of California Press.
- Boswell, Christina. 2007. "Theorizing Migration Policy: Is There a Third Way?" *International Migration Review* 41(1): 75-100.
- Brettell, Caroline B. and James F. Hollifield eds. 2000. *Migration Theory: Talking Across Disciplines*. New York and London: Routledge.
- Brubaker, Rogers. 1992. *Citizenship and Nationhood in France and Germany*. Cambridge: Harvard University Press.
- Castles, Stephen and Mark J. Miller. 2009. *The Age of Migration: International Population Movements in the Modern World*, 4th Edition. London: The Guilford Press.
- Chun, Kyung Ock and Kwang-Il Yoon. 2013. “MIPEX and Korea: Assessment and Lessons.” *OMNES: The Journal of Multicultural Society* 4(2): 27-58.
- Cornelius, Wayne, A., Takeyuki Tsuda, Philip L. Martin and James F. Hollifield eds. 2004. *Controlling Immigration: A Global Perspective, 2nd Edition*. Stanford: Stanford University Press.

- Freeman, Gary. 1995. "Modes of Immigration Politics in Liberal Democratic States." *International Migration Review* 29(4): 881-902.
- Goodman, Sara Wallace. 2010. "Integration Requirements for Integration's Sake? Identifying, Categorising and Comparing Civic Integration Policies." *Journal of Ethnic and Migration Studies* 36(5): 753-772.
- Gurowitz, Amy. 1999. "Mobilizing International Norms: Domestic Actors, Immigrants, and the Japanese State." *World Politics*. 51(3): 413-445.
- International Organization for Migration. 2015. *World Migration Report 2015*. Geneva: IOM.
- Joppke, Christian. 2007. "Beyond National Models: Civic Integration Policies for Immigrants in Western Europe." *West European Politics* 30(1): 1-22.
- Joppke, Christian. 2012. "Immigrant Integration after Multiculturalism." Paper presented at East-West Dialogue on Multiculturalism. Seoul, Korea.
- Kymlicka, Will. 2012. "Multiculturalism: Success, Failure, and the Future." Paper presented at East-West Dialogue on Multiculturalism. Seoul, Korea.
- Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino and J. Edward Taylor. 1993. "Theories of International Migration: A Review and Appraisal." *Population and Development Review* 19(3): 431-466.
- Nicholls, Water and Justus Uitermark. 2013. "Post-Multicultural Cities: A Comparison of Minority Politics in Amsterdam and Los Angeles, 1970-2010." *Journal of Ethnic and Migration Studies* 39(10): 1555-1575.

- Schain, Martin A. 2012. "The Comparative Politics of Immigration." *Migration and Citizenship: Newsletter of American Political Science Association* 1(1): 481-497.
- Schmidtke, Oliver. 2014. "Beyond National Models?: Governing migration and integration at the regional and local levels in Canada and Germany." *Comparative Migration Studies* 2(1): 77-99.
- Soysal, Yasemin Nuhoglu. 1994. *Limits of Citizenship: Migrants and Postnational Membership in Europe*. Chicago: The University of Chicago Press.
- Tichenor, Daniel. 2002. *Dividing Lines: The Politics of Immigration Control in America*. Princeton: Princeton University Press.
- Triadafilopoulos, Triadafilos. 2012. "From National Models to Indices: Immigrant Integration in Political Science." *Migration and Citizenship: Newsletter of American Political Science Association* 1(1): 22-29.
- United Nations Department of Economic and Social Affairs. 2014. *World Urbanization Prospects: The 2014 Revision*. New York: United Nations.
- Wimmer, Andreas and Nina Glick Schiller. 2002. "Methodological nationalism and beyond: nation-state building, migration and the social sciences." *Global Network* 2(4): 301-334.

Political Science Approach to International Migration Studies and Methodological Issues

Byounga Lee*

ABSTRACT

This article aims to examine political science approach to international migration studies. Unlike economics and sociology, political science lags behind developing its own perspective on international migration. However, political science has developed its own theory based on three keywords such as interests, rights, and institutions. Built on three keywords, political science has specified political economy theory, rights-based theory which puts emphasis on international norms and judicial powers, and neo-institutionalism theory. I also argue that political science needs to go beyond ‘methodological nationalism,’ which mainly takes nation-state as a main unit of analysis. I suggest that we should consider other sub-national units as an alternative unit of analysis in examining the complexity and transnationality embedded in international migration itself. Finally, this paper claims that we need to utilize various indices when we evaluate each state’s immigration policy.

Keywords: International Migration Studies, Political Science Approach, Political Economy, Neo-Institutionalism, International Norms, Methodological Nationalism

투고일: 2017.04.14. 심사일: 2017.04.14. 게재확정일: 2017.04.22.

* University of Seoul

[DOI] <http://dx.doi.org/10.21487/jrm.2017.05.2.1.53>

【연구논문】

SNS 활용과 선거캠페인 효과

장우영*·이관률**

논문요약

소셜(social)의 의미가 함축하듯이 오늘날의 커뮤니케이션은 사적 친교에 토대한 ‘개인 중심성’과 ‘연결망 확장성’의 특징을 확연하게 드러내고 있다. 즉 집단 귀속에서 벗어나 개인 친교를 확장하는 네트워킹 기반의 소통이 일상에서 실시간으로 이루어지고 있는 것이다. 이 연구는 20대 총선을 사례로 SNS 선거캠페인 효과를 분석한다. 방법론적 측면에서 이 연구는 선거운동기간의 SNS(트위터) 빅데이터를 활용해서 캠페인 효과를 경험적으로 분석한다. 구체적으로 이 연구는 후보의 트위터 활동 유형을 분류한 뒤, 선거 결과에 영향을 미치는 후보 특성 변수와 선거구 특성 변수를 조합하여 SNS 효과를 집단간에 비교분석하였다. 이 연구는 실험적인 분석으로서 결과를 일반화하기는 어려우나, 트위터 캠페인이 후보와 선거구 변수와 결합해서 득표율과 당락에 영향을 미칠 수 있다는 것을 보여준다.

주제어: 빅데이터, 선거캠페인, 소셜네트워크, 정치적 효과, 커뮤니케이션

* 대구가톨릭대학교

** 충남연구원

I. 연구문제와 분석의 시각

1990년대 중반 인터넷이라는 글로벌 네트워크가 국경을 넘어 빠른 속도로 확장되기 시작한 이래 뉴미디어 생태계는 눈부시게 다변화되었다. 국내 변화상을 일별하면 2000년대 들어서는 웹이 일상생활 속에 뿌리 깊게 정착되었고, 동시에 큰 규모의 웹사이트의 성장세가 둔화되는 반면 싸이월드와 블로그 같은 1인 미디어가 득세하기 시작하였다. 이어서 2010년대에 들어서는 트위터와 페이스북 등 글로벌 SNS(Social Network Service)를 중심으로 1인 미디어 시장이 재편되었다. 그리고 근래에는 모바일 혁명에 편승하여 개인 취향별로 다양한 1인 미디어가 탑재되면서 다면적으로 실시간 소통을 촉진하고 있다(조희정, 2013).

소셜(social)의 의미가 함축하듯이 오늘날의 커뮤니케이션은 사적 친교에 토대한 ‘개인 중심성’과 ‘연결망 확장성’의 특징을 확연하게 드러내고 있다. 즉 집단 귀속에서 벗어나 개인 친교를 확장하는 네트워킹 기반의 소통이 일상에서 실시간으로 이루어지고 있는 것이다. 이는 종래의 매스미디어 기반의 집체적 소통과는 현격하게 대비되는 것으로, 그만큼 미디어 이용자의 시민 정체성과 사회적 자아(social ego)가 크게 신장되었음을 함의한다(임혁백 외, 2016). 아울러 개인 간 연결과 상호의존을 촉진하는 정보통신기술과 뉴미디어 발전이 이러한 변화에 지대한 역할을 하고 있다는 점에는 달리 생각할 여지가 없을 것이다.

이런 맥락에서 미디어 기술과 이용자 권능(empowerment)의 발전이 정치과정에서 어떤 역할을 하고 있는지는 사회과학의 중요한 논제로 여겨진다. 일찍이 샤피로(Shapiro, 1999)가 두 요인이 사회정치적 변화에 미치는 파급을 ‘통제혁명’(control revolution)이라고 주장했듯이, 뉴미디어 환경에서는 대중과 분리되었던 기성 사회 엘리트들이 대중의 통제-감시와 상호작용 등- 아래 정치과정을 재구조화하게 된다. 이 과정에서 시민참여는 필수불가결한 요소인데, 그 범위와 방식은 기술과 정치의 상호 수용성 수준에 달려 있다.

이후 SNS가 부상하는 환경에서 카스텔(Castells, 2009)은 이런 통제혁명의 흐름을 네트워크 개인주의(networked individualism)로 이념화했고, 브런스(Bruns,

2005)는 통제혁명의 핵심 기능을 게이트와칭(gate-watching)이라고 명명했다. 즉 고립된 개인에서 벗어나 네트워크로 연결된 비판적 시민들이 통제혁명의 주체로서, 기성 언론과 대의정치 집단의 게이트키퍼(gate-keeping)을 상쇄하는 게이트워킹을 통해 주도적으로 정치과정의 재구조화와 사회변화를 추동해간다는 것이다. 특히 셔키(Shirky, 2010)는 웹 기반의 네트워크에서 개인 기반의 소셜네트워크로 무게중심이 이동함으로써, 개인 수준의 참여가 활발해지고 행위의 효능감이 유례없이 증대하였다고 평가한다.

그렇다면 SNS 매개 정치의 순기능은 구체적으로 어떻게 나타나고 있을까? 이원태(2010)는 해외의 SNS 선거캠페인 사례분석을 통하여 다음과 같은 함의를 제공하고 있다. 첫째, SNS는 언제 어디서나 정치적 소통이 가능한 환경을 제공하고 있다는 점이다. 특히 SNS는 개인이 정치적 의견을 표명할 수 있는 기회와 여건을 제공할 뿐만 아니라, 자신과 유사한 이념이나 이해를 가진 사람들을 효율적으로 연결하고 조직화하는 비용을 크게 절감시킨다. 둘째, SNS를 매개로 이용자들이 감정 공유를 토대로 사회이슈 형성을 촉진하는 이른바 정치의 일상화를 보편화시킬 수 있다. 생활이슈 중심의 의사소통 방식은 정치참여 주도권을 기존 제도정치와 대중매체에서 일반 시민과 유권자들로 옮겨가게 만들 수 있다. 셋째, 선거 공간에서 유권자들의 자발적 정치참여는 아래로부터의 의제설정을 통해 선거정치를 재편할 수 있는 영향력을 가진다. 이 경우 집단분극화(group bipolarization)의 문제가 동반되기도 하지만, 공적 사안에 대한 시민 관심도와 효능감이 크게 증대하게 된다.

SNS 매개 정치에서 매우 큰 관심사 중 하나가 선거캠페인이다. 특히 선거결과에 대한 SNS의 실증적 효과 여부는 크게 이목을 잡아당긴다. 즉 전통적인 선거결과 결정요인의 범주에 SNS가 포함될 수 있는지, 나아가 그렇다면 그 영향력의 강도는 어떤 수준인지가 분석주제로 탐문될 수 있을 것이다. 다시 말해서 기존의 선거 연구가 주로 후보와 선거구 특성을 주된 요인으로 고려해왔다는 점에서, 유권자 참여와 연동된 SNS 캠페인 전략이 선거 결과에 미치는 영향력 여부는 상당히 흥미로운 논제로서, 선거 연구의 지평을 확대하는데 기여할 수 있을 것이다(장우영·류석진, 2013; 장우영, 2014).

아울러 SNS 캠페인 효과 분석은 빅데이터(big data) 환경에 대한 학문적 적응과 적절한 활용을 요한다. 대개의 경험적 선거 연구들이 유권자 설문조사와 같은 응답자 진술의 맥락에서 이루어지고 있는 것이 현실이다. 가령 대표적인 연구들(하승태, 2012; 이창호·정낙원, 2014; 도묘연, 2015; 박세미·황하성, 2016)을 살펴보면 주로 유권자의 SNS 이용도나 신뢰도를 변수로 활용하는 차원에서 분석이 이루어지고 있다는 것을 알 수 있다. 이런 점에서 데이터 기반의 경험적 분석의 중요성이 강조될 수 있다. SNS는 개인과 개인 간의 데이터 유통의 핵심 채널로서, 실시간의 빅데이터 흐름과 활용 속에서 SNS 캠페인이 이루어지고 있다는 점에 주목할 필요가 있다. 따라서 빅데이터를 분석 대상으로 포착하고 변수로 조작함(operating)으로써 후보와 유권자를 가교하는 빅데이터 분석의 효용성을 제고할 수 있다. 첨언하면 국내 사회과학계의 빅데이터 분석이 주로 추세분석에 머물러 있다는 현실도 적절한 빅데이터 분석모형의 탐구를 고무하고 있다.

이 연구는 20대 총선을 사례로 SNS 선거캠페인 효과를 분석한다. 이 연구는 선거운동 기간에 발생한 트위터 빅데이터를 활용해서 캠페인 효과를 실증적으로 분석하는데서 그 의의를 찾을 수 있다. 특히 장우영·류석진(2013) 및 장우영(2017)외에 이와 관련한 선행연구를 발견할 수 없다는 점에서, 이 연구는 빅데이터 기반의 경험적 분석을 발전시키는 단초가 될 것으로 기대한다. 이 연구의 분석절차는 다음과 같다. 우선 20대 총선에서 트위터 캠페인에 참가한 후보들의 빅데이터를 수집하여 플랫폼별로-트윗, 활동기간, 팔로어, 리트윗- 정량화하였다. 다음으로 이 중 후보 캠페인에 해당하는 트윗과 활동기간을 기준으로 SNS 캠페인 활동을 유형화하였다. 부연하면 트윗과 활동기간은 균일한 분포를 보인 반면, 유권자 캠페인에 해당하는 팔로어와 리트윗은 균일성이 확보되지 않아 유형화의 의미를 찾을 수 없었다. 이어서 각 유형별로 후보 특성 변수와 선거구 특성 변수를 조합하여 SNS 효과를 집단간 비교분석하였다. 후보 특성 변수로는 현직 여부, 소속 정당, 정당 규모, 이념, 소득수준을 활용하였다. 그리고 선거구 특성 변수로는 수도권 여부, 지역주의 여부, 재정자립도를 활용하였다. 요컨대, SNS 활동 유형을 이러한 두 차원의 전통적인 변수군과 개별적으로 조합하여 득표율과 당선여부에 미치는 영향을 집단간 비교분석하였다.

II. 분석전략

1. 20대 총선 트위터 캠페인 현황

20대 총선에서 트위터 캠페인에 참가한 후보들은 612명으로 집계되는데 그 특징을 요약하면 다음과 같다. 첫째, 이 중 수도권 선거구 입후보자는 329명, 비수도권 입후보자는 283명으로 나타났다. 이는 비수도권에 비해 수도권 후보들이 상대적으로 트위터 캠페인 참여가 활발하였음을 보여준다. 둘째, 정당의 경우 더민주당(206명), 새누리당(186명), 국민의당(103명), 정의당(25명) 순으로 나타나 양대 정당과 제3당 중심으로 트위터 캠페인이 수행되었음을 알 수 있다. 셋째, 당선자수를 기준으로 양대 정당의 점유율이 각각 46.5%(더민주당)와 39.0%(새누리당)로 나타나 트위터의 선거캠페인 활용과 수혜가 이들에게 편중되었음을 알 수 있다. 이에 반해 국민의당이 호남지역 후보들을 중심으로 활발하게 트위터 캠페인을 수행했지만, 전체적으로 양당 외의 군소정당들과 무소속의 트위터 캠페인 활동과 수혜는 저조하였다.

〈표 1〉 20대 총선 트위터 캠페인 현황

소속정당		트위터 캠페인 참가 후보	
		후보수(%)	당선자수(%)
새누리당	수도권	99 (16.1)	31 (13.9)
	비수도권	87 (14.2)	56 (25.1)
더민주당	수도권	113 (18.4)	78 (34.9)
	비수도권	93 (15.1)	26 (11.6)
국민의당	수도권	66 (10.7)	2 (0.8)
	비수도권	37 (6.0)	18 (8.0)
정의당	수도권	12 (1.9)	1 (0.4)
	비수도권	13 (2.1)	1 (0.4)

기타 정당	수도권	19 (3.1)	0 (0.0)
	비수도권	8 (1.3)	0 (0.0)
무소속	수도권	20 (3.2)	2 (0.8)
	비수도권	45 (7.3)	8 (3.5)
합계		612 (100.0)	223 (100.0)

2. 비교집단의 유형화와 분석방법

SNS 캠페인 집단을 유형화할 수 있는 변수는 후보자 캠페인을 지시하는 (indicating) 선거트윗수와 트위터 활동기간, 그리고 유권자 캠페인을 나타내는 선거팔로어수와 리트윗수로 분류할 수 있다. 이 네 가지 변수들의 평균값과 중앙값, 그리고 분위수를 살펴보면 다음과 같다. 우선 후보자 측면의 선거트윗수의 평균값은 37.96회, 중앙값은 17.00회이다. 그리고 트위터 활동기간의 평균값은 53.19개월, 중앙값은 60.00개월이다. 이 변수들의 분위수를 살펴보면, 중앙값을 기준으로 최대값과 최소값 양측으로 균등한 분포를 이루고 있음을 알 수 있다, 다음으로 유권자 측면의 선거팔로어수의 평균값은 158.41회, 중앙값은 19.00회이다. 그리고 리트윗수의 평균값은 1,179.47회, 중앙값은 30.50회이다. 이 변수들의 분위수를 살펴보면, 대부분의 값들이 최대값을 중심으로 분포되어 있어 분포가 한 쪽으로 편중되어 있음을 알 수 있다.

〈표 2〉 트위터 변수의 평균값, 중앙값 및 분위수

구분	평균값	중앙값	분위수		
			최대값	574.0	
후보 캠페인	선거트윗수(회)	37.96	17.00	75%	45.0
				50%	17.0
				25%	0.0
				최소값	-104.0
				트위터	53.19

	활동기간(월)			75%	70.0
				50%	60.0
				25%	49.0
				최소값	0.0
유권자 캠페인	선거팔로어수(회)	158.41	19.00	최대값	7,759.0
				75%	97.5
				50%	19.0
				25%	3.0
	리트윗수(회)	1179.47	30.50	최대값	53,517.0
				75%	450.0
				50%	30.5
				25%	0.0
				최소값	0.0

일반적으로 특정 변수를 기준으로 집단을 유형화할 경우 집단별로 균등하게 표본이 배분될 수 있도록 해야 한다. 왜냐하면 집단 간의 평균 혹은 빈도의 차이를 비교할 때, 각 집단의 규모가 유사해야 하기 때문이다. 이러한 조건을 고려하면, 집단의 유형화는 평균값과 중앙값의 차이가 적고, 중앙값 보다 높은 값과 낮은 값이 균일하게 분포되어 있는 선거트윗수와 트위터 활동기간을 기준변수로 설정하는 것이 바람직할 것이다. 따라서 이 연구에서는 선거트윗수 20회를 기준으로 집단을 유형화하고, 트위터 활동기간은 60개월을 기준으로 집단을 유형화하고자 한다. 이 분류 기준은 평균값과 중앙값, 그리고 개별 변수의 분포를 고려하여 연구자가 설정한 것이다. 이 기준으로 집단을 구분하면, <표 2>과 같은 집단의 유형으로 구분되고, 각 집단의 유형별 표본의 분포는 비교적 균일하게 분포되어 있음을 알 수 있다. 즉 선거트윗수가 많고 트위터 활동기간이 긴 제1유형(n=140), 선거트윗수가 적고 트위터 활동기간이 긴 제2유형(n=171), 선거트윗수가 많고 트위터 활동기간이 짧은 제3유형(n=143), 선거트윗수가 적고 트위터 활동기간이 짧은 제4유형(n=158)으로 구분된다.

〈표 3〉 비교집단의 유형화

n=612		선거트윗수		
		20회 이상	20회 미만	
트위터 활동기간	6개월 이상	n=140	n=171	n=311
		〈제1유형〉	〈제2유형〉	
	6개월 미만	n=143	n=158	n=301
		〈제3유형〉	〈제4유형〉	
		n=283	n=329	

이 연구에서는 SNS 활동 유형과 후보자 및 선거구의 특성을 종합적으로 반영해서 이 변인들이 득표율(비율척도)과 당선여부(명목척도)에 미치는 영향을 분석하고자 한다. 따라서 득표율은 분산분석(집단 간의 평균 차이 분석)을 수행하고 사후검증을 위해 던칸 테스트(duncan test)를 병행하고자 한다. 그리고 당선여부는 χ^2 분석(집단 간의 빈도 차이 분석)을 수행하고자 한다. 분석의 주요 내용은 다음과 같다.

〈표 4〉 후보자 및 선거구 특성 변수

후보자 특성	선거구 특성
현직 여부 (19대 현직 / 19대 비현직)	수도권 여부 (수도권 / 비수도권)
소속 정당 (새누리당 / 기타)	지역주의 여부 (영호남 / 기타)
정당 규모 (대정당 / 군소정당)	재정자립도 (30% 이상 / 30% 미만)
이념 (보수 / 진보)	
소득 (10억원 이상 / 10억원 미만)	

우선 후보자 특성에서는 현직, 여야, 소득, 정치이념, 정당규모의 측면에서 분석하고자 한다. 우선 현직에서는 19대 국회에서 현직 의원과 아닌 경우로 구분하고, 여야에서는 새누리당과 그 외 정당으로 구분한다. 그리고 소득에서는 후보자의 재산이 10억원 이상인 경우와 10억원 미만인 경우로, 이념은 보수와 진보, 정당규모는 양대 거대정당(새누리당과 민주당) 및 그 외의 군소정당으로 구분한

다. 선거구 특성에서는 선거지역은 수도권 지역과 비수도권 지역으로, 재정자립도는 30% 이상인 지역구와 30% 미만인 지역구로, 그리고 지역주의는 영호남 지역주의와 그 외의 지역으로 구분해 분석하기로 한다.

이러한 접근방법은 SNS의 활동유형과 후보자의 특성 및 선거구 특성이 결합하여 득표율과 당선여부에 어떠한 영향을 미치는지 확인할 수 있다는 측면에서 의의가 있다. 즉 단순히 SNS 활동이 득표율과 당선여부에 영향을 미친다는 것이 아니라, 후보자의 특성과 선거구 특성을 고려할 때, 어떤 조건에서 득표율과 당선여부에 영향을 미치는지를 세부적으로 확인할 수 있다는 장점이 있다.

한편 소득과 재정자립도의 구분한 기준을 살펴보면 다음과 같다. 우선 소득의 평균값이 2,699.0백만원이고, 중앙값이 770.5백만원이다. 따라서 소득의 1,000백만원을 기준으로 집단을 구분하는 것이 적절할 것이다. 그리고 재정자립도의 경우 평균값이 29.56%이고, 중앙값이 24.40%이기 때문에 30%를 기준으로 집단을 구분하는 것이 적절할 것으로 판단된다.

〈표 5〉 소득과 재정자립도의 평균값, 중앙값 및 분위수

구분	평균값	중앙값	분위수	
			최대값	263,773.0
소득(백만원)	2,699.0	770.5	75%	1,803.2
			50%	770.5
			25%	352.4
			최소값	-1,418.0
			최대값	60.0
재정자립도(%)	29.56	24.40	75%	39.6
			50%	27.4
			25%	19.5
			최소값	5.6
			최대값	60.0

III. SNS 활동 유형별 집단간 비교분석

SNS를 활용한 후보자의 평균 득표율은 33.67%이고, 제1유형의 득표율은 38.84%, 2유형은 31.93%, 3유형은 32.86%, 그리고 4유형은 31.76%이다. 분산 분석의 결과에 의하면, 제1유형은 다른 유형에 비해서 득표율이 높은 것으로 나타나지만, 2유형, 3유형, 4유형 간에는 통계적으로 유의미한 차이를 확인할 수 없다($F=5.71$, $prob=0.0007$).

한편 SNS 유형별로 당선여부를 살펴보면, 제1유형에서는 당선자의 분포(75명)가 높다. 그러나 제2유형과 제3유형과 제4유형의 경우는 낙선자(각각 109명, 102명, 113명)의 분포가 상대적으로 더 높게 나타남을 알 수 있다($\chi^2=25.7905$, $df=3$, $prob=0.0001$).

이상을 종합해 볼 때, 단순히 트위터 활동기간이 많아야 할 뿐 아니라, 선거 트윗수가 많아야 득표율과 당선여부에 유의미한 영향을 줄 수 있다고 판단할 수 있다. 즉 트위터 활동기간이 짧거나 선거트윗수가 많지 않은 경우에는 득표율과 당선여부에 영향을 미치지 않음을 확인할 수 있다. 따라서 제1유형의 SNS 활동만이 득표율과 당선여부를 결정짓는 유효한 전략이라고 할 수 있을 것이다.

〈표 6〉 SNS 유형별 득표율 및 당선여부

(단위: %, 명)

구분	득표율	당선여부		
		전체	당선	낙선
전체	33.67	612 (100.0)	223 (100.0)	389 (100.0)
1유형	38.84 ^a	140 (22.9)	75 (33.6)	65 (16.7)
2유형	31.93 ^b	171 (27.9)	62 (27.8)	109 (28.0)
3유형	32.86 ^b	143 (23.4)	41 (18.4)	102 (26.2)

4유형	31.71 ^b	158 (25.8)	45 (20.2)	113 (29.0)
F값/ χ^2 (prob)	5.71(0.0007)	$\chi^2=25.7905$, df=3, prob=0.0001		

1 후보자 특성 요인

1) 현직

19대 현직 여부(2개 집단)와 SNS 유형(4개 집단)을 통합해서 모두 8개 집단으로 구분했다. 현직의 경우 SNS 유형에 따른 득표율의 차이가 없는 것으로 나타났다. 반면, 현직이 아닌 경우 제1유형(33.83%)이 제2유형(24.62%)과 제4유형(26.16%)에 비해서 득표율이 높은 것으로 나타났다. 다만 제3유형(29.54%)은 SNS 유형에 따른 득표율의 차이가 없는 것으로 나타났다(F=32.18, prob=0.0001).

이러한 경향은 당선여부에서도 동일하게 나타나고 있다. 즉 현직인 경우는 SNS 유형에 따라 당선여부의 차이가 없는 반면, 비현직인 경우는 낙선의 빈도가 상대적으로 높은 것으로 나타났다($\chi^2=171.251$, df=7, prob=0.0001).

이상을 종합해 볼 때, 득표율과 당선여부에는 SNS 활동유형 보다는 현직/비현직이 더 큰 영향을 미치는 것으로 판단할 수 있다. 다만 비현직인 경우 트위터 활동기간이 길고, 선거트윗수가 많은 경우가 그렇지 않은 유형에 비해서 득표율이 높게 나타났다.

〈표 7〉 현직 유형과 SNS 유형별 득표율 및 당선여부

(단위: %, 명)

구분	득표율	당선여부		
		전체	당선	낙선
전체	33.67	612	223	389

			(100.0)	(100.0)	(100.0)
19대 현직	1유형	45.91 ^a	58 (9.5)	44 (19.7)	14 (3.6)
	2유형	45.80 ^a	59 (9.6)	45 (20.2)	14 (3.6)
	3유형	46.52 ^a	28 (4.6)	20 (9.0)	8 (2.1)
	4유형	49.23 ^a	38 (6.2)	26 (11.7)	12 (3.1)
19대 비현직	1유형	33.83 ^b	82 (13.4)	31 (13.9)	51 (13.1)
	2유형	24.62 ^c	112 (18.3)	17 (7.6)	95 (24.4)
	3유형	29.54 ^{bc}	115 (18.8)	21 (9.4)	94 (24.2)
	4유형	26.16 ^c	120 (19.6)	19 (8.5)	101 (26.0)
F값/ χ^2 (prob)		32.18(0.0001)	$\chi^2=171.251$, df=7, prob=0.0001		

2) 소속 정당: 여야

여당인 경우 SNS 활동유형에 따라 득표율의 통계적 차이가 없는 것으로 나타났다. 그러나 야당인 경우는 SNS 활동유형에 따라 득표율의 통계적 차이가 일부 나타났다. 즉 야당인 경우 제1유형(37.49%)은 다른 유형에 비해서 득표율이 높게 나타났고, 제3유형(30.90%)이 제2유형(24.48%)에 비해서 득표율이 높은 것으로 나타났다(F=18.98, prob=0.0001).

이러한 경향은 당선여부에서도 동일하게 나타나고 있다. 즉 여당인 경우는 SNS 활동유형과 관계없이 모두 당선의 빈도가 높은 반면, 야당인 경우는 SNS 활동의 제1유형인 경우만 당선의 빈도(61명)가 높은 것으로 나타났고, 그 외의 경우는 모두 낙선의 빈도가 높은 것으로 나타났다($\chi^2=58.2677$, df=7, prob=0.0001).

여기서 한 가지 흥미로운 점은, 야당의 경우는 제3유형의 득표율이 제2유형보다 높다는 것이다. 따라서 야당 후보자의 경우는 트위터 활동기간이 길어야 하지만, 선거트윗수가 활동기간에 비해서 더 중요함을 보여주고 있다.

이상을 종합해 볼 때, 여당인 경우는 SNS 활동이 득표율과 당선여부에 크게 영향을 미치지 않는 것으로 나타났다. 다만 야당인 경우 트위터 활동기간이 길 수록, 그리고 선거트윗수가 많을수록 득표율과 당선의 가능성이 높다고 할 수 있다.

〈표 8〉 여야 유형과 SNS 유형별 득표율 및 당선여부

(단위: %, 명)

구분		득표율	당선여부		
			전체	당선	낙선
전체		33.67	612 (100.0)	223 (100.0)	389 (100.0)
여당	1유형	42.71 ^{ab}	36 (5.9)	14 (6.3)	22 (5.7)
	2유형	41.70 ^{ab}	74 (12.1)	37 (16.6)	37 (9.5)
	3유형	40.94 ^{ab}	28 (4.6)	10 (4.5)	18 (4.6)
	4유형	45.01 ^a	48 (7.8)	25 (11.2)	23 (5.9)
야당	1유형	37.49 ^b	104 (17.0)	61 (27.4)	43 (11.1)
	2유형	24.48 ^d	97 (15.9)	25 (11.2)	72 (18.5)
	3유형	30.90 ^c	115 (18.8)	31 (13.9)	84 (21.6)
	4유형	25.91 ^{cd}	110 (18.0)	20 (9.0)	90 (23.1)
F값/ χ^2 (prob)		18.98(0.0001)	$\chi^2=58.2677$, df=7, prob=0.0001		

3) 정당 규모

새누리당과 민주당과 같은 대정당의 경우는 SNS 활동유형에 따라 득표율의 통계적 차이가 없는 것으로 나타났다. 다만 군소정당의 경우 제1유형(25.91%)이 다른 유형에 비해서 득표율이 높은 것으로 나타났다(F=65.12, prob=0.0001).

그리고 당선여부에서는 대정당의 경우는 제1유형(61명)과 제2유형(54명)에서

당선 빈도가 높게 나타난 반면, 제3유형(37명)과 제4유형(39명)에서는 상대적으로 낮게 나타났다. 그러나 군소정당의 경우는 모든 유형에서 낙선의 빈도가 높게 나타났다($\chi^2=98.8454$, $df=7$, $prob=0.0001$).

이상을 종합해 볼 때, 대정당의 경우는 SNS 활동유형에 따라 득표율은 차이가 없으나, 제1유형과 제2유형의 전략을 취하는 것이 당선가능성을 높일 수 있을 것이다. 반면 군소정당의 경우는 제1유형이 득표율을 높일 수는 있지만, 실제 제1유형의 SNS 활동을 하더라도 당선가능성은 낮은 것으로 판단된다.

〈표 9〉 정당규모 유형과 SNS 유형별 득표율 및 당선여부

(단위: %, 명)

구분		득표율	당선여부		
			전체	당선	낙선
전체		33.67	612 (100.0)	223 (100.0)	389 (100.0)
대정당	1유형	44.19 ^a	99 (16.2)	61 (27.4)	38 (9.8)
	2유형	40.89 ^a	108 (17.7)	54 (24.2)	54 (13.9)
	3유형	39.50 ^a	99 (16.2)	37 (16.6)	62 (15.9)
	4유형	43.31 ^a	86 (14.1)	39 (17.5)	47 (12.1)
군소 정당	1유형	25.91 ^b	41 (6.7)	14 (6.3)	27 (6.9)
	2유형	16.58 ^c	63 (10.3)	8 (3.6)	55 (14.1)
	3유형	17.92 ^c	44 (7.2)	4 (1.8)	40 (10.3)
	4유형	17.86 ^c	72 (11.8)	6 (2.7)	66 (17.0)
F값/ χ^2 (prob)		65.12(0.0001)	$\chi^2=98.8454$, $df=7$, $prob=0.0001$		

4) 이념

정치이념이 보수인 경우는 SNS 활동유형에 따른 득표율의 통계적 차이가 없

는 것으로 나타났다. 다만 진보인 경우는 제1유형(37.76%)이 다른 유형에 비해서 득표율이 높은 것으로 나타났다. 그리고 진보인 경우는 야당의 경우에서도 동일하게 제3유형(31.29%)이 제2유형(24.68%)에 비해서 득표율이 높은 것으로 나타났다. 그러나 제2유형과 제4유형(25.21%) 간에는 득표율의 통계적 차이가 없는 것으로 나타났다($F=14.44$, $prob=0.0001$).

당선여부에 있어서는 보수는 제2유형(40명)이 당선가능성이 가장 많은 것으로 나타났다. 반면 제1유형(15명), 제3유형(10명), 제4유형(28명)의 경우 당선 빈도가 상대적으로 낮았다. 그리고 진보의 경우는 제1유형(60명)에서 당선 빈도가 가장 높게 나타났고, 그 외 유형에서는 낙선의 빈도(각각 67명, 80명, 76명)가 높게 나타났다($\chi^2=52.9802$, $df=7$, $prob=0.0001$).

이상을 종합해 볼 때, 정치이념 상 보수는 어떠한 SNS 활동유형을 갖더라도 득표율에서는 통계적으로 유의미한 차이가 없고, 다만 당선을 위해서는 제2유형의 전략을 갖는 것이 바람직할 것으로 판단된다. 반대로 진보의 경우는 SNS 활동에 따라 득표율이 통계적으로 차이가 있는 것으로 나타났고, 당선을 위해서는 제1유형의 전략을 갖는 것이 합리적인 것으로 판단된다.

〈표 10〉 정치이념 유형과 SNS 유형별 득표율 및 당선여부

(단위: %, 명)

구분		득표율	당선여부		
			전체	당선	낙선
전체		33.67	612 (100.0)	223 (100.0)	389 (100.0)
보수	1유형	41.52 ^a	40 (6.5)	15 (6.7)	25 (6.4)
	2유형	39.80 ^a	82 (13.4)	40 (17.9)	42 (10.8)
	3유형	38.33 ^a	32 (5.2)	10 (4.5)	22 (5.7)
	4유형	41.01 ^a	65 (10.6)	28 (12.6)	37 (9.5)
진보	1유형	37.76 ^a	100 (16.3)	60 (26.9)	40 (10.3)

2유형	24.68 ^c	89 (14.5)	22 (9.9)	67 (17.2)
3유형	31.29 ^b	111 (18.1)	31 (13.9)	80 (20.6)
4유형	25.21 ^c	93 (15.2)	17 (7.6)	76 (19.5)
F값/ χ^2 (prob)	14.44(0.0001)	$\chi^2=52.9802$, df=7, prob=0.0001		

5) 소득수준

후보자의 소득이 10억원 이상인 경우는 SNS 활동유형에 따라 득표율의 통계적 차이가 없는 것으로 나타났다. 다만 소득이 10억원 미만인 경우는 제1유형의 경우 득표율(36.29%)이 그 외 유형보다 높은 것으로 나타났다. 그리고 소득이 10억원 미만이 경우는 제2유형(26.12%), 제3유형(30.09%), 제4유형(27.45%) 간의 득표율이 통계적으로 유의미한 차이가 없는 것으로 나타났다(F=10.14, prob=0.0001).

이는 당선여부에서도 유사한 경향을 보이고 있다. 후보자 소득이 10억원 이상인 경우는 제1유형(32명), 제2유형(39명)에서 당선의 빈도가 높게 나타났고, 반면 제2유형(37명)과 제4유형(40명)에서는 낙선의 빈도가 높다. 그러나 10억원 미만인 경우는 제1유형에서 당선과 낙선의 빈도가 동일하였고, 그 외의 유형에서는 모두 낙선의 빈도가 높은 것으로 나타났다($\chi^2=48.8635$, df=7, prob=0.0001).

이상을 종합해 볼 때, 후보자의 소득이 10억원 이상인 경우는 SNS 활동유형이 득표율과 당선여부에 크게 영향을 미치지 않는 것으로 나타났다. 다만 10억원 미만인 경우는 제1유형이 다른 유형에 비해서 득표율과 당선가능성을 높일 것으로 판단된다.

〈표 11〉 소득 유형과 SNS 유형별 득표율 및 당선여부

(단위: %, 명)

구분		득표율	당선여부		
			전체	당선	낙선
전체		33.67	612 (100.0)	223 (100.0)	389 (100.0)
10억 이상	1유형	42.89 ^a	54 (8.8)	32 (14.4)	22 (5.7)
	2유형	39.19 ^{ab}	76 (12.4)	39 (17.5)	37 (9.5)
	3유형	35.84 ^b	69 (11.3)	26 (11.7)	43 (11.1)
	4유형	37.97 ^{ab}	64 (10.5)	24 (10.8)	40 (10.3)
10억 미만	1유형	36.29 ^{ab}	86 (14.1)	43 (19.3)	43 (11.1)
	2유형	26.12 ^c	95 (15.5)	23 (10.3)	72 (18.5)
	3유형	30.09 ^c	74 (12.1)	15 (6.7)	59 (15.2)
	4유형	27.45 ^c	94 (15.4)	21 (9.4)	73 (18.8)
F값/ χ^2 (prob)		10.14(0.0001)	$\chi^2=48.8635$, df=7, prob=0.0001		

2. 선거구 특성

1) 수도권 여부

선거지역을 수도권과 비수도권으로 구분해 보면, 수도권에서는 SNS 활동유형에 따라 득표율의 통계적 차이가 있는 반면, 비수도권에서는 득표율의 통계적 차이가 없는 것으로 나타났다. 수도권의 경우 제1유형(37.95%)은 다른 유형에 비해서 득표율이 높게 나타났고, 그 외 유형에는 통계적 차이가 없는 것으로 나타났다(F=4.52, prob=0.0001).

한편 선거지역에 따른 당선여부를 살펴보면, 수도권의 경우 제1유형만

당선의 빈도(49명)가 높은 반면, 그 외 유형에서는 당선의 빈도가 상대적으로 낮았다. 이는 비수도권에서도 동일한 형태를 보이고 있다($\chi^2=33.2841$, $df=7$, $prob=0.0001$).

이상을 종합해 볼 때, 제1유형의 SNS 활동이 수도권에서만 득표율에 영향을 미치는 것으로 나타났다. 그리고 제1유형은 선거지역에 관계없이 모두 당선의 빈도를 높이는 것으로 나타났다. 따라서 제1유형은 선거지역과 관계없이 당선가능성을 보이는 것으로 나타났고, 득표율은 수도권에서만 유의미한 것으로 나타났다.

〈표 12〉 선거지역 유형과 SNS 유형별 득표율 및 당선여부

(단위: %, 명)

구분		득표율	당선여부		
			전체	당선	낙선
전체		33.67	612 (100.0)	223 (100.0)	389 (100.0)
수도권	1유형	37.95 ^a	89 (14.5)	49 (22.0)	40 (10.3)
	2유형	29.11 ^{cd}	88 (14.4)	25 (11.2)	63 (16.2)
	3유형	31.10 ^{bcd}	87 (14.2)	26 (11.7)	61 (15.7)
	4유형	27.77 ^d	65 (10.6)	14 (6.3)	51 (13.1)
비 수도권	1유형	40.39 ^a	51 (8.3)	26 (11.7)	25 (6.4)
	2유형	34.92 ^{abc}	83 (13.6)	37 (16.6)	46 (11.8)
	3유형	35.60 ^{ab}	56 (9.2)	15 (6.7)	41 (10.5)
	4유형	34.47 ^{abc}	93 (15.2)	31 (13.9)	62 (15.9)
F값/ χ^2 (prob)		4.52(0.0001)	$\chi^2=33.2841$, $df=7$, $prob=0.001$		

2) 지역주의 여부

지역주의는 영호남 지역주의와 그 외 지역주의로 구분해 볼 수 있다. 영호남 지역주의인 경우는 SNS 유형별로 득표율이 통계적으로 유의미한 차이가 있는 것으로 나타났다. 즉 영호남 지역주의의 경우 제1유형(41.46%)이 다른 유형에 비해서 통계적으로 높은 득표율을 보이고 있고, 다른 유형 간에는 통계적 차이가 없는 것으로 나타났다. 반면 그 외 지역의 경우 제1유형(37.82%)과 제3유형(32.08%)의 득표율이 유사하게 나타났고, 제1유형은 제2유형(30.75%)과 제4유형(30.09%)과 통계적 차이가 있는 것으로 나타났다($F=32.18$, $prob=0.0001$).

한편 당선여부의 경우, 영호남 지역주의인 경우는 제2유형(25명)만 당선의 빈도가 상대적으로 높고, 그 외 유형은 당선의 빈도가 상대적으로 낮게 나타나고 있다. 반면 그 외 지역에서는 제1유형(56명)은 당선의 빈도가 다른 유형에 비해서 높게 나타났다.

이상을 종합해 볼 때, 지역주의에 관계없이 제1유형은 득표율을 높이게 하고 있다. 그리고 당선여부에서는 영호남 지역주의인 경우는 제2유형이, 그 외 지역에서는 제1유형의 당선가능성이 높은 것으로 나타났다.

〈표 13〉 지역주의 유형과 SNS 유형별 득표율 및 당선여부

(단위: %, 명)

구분		득표율	당선여부		
			전체	당선	낙선
전체		33.67	612 (100.0)	223 (100.0)	389 (100.0)
영호남	1유형	41.46 ^a	39 (6.4)	19 (8.5)	20 (5.1)
	2유형	34.23 ^{bc}	58 (9.5)	25 (11.2)	33 (8.5)
	3유형	35.02 ^{bc}	38 (6.2)	10 (4.5)	28 (7.2)
	4유형	34.04 ^{bc}	65 (10.6)	22 (9.9)	43 (11.1)

그 외	1유형	37.82 ^{ab}	101 (16.5)	56 (25.1)	45 (11.6)
	2유형	30.75 ^c	113 (18.5)	37 (16.6)	76 (19.5)
	3유형	32.08 ^{bc}	105 (17.2)	31 (13.9)	74 (19.0)
	4유형	30.09 ^c	93 (15.2)	23 (10.3)	70 (18.0)
F값/ χ^2 (prob)		3.28(0.0020)	$\chi^2=29.6130$, df=7, prob=0.0001		

3) 재정자립도

재정자립도가 30% 이상인 선거구에서는 제1유형(37.98%)이 그 외의 유형에 비해서 득표율이 통계적으로 유의미한 것으로 나타났다. 그리고 재정자립도가 30% 미만인 선거구에서는 제1유형(39.56%)이 제2유형(32.89%)과 제4유형(31.87%)에 비해서 득표율이 높은 것으로 나타났다. 다만 제3유형(34.45%)과는 통계적으로 유의미한 차이가 없는 것으로 나타났다(F=2.85, prob=0.0063).

한편 당선여부에서는 재정자립도의 유형과 관계없이 제1유형(각각 33명, 42명)이 다른 유형에 비해서 당선빈도가 높은 것으로 나타났다. 즉 제1유형을 제외한 제2유형, 제3유형, 제4유형은 모두 당선가능성이 낮은 것으로 나타났다($\chi^2=28.4133$, df=7, prob=0.0002).

이상을 종합해 볼 때, 재정자립도가 높은 선거구에서는 제1유형의 SNS 활동이 다른 유형 보다 득표율을 높이는 반면, 재정자립도가 낮은 선거구에서는 SNS 활동유형별 득표율 차이가 크지 않을 것으로 예상된다. 다만 재정자립도와 관계없이 제1유형이 다른 유형에 비해서 당선가능성을 높일 것으로 예상된다.

〈표 14〉 재정자립도 유형과 SNS 유형별 득표율 및 당선여부

(단위: %, 명)

구분		득표율	당선여부		
			전체	당선	낙선
전체		33.67	612 (100.0)	223 (100.0)	389 (100.0)
30% 이상	1유형	37.98 ^{ab}	64 (10.5)	33 (14.8)	31 (8.0)
	2유형	30.54 ^c	70 (11.4)	21 (9.4)	49 (12.6)
	3유형	30.67 ^c	60 (9.8)	19 (8.5)	41 (10.5)
	4유형	31.47 ^c	62 (10.1)	18 (8.1)	44 (11.3)
30% 미만	1유형	39.56 ^a	76 (12.4)	42 (18.8)	34 (8.7)
	2유형	32.89 ^{bc}	101 (16.5)	41 (18.4)	60 (15.4)
	3유형	34.45 ^{abc}	83 (13.6)	22 (9.9)	61 (15.7)
	4유형	31.87 ^c	96 (15.7)	27 (12.1)	69 (17.7)
F값/ χ^2 (prob)		2.85(0.0063)	$\chi^2=28.4133$, df=7, prob=0.0002		

IV. 결론 및 함의

이상에서 살펴본 분석내용을 간략히 정리하면 다음과 같다. 첫째, SNS 활동유형에 따라 득표율과 당선여부의 통계적 차이가 있는 것으로 나타났다. 즉 제1유형은 다른 유형에 비해서 득표율과 당선여부에 있어서 더 높게 나타났다. 다만 제2유형과 제3유형, 제4유형간에는 통계적 차이가 없다. 따라서 득표율이 높고, 당선이 되기 위해서는 선거트윗수가 많고 트위터 활동기간을 많이 해야 한다.

둘째, 후보자의 특성이 반영된 경우, SNS 활동유형은 득표율과 당선여부에 유

의미한 영향력을 미치지 못하는 것으로 판단된다. 즉 현직인 경우, 여당인 경우, 소득이 10억원 이상인 경우, 보수인 경우, 대정당인 경우는 그렇지 않은 경우에 비해서 득표율과 당선여부에서 절대적으로 유리하기 때문이다. 그럼에도 불구하고, 비현직인 경우, 야당인 경우, 10억원 미만인 경우, 진보인 경우, 군소정당이 경우는 제1유형이 다른 유형에 비해서 득표율을 높일 수 있는 유효한 전략인 것으로 나타났다. 그리고 야당인 경우는 제2유형 보다 제3유형이 더 적절한 전략으로 나타났다. 한편 당선여부의 경우 야당인 경우, 10억원 미만인 경우, 진보인 경우는 제1유형이 효과적인 전략인 것으로 나타났으나, 비현직과 군소정당은 어떠한 SNS 활동유형을 선택해도 무조건 낙선하는 것으로 나타났다.

셋째, 선거구 특성이 반영된 경우는 SNS 활동유형은 득표율과 당선여부에 유의미한 영향을 미치는 것으로 판단된다. 우선 수도권에서는 제1유형이 득표율에 유의미한 영향력을 미치는 반면, 비수도권에서는 SNS 활동유형별 차이가 없는 것으로 나타났다. 그리고 당선여부에서는 수도권과 비수도권의 차이 없이 다만 제1유형만이 효과적인 것으로 나타났다. 이는 재정자립도에서도 동일한 경향을 보이고 있다. 그리고 지역주의에 관계없이 제1유형이 득표율이 높은 것으로 나타났고, 당선여부에서는 영호남 지역주의인 경우는 제2유형이, 그리고 그 외 지역에서는 제1유형이 유리한 것으로 나타났다.

〈표 15〉 분석결과의 종합

구분		득표율	당선여부
SNS 유형		제1유형이 유리	제1유형이 유리
후보자 특성	현직	현직여부가 중요 비현직은 제1유형이 유리	현직여부가 중요 비현직은 무조건 불리
	여야	여당여부가 중요 야당은 제1유형이 유리, 제2유형 보다 제3유형이 유리	여당여부가 중요 야당은 제1유형이 유리
	소득	10억원 이상이 중요	10억원 이상이 중요

		10억원 미만은 제1유형이 유리	10억원 미만은 제1유형이 유리
	정치이념	보수가 중요 진보는 제1유형이 유리	보수가 중요 보수는 제2유형이 중요 진보는 제1유형이 유리
	정당규모	대정당이 중요 군소정당은 제1유형이 유리	대정당은 제1유형과 제2유형이 유리 군소정당은 무조건 불리
선거구 특성	선거지역	수도권에서만 제1유형이 유리 비수도권에서는 차이 없음	선거지역 관계없이 제1유형이 유리
	재정자립도	재정자립도가 높은 경우 제1형이 유리 재정자립도가 낮은 경우 차이 없음	재정자립도와 관계없이 제1유형이 유리
	지역주의	지역주의는 중요하지 않음 제1유형이 유리	영호남 지역주의는 제2유형이 유리 그 외 지역은 제1유형이 유리

상기의 분석 결과를 종합하면, SNS 상 선거트윗수가 많고 트위터 활동기간이 많을수록 득표율과 당선가능성이 높다고 할 수 있다. 그러나 득표율과 당선여부에는 후보자 특성이 결정적인 영향을 미치지 때문에, SNS 활동이 득표율과 당선여부를 결정짓는 요소로 작동하지 못하고 있다. 다만 비현직, 야당, 10억 미만의 소득, 진보, 군소정당인 경우는 SNS 활동이 득표율과 당선여부에 유의미한 영향을 미치는 것으로 나타나고 있다. 이런 맥락에서 볼 때, 한국 정치구조의 다원화라는 관점에서 SNS 상의 선거운동을 확대할 수 있는 제도개선이 요구된다고 하겠다. 그리고 선거구 특성이 반영된 경우에도 제1유형이 유리하기 때문에 SNS 활동은 장기간 활동해야 하는 것도 중요하고, 선거트윗수가 많은 것도 함께 중요함을 알 수 있다. 그리고 SNS 활동에서 후보자 측면과 고려되고 있는데, 향후에는 유권자의 측면이 반영될 수 있도록 선거팔로어와 리트윗을 활성화할 수 있는 캠페인 전략도 함께 마련되어야 할 것이다.

참고문헌

- 도묘연. 2015. “SNS 이용자 특성과 선거참여 행태 분석: 18대 대선의 사례.” 『국제정치연구』 제18권 1호: 175-196.
- 박세미·황하성, 2016. “정치인들의 선거캠페인 수단으로서의 SNS 활용.” 『한국인터넷정보학회논문지』 제17권 2호: 87-95.
- 이원태. 2010. 『KISDI Premium Report: 트위터의 정치사회적 영향과 시사점』. 과천: 정보통신정책연구원.
- 이창호·정낙원. 2014. “소셜미디어 이용정도 및 이용동기가 사회자본에 미치는 영향: 개방형·폐쇄형 소셜미디어 비교.” 『한국언론정보학보』 제65권: 5-26.
- 임혁백·송경재·장우영. 2016. 『빅데이터 기반 헤테라키 민주주의: 현황과 전망』. 대구: 한국정보화진흥원.
- 장우영·류석진. 2013. “소셜네트워크 캠페인의 정치적 효과.” 『한국정치학회보』 제47권 4호: 93-113.
- 장우영. 2014 “소셜네트워크 캠페인과 유권자 참여 효과: 18대 대선의 경험적 분석.” 『세계지역연구논총』 제32권 3호: 37-55.
- 장우영. 2017 “20대 총선의 소셜미디어 캠페인 특징과 효과: 수도권 사례.” 『입법과정』 제6권 2호: 57-78.
- 조희정. 2013. 『민주주의의 기술』. 파주: 한국학술정보.
- 하승태. 2012. “소셜네트워크 서비스 이용이 수용자의 정치사회적 태도에 미치는 영향.” 『언론과학연구』. 제12권 4호: 575-606.
- Bruns, A. 2005. *Gatewatching: Collaborative Online News Production*. New York: Peter Lang.
- Schapiro, A. L. 1999. *The Control Revolution: How the Internet Is Putting Individuals in Charge and Changing the World We Know*. New York: Public Affairs.

Castells, M. 2009. *Communication Power*. Oxford: Oxford University Press.

Shirky, C. 2010. *Cognitive Surplus*. New York: Penguin Press.

The SNS Use and Electoral Campaign Effects

Woo-young Chang* · Kwak Ryul Lee**

ABSTRACT

Today's social media communication clearly reveals the characteristics of 'individual centrism' and 'network extensibility' that are based on private friendship. In other words, networking based communication that extends individual friendship away from group is being performed in real time in everyday life. This study analyzes the effect of the SNS election campaign in the case of the Korean 20th general election. From a methodological point of view, this study empirically analyzes campaign effects using social media(twitter) big data of campaign period. Specifically, this study classifies the twitter activity types of candidates and then compares the social media effects among the groups by combining candidate characteristics and electoral characteristics that affect election results. This research is an pilot study that it is difficult to generalize the results, but it shows that the twitter campaign can affect the votes and winnings in combination with the candidate and electoral variables.

Keywords: Big Data, Communication, Electoral Campaign, SNS, Political Effects

투고일: 2017.04.14. 심사일: 2017.04.14. 게재확정일: 2017.04.22.

* Daegu Catholic University

** Chungnam Institute

[DOI] <http://dx.doi.org/10.21487/jrm.2017.05.2.1.79>

【연구경향】

An Introduction to Mixed Logit Model

Byung-Jae Lee*

ABSTRACT

The logit models have been quite useful in analyzing the discrete dependent variable. Specifically, multinomial logit(MNL) and conditional logit(CL) models have been the workhorses in analyzing discrete choice data with polychotomous nominal categories. However, there are some limitation in these models which restrict the applicability of these models. Three well-known limitations in these models are the assumptions of homogeneous preference among individuals, independence of irrelevant alternatives(IIA) and no correlation across time and individuals. One of the recent advances in categorical data analysis is mixed logit(MXL) model, or random parameter model. MXL obviates the well-known limitations of standard logit models by allowing heterogeneous preferences, unrestricted substitution patterns, and correlation in unobserved factors over time. The purpose of this paper is to introduce MXL and its applications.

Keywords: Mixed Logit, Multinomial Logit, Conditional Logit

* Yonsei University. This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government(NRF-2016S1A3A2925033).

I. Introduction

The logit models have been quite useful when the dependent variable is discrete, i.e., when the dependent variable has more than one mutually exclusive category(Agresti 2012; Hilbe 2009; Powers and Xie 2008). More specifically, multinomial logit(MNL) and conditional logit(CL) models have been the workhorses in analyzing discrete choice data with polychotomous nominal categories. Since the methodological innovation by McFadden(1973), plenty of works that applied these models have appeared for analyzing individual choice in many fields, including electoral choice in multi-candidate election setting in political science. The primary advantage of these models is that the researcher could incorporate the alternative-specific variables as well as the variables on the people making choices. However, there are some limitation in these models which restrict the applicability of them. Three well-known limitations in these models are the assumptions of homogeneous preference among individuals, independence of irrelevant alternatives(IIA) and no correlation across time and individuals(Glasgow 2001; Train 2009)

One of the recent advances in categorical data analysis is mixed logit(MXL) model, or random parameter model. MXL obviates the well-known limitations of standard logit models by allowing heterogeneous preferences, unrestricted substitution patterns, and correlation in unobserved factors over time. Like probit, the MXL has been known for many years but has become only very recently available for practical data analysis since the development of simulation techniques.

The main purpose of this paper is to introduce MXL and illustrate its applications with some simple examples. To motivate, random utility, multinomial logit and conditional logit models are introduced in the

second and the third section. The fourth section describes the reasoning and estimation procedures for MXL, and the fifth section discusses the applications of MXL. This paper concludes with a brief discussion of the possible research topics that could use MXL models.

II. Random Utility and Multinomial Logit Model

Most of the logit models are based on random utility maximization behavior. These models are based on the assumption that individual choice among alternatives can be described by a utility function, which depends on the attributes of the alternatives and the characteristics of the individual (Long and Freese 2014). For example, individual voter selects the candidate or party that yields that highest utility. The decision maker i faces a choice among J alternatives. The utility of person i from alternative j is expressed as follows:

$$U_{ij} = V_{ij} + \epsilon_{ij}$$

where V_{ij} is the systematic portion of utility, ϵ_{ij} is the unobserved stochastic portion of utility. In turn, V_{ij} is a function of observable attributes of the alternatives, x_{ij} and of the decision maker, z_i . The probability that individual i chooses alternative j is the probability that the utility of alternative j exceeds the utility of all other alternatives.

$$P_{ij} = \Pr(U_{ij} > U_{ik}) \forall K = \Pr(V_{ij} + \epsilon_{ij} > V_{ik} + \epsilon_{ik} \forall K)$$

Based on the different assumptions about the distribution of random terms, different choice models follow.

Multinomial logit model(MNL) assumes that the unobserved portion of utility are identically and independently distributed(IID) that follows type I extreme value distribution. It is quite well known that the choice probabilities of MNL have the IID property.

An example of MNL is the choice of transportation mode: car, plane, and train. Here, the satisfaction from each transportation mode, V_j , depends linearly on cost and time.

$$\begin{cases} V_1 = \alpha_1 + \beta x_1 + \gamma z_1 \\ V_2 = \alpha_2 + \beta x_2 + \gamma z_2 \\ V_3 = \alpha_3 + \beta x_3 + \gamma z_3 \end{cases}$$

In this case, the probability of choosing alternative J is increasing with V_j , which in turn depends on cost(x) and time(z). For estimation, we need to transform the satisfaction index, which can take any real value so that it is restricted to the unit interval and can be interpreted as a probability. The multinomial logit model is obtained by applying such a transformation to the V_{j_s} . For each probability, we get:

$$\begin{cases} P_1 = \frac{e^{V_1}}{e^{V_1} + e^{V_2} + e^{V_3}} \\ P_2 = \frac{e^{V_2}}{e^{V_1} + e^{V_2} + e^{V_3}} \\ P_3 = \frac{e^{V_3}}{e^{V_1} + e^{V_2} + e^{V_3}} \end{cases}$$

Here, the two characteristics of probabilities are satisfied, which are:

- $0 \leq P_j \leq 1 \forall i = 1, 2, 3$
- $\sum_{j=1}^3 P_j = 1$

Once fitted, a logit model is useful for predicting the probabilities of choice for an individual(Croissant n.d.).

III. Conditional Logit Model and its Limitations

If we make the assumption that the random terms are IID type I extreme value distribution, we obtain the conditional logit model.

$$P_{ij} = \frac{\exp(x_{ij}^T \beta + z_i^T \gamma_j)}{\sum_{j=1}^m \exp(x_{ij}^T \beta + z_i^T \gamma_j)}, j = 1, \dots, m.$$

where \mathbf{x}_{ij} are alternative-specific regressors and \mathbf{z}_i are case-specific regressors. For identification purpose, one of the γ_j , as in typical multinomial model, need to be set at zero in estimation.

Unlike MNL, Conditional logit(CL) model requires the inclusion of alternative-specific variables, while individual- or case-specific variables are sufficient for multinomial logit model. In case that we have a dataset that includes alternative-specific variables, such as prices and quality measures for all alternatives or evaluation of all candidates by respondents, not just the chosen alternatives, then we can use conditional logit model.¹⁾

1) The parameters for conditional logit models are usually estimated with the data in long form(instead of wide form), with one observation providing the data for each alternative for each individual.

As an illustration, here is a conditional logit model using the dataset on mode of transportation(Cameron and Trivedi 2010). The models that incorporate alternative-specific variables require the dataset in which each row of the dataset represents one alternative for one person. So, if we have N individuals choosing among J alternatives, then the data need to have $N \times J$ rows. In the following example, the variable `mode` distinguishes the modes of transportation(1=train, 2=bus, 3=car), and the variable `id` indicates different individuals. `choice` variable indicates which mode is selected(and not selected) by the individual. So, for the first two individuals, the data are as follows:

	id	mode	choice	time	hinc
1	1	Train	Not selected	406	35
2	1	Bus	Not selected	452	35
3	1	Car	Selected	180	35
4	2	Train	Not selected	398	30
5	2	Bus	Not selected	452	30
6	2	Car	Selected	255	30

Here, alternative-specific variable `time` indicates the amount of travel time for each mode of transportation for each individual. The first row indicates that it would take 406 minutes for individual 1 to travel by train, and the alternative train is not selected. Both individual 1 and 2 selected a car as his/her mode of transportation. The variable `hinc`, household income, is case-specific variable and the same value are repeated three times for the same individual. Below is the Stata output for a conditional logit model with combination of alternative-specific (`time`) and individual-specific(`hinc`) variables. For an easier interpretation, odds ratios are presented instead of coefficients.

〈Table 1〉 Conditional Logit Model for mode of transportation

Variables		Odds Ratio
Time		0.98*** (0.002)
Train		Base category
Bus	income	1.03* (0.017)
	intercept	0.25** (0.144)
Car	income	1.05*** (0.016)
	intercept	0.05*** (0.052)
Log-likelihood		-84.92
N		456

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$, standard errors in parentheses.

In Table 1, the coefficient for time indicates the effect of time on the log odds that an alternative is selected. The coefficient of time is negative, which indicates that the chances of an alternative being selected decrease as the amount of time required to travel using that alternative increases. In other words, the odds ratio for time variable, 0.98, indicates that increasing the time of travel by 1 minute for a given mode of transportation decreases the odds of using that mode by a factor of 0.98(2%), controlling other variables. And a unit increase in household income increases the odds of traveling by bus versus traveling by train by 3.2%, controlling other variables. A unit increase in income increases the odds of traveling by car versus traveling by train by 4.9%, controlling other variables.

IV. Moving Beyond the Conditional Logit Model²⁾

Although MNL and CL are quite useful in analyzing the choice situations, there are some problems such as IIA property and the assumption of preference homogeneity. MXL considers the coefficients themselves to be random variables and then estimate the multidimensional integrals that define the choice probabilities using Monte Carlo simulation. In other words, the procedure draws the random numbers from the assumed joint probability distributions of the coefficients, and compute the conditional choice probabilities. This process is repeated many times, and the results are averaged. Then, the average is an unbiased estimate of the unconditional choice probabilities. Let me describe these processes in detail.

For the standard logit model, the probabilities are:

$$P_{ij} = \frac{e^{\beta^T x_{ij}}}{\sum_j e^{\beta^T x_{ij}}}$$

Now suppose that the coefficients are individual specific. Then, the probabilities are:

$$P_{ij} = \frac{e^{\beta_i^T x_{ij}}}{\sum_j e^{\beta_i^T x_{ij}}}$$

In this situation, we have two possible estimation strategies. First, estimate the coefficients for each individual in the sample, or second,

2) This section is largely based on chapter 6 of Train 2009 and Croissant n.d.

consider the coefficients as random variables. The first approach is not much interesting, since it would require numerous observations for each individual. And, we are usually interested in the values of the coefficients for a given individual. From the second approach, the random coefficient model is developed.

1. Random Coefficients

The probability that an individual will choose alternative j is:

$$P_{ij}|\beta_i = \frac{e^{\beta_i^T x_{ij}}}{\sum_j e^{\beta_i^T x_{ij}}}$$

This formula describes the probability for individual i to choose alternative j conditional on the vector of individual-specific coefficients β_i . To get the unconditional probability, we just need to compute the average of these conditional probabilities for all values of β_i .

Suppose that, $V_{ij} = \alpha + \beta_i x_{ij}$, i.e., there is only one individual-specific coefficient and that the density of β_i is $f(\beta, \theta)$, where θ is the vector of parameters of the distribution of β . The unconditional probability is calculated then:

$$P_{ij} = E(P_{ij}|\beta_i) = \int_{\beta} (P_{ij}|\beta) f(\beta, \theta) d\beta$$

which is a one-dimensional integral that can be efficiently estimated by quadrature methods.

If $V_{ij} = \beta_i^T x_{ij}$ where β_i is a vector of length K and $f(\beta, \theta)$ is the

joint density of the K individual-specific coefficients, the unconditional probability is:

$$P_{ij} = E(P_{ij}|\beta_i) = \int_{\beta_1} \int_{\beta_2} \dots \int_{\beta_K} (P_{ij}|\beta) f(\beta, \theta) d\beta_1 d\beta_2 \dots d\beta_K$$

This is a K -dimensional integral which cannot easily be estimated by quadrature methods. In this kind of situations, the only practically possible method is to use simulations. In simulation, R draws of the parameters are taken from the distribution of β , then the probability is computed for every draw and the unconditional probability, which is the expected value of the conditional probabilities, is estimated by averaging R probabilities.

2. Unrestricted Substitution Patterns

One of the strong assumptions required for MNL is IIA assumption. Although multinomial probit model does not require IIA assumption, the mixed logit model provides more flexible alternatives. Mixed logit does not depend on IIA property or the restrictive substitution patterns of logit. The ratio of mixed logit probabilities, P_{ij}/P_{ik} , depends on all the data, including attributes of alternatives other than j or k . The denominators of the logit formula are inside the integrals and therefore do not cancel out. The percentage change in the probability for one alternative given a change in the m^{th} attribute of another alternative is:

$$\begin{aligned} E_{ij} x_{ij}^m &= - \frac{1}{P_{ij}} \int \beta^m L_{ij}(\beta) L_{ij}(\beta) f(\beta) d\beta \\ &= - \int \beta^m L_{ik}(\beta) \left[\frac{L_{ij}(\beta)}{P_{ij}} \right] f(\beta) d\beta \end{aligned}$$

Here β^m is the m^{th} element of β . The elasticity varies across individuals. A ten-percent reduction for one alternative does not need to imply a ten-percent reduction in other alternatives. Instead, the substitution pattern depends on the specification of the variables and mixing distribution, which can be empirically determined. Note that the percentage change in probability depends on the correlation between $L_{ij}(\beta)$ and $L_{ik}(\beta)$ over different values of β , which is determined by the researcher's specification of variables and mixing distribution.

3. Correlation in unobserved factors over time

Standard logit does not incorporate any unobserved factors that persist over time for a given decision maker. This might be problematic for longitudinal data analysis. By applying a standard logit model to longitudinal data, we are assuming that the unobserved factors that affect a person's choice are new every time the person makes the choice, which is very unrealistic assumption. To take into account both random preference variation and correlation in unobserved factors over time, the utility for respondent i for alternative j at time t is specified as follows:

$$U_{ijt} = \beta_i X_{ijt} + \epsilon_{ijt}$$

where subscript t indicates time dimension. However, we still assume the IID extreme value distribution of ϵ over time, individual and alternative. The correlation over time over alternatives arises from the common effect of β s, which enter utility in each time period and each alternative.

Assume that β s are $N(\bar{\beta}, \sigma^2)$. Then, the utility equation becomes:

$$U_{ijt} = (\bar{\beta} + \sigma\eta_i)X_{ijt} + \epsilon_{ijt}$$

Here, η is a draw from the standard normal density. The equation becomes

$$U_{ijt} = \bar{\beta} X_{ijt} + (\sigma\eta_i X_{ijt} + \epsilon_{ijt})$$

$$U_{ijt} = \bar{\beta} X_{ijt} + e_{ijt}$$

where, $e_{ijt} = \sigma\eta_i X_{ijt} + \epsilon_{ijt}$. The second term in the right side of the equation, $\sigma\eta_i X_{ijt}$, is not independent over time or alternatives, while ϵ_{ijt} is independent. Then, the covariance between alternative j and k is,

$$Cov(e_{ijt}, e_{ikt}) = \sigma^2 (X_{ijt} X_{ikt})$$

and the covariance between time t and q is

$$Cov(e_{ijt}, e_{ijq}) = \sigma^2 (X_{ijt} X_{ijq})$$

By specifying the X 's properly, we can obtain any pattern of covariance over time and alternatives.

Conditional on β_i , the probability of the sequence of choices by person is simply the product of the logit probability of each individual choice by that individual.

$$L_i(\beta_i) = \prod_t \frac{e^{\beta_i X_{ijt}}}{\sum_j e^{\beta_i X_{ijt}}}$$

since ϵ_{ijt} is independent over time. Then the unconditional probability

of the sequence of choice is simply the integral of this product of logits over the density of β .

$$P_{ij} = \int L_n(\beta) f(\beta | \theta) d\beta$$

4. Simulation

The probabilities for the random parameter logit are integrals with no closed form. Moreover, the degree of integration is the number of random parameters. In practice, these models with random parameters are estimated using simulations, in which the expected value is replaced by an arithmetic mean. More precisely, the computation process is as follows:

1. Begin with an initial hypothesis about the distribution of the random parameters
2. Draw R numbers from this distribution,
3. For each draw β^r , compute the probability $P_{ij}^r = \frac{e^{\beta^r x_{ij}}}{\sum_j e^{\beta^r x_{ij}}}$
4. Compute the average of these probabilities: $\overline{P_{ij}} = \sum_{r=1}^n \frac{P_{ij}^r}{R}$
5. Compute the log-likelihood for these probabilities
6. Iterate the process until the maximum is reached.

To estimate a model using simulations, one needs to draw pseudo-random numbers from a specified distribution. For this purpose, what we actually need is a function that draws pseudo-random numbers from a uniform distribution between 0 and 1. These numbers are then

transformed using the quantile function of the required distribution. For example, suppose we need to draw random numbers from the Gumbell distribution. The cumulative distribution of a Gumbell variable is expressed as $F(x) = e^{-e^{-x}}$ the quantile function is obtained by inverting this function:

$$\Rightarrow F^{-1}(x) = -\ln(-\ln x)$$

And R draws from a Gumbell distribution are obtained by computing $F^{-1}(x)$ for R draws from the uniform distribution between 0 and 1. The problem is that there may not be a good coverage of the relevant interval for sufficiently numerous draws to be made. More deterministic methods like Halton draws are more practical and useful. Halton draws are made from Halton sequence.

Halton sequence is the sequence used to generate points in space for numerical methods such as MCMC simulations. An interesting feature of Halton sequences is that although they are actually deterministic, they appear to be random for many purposes. It produces well-spaced draws from the unit interval, and can be used in a way that provides negative correlation between simulated probabilities for individuals. Based on a particular prime number, the sequence is constructed based on finer and finer prime-based divisions of sub-intervals of unit interval. Using a prime number 3 as an example, the process is as follows.

1. Break the interval into equal parts. Then start of sequence is 0, 1/3, 2/3.
2. Break each sub-interval into 3 equal parts. Enter the break points into the sequence in a particular way and enter all lowest break points, then enter all the highest breakpoints. Now the sequence

becomes $1/3, 2/3, 1/9, 4/9, 7/9, 2/9, 5/9, 8/9$

3. Break each sub-sub-interval into 3 equal parts, starting with all the lowest sub-intervals. Enter break points into sequence in particular way, first all the lower breaks, then the upper breaks. Now sequence becomes $1/3, 2/3, 1/9, 4/9, 7/9, 2/9, 5/9, 8/9, 1/27, 10/27, 19/27, 4/27, 13/27, 22/27, 7/27, 16/27, 25/27, 2/27, 11/27, 20/27, 5/27, 14/27, 23/27, 8/27, 17/27, 26/27$
4. Continue this until a sequence of desired length is obtained.

V. Applications

1. Household's choice of electricity supplier

The first example uses the dataset compiled by Huber and Train(2001) on household's choice of electricity supplier. In this experimental data, residential electricity customers were presented with a series of experiments with four alternative electricity suppliers. The following alternative-specific variables are included in the dataset.

- price: price in cents per kWh if fixed price, 0 if TOD or seasonal rates
- contract: contract lengths in years
- local: whether a local company(0 or 1)
- wknown: whether a well-known company(0 or 1)
- tod: time of day rates(0, 1)
- seasonal: seasonal rates(0, 1)

The following is the mixed logit results, with contract, local, wknown, tod and seasonal being assumed to be normally distributed random variables. The coefficient of price is fixed. The following estimation is performed using mixlogit command in Stata, which is a user-written program by Hole(2007).

〈Table 2〉 Mixed Logit Model for Household Electricity

	Variables	Coefficients
Mean	price	-0.95*** (0.07)
	contract	-0.27*** (0.05)
	local	2.14*** (0.23)
	wknown	1.55*** (0.18)
	tod	-9.32*** (0.61)
	seasonal	-9.35*** (0.61)
SD	contract	0.39*** (0.04)
	local	1.87*** (1.87)
	wknown	1.24*** (1.24)
	tod	2.47*** (2.47)
	seasonal	2.26*** (2.26)
Log-likelihood	-1105.2832	
N	4,780	

***p < 0.01, ** p < 0.05, * p < 0.10, standard errors in parentheses.

The upper panel in Table 2 shows the mean coefficients and the lower panel estimated standard deviations of the coefficients. All

estimates are highly significant, which means that these coefficients really vary across individuals. The mean coefficients for contract, tod, and seasonal are negative, and the coefficients for local and wknown are positive. The signs in the coefficients indicate the direction of the average effects of these variables on the dependent variable. Customers prefer shorter contract length, non-tod and non-seasonal rates, while they prefer local and well-known companies. Because we assume normal distribution for these coefficients, we can get some information that cannot be obtained in standard logit. For example, using the mean coefficients, we can determine the amount that a customer with average coefficients for fixed price and contract length is willing to pay for an extra year of contract length, which is $0.266/0.958 = 0.277$. This means that a customer with mean coefficients for price and contract length is willing to pay 0.277 cents per kWh extra to have a contract that is one year shorter.

2. Train Traveling

The second example is to use the Train data, which is included in R package “mlogit.” The data show the stated preferences for train traveling and were collected in Netherlands in 1987. The unit of observations is individual and the number of observations is 2,929. The dependent variable is either choice 1 or choice 2. The other variables in the dataset are as follows:

- Price: fare of the train(1 or 2)
- Time: travel time in minutes
- Comfort: comfort level(1, 2, or 3 in decreasing order)
- Change: number of changes(0, 1, 2, 3, or 4)

In this application, two mixed logit models, correlated(MXLC) and uncorrelated models(MXLU), are estimated. Time, change, and comfort variables are assumed to be randomly distributed. Censored normal and normal distribution are used for time and change, respectively, while log normal distribution for comfort are used for random draws for each parameter. For comparison, multinomial logit results are also included.

<Table 3> Multinomial Logit, Correlated Mixed Logit and Uncorrelated Mixed Logit Models for Train

	MNL	MXLC	MXLU
intercept	-0.030 (0.04)	-0.032 (0.048)	-0.035*** (0.041)
Price	0.001*** (0.00)	0.001*** (0.000)	0.002*** (0.000)
time	0.028*** (0.003)	0.027*** (0.001)	0.033*** (0.001)
change	0.326*** (0.060)	0.349*** (0.043)	0.410*** (0.03)
comfort	0.947*** (0.947)	-0.130*** (0.018)	-0.077*** (0.006)
Time.time		0.112** (0.001)	
Time.change		-0.200*** (0.061)	
Time.comfort		0.152*** (0.002)	
Change.change		1.607*** (0.059)	
Change.comfort		2.060*** (0.007)	
Comfort.comfort		1.451*** (0.006)	
Sd.time			0.111*** (0.000)
Sd.change			2.189*** (0.03)
Sd.comfort			3.428*** (0.004)
Log-Likelihood	-1723.8	-1556.2	-1589.6
Mcfadden's R ²	0.151	0.233	0.217

***p < 0.01, ** p < 0.05, * p < 0.10, standard errors in parentheses.

Table 3 shows the results of multinomial logit (MNL), MXLC (correlated mixed logit), and MXLU (uncorrelated mixed logit). In terms of signs of the coefficients, the notable difference is in the sign of comfort variable. In MNL, it does not provide much information other than the positive sign. The signs of the coefficients for comfort variable in two mixed logit models are negative, which is the opposite to the one in MNL. The standard deviation of the coefficient for comfort shows large variation of the effect of comfort on the choice. The distribution in Table 4 shows that the mean is much larger than the median, which indicates that some respondents give very high weights on the comfort, although most people do not.

〈Table 4〉 Distribution of Mean Coefficients for MXLC and MXLU

	time		change		comfort	
	MXLC	MXLU	MXLC	MXLU	MXLC	MLXU
Min	0	0	-	$-\infty$	0	0
1 st quartile	0.00	0.000	-0.743	-1.070	0.160	0.092
Median	0.027	0.033	0.349	0.406	0.878	0.926
Mean	0.060	0.063	0.349	0.406	21.238	329.307
3 rd quartile	0.103	0.108	1.441	1.882	4.819	9.345
Max	∞	∞	∞	∞	∞	∞

3. Latent Class Mixed Logit Model

So far, we have assumed that the distribution of the coefficients in the model is continuous. However, the coefficients could be discrete and we need more elaborate model in these situations. In latent class model, each respondent is assumed to belong to a class q , in which preferences vary across individuals, but not within classes. In this case, the

probability of a particular sequence of choices is given by:

$$S_i = \sum_{q=1}^Q H_{iq} \prod_{t=1}^T \prod_{j=1}^J \left[\frac{\exp(x_{ijt}^T)\beta_q}{\sum_{j=1}^J \exp(x_{ijt}^T)\beta_q} \right]^{y_{ijt}}$$

And the probability of belonging to class q , i.e., H_{iq} in the above equation, is usually specified as:

$$H_{iq} = \frac{\exp(Z_i^T \gamma_q)}{\sum_{q=1}^Q \exp(Z_i^T \gamma_q)}$$

where γ_q is set to zero. The log-likelihood for this model is

$$S_i = \sum_{t=1}^I \ln \left\{ \sum_q H_{iq} \prod_{t=1}^T \prod_{j=1}^J \left[\frac{\exp(x_{ijt}^T)\beta_q}{\sum_{j=1}^J \exp(x_{ijt}^T)\beta_q} \right]^{y_{ijt}} \right\}$$

This can be maximized directly using standard methods, or indirectly using the EM algorithm. Following is an example of the three classes latent model. For the estimation, the same dataset as in the first example, the choice of household's electricity supplier, is used. For the adequate number of classes, the standard methods of model comparison, such as Akaike Information Criteria(AIC) or Bayesian Information Criteria(BIC), can be used. Here, three latent classes model are specified for estimation. The interpretation of the results is same as in the first example, although the coefficients are estimated only for the people in the same class.

〈Table 5〉 Latent Class Mixed Logit Model for Household Electricity

	Variables	Coefficients
Choice1	Price	-0.803*** (0.118)
	Contract	-0.507*** (0.059)
	Local	0.499*** (0.205)
	Wknown	0.357** (0.182)
	Tod	-5.992*** (0.932)
	seasonal	-6.647*** (0.993)
Choice2	Price	-0.211** (0.085)
	Contract	0.023 (0.030)
	Local	3.087*** (0.250)
	Wknown	2.308*** (0.234)
	Tod	-1.878** (0.735)
	seasonal	-1.965** (0.764)
Choice3	Price	-1.140*** (0.141)
	Contract	-0.230*** (0.065)
	Local	1.675*** (0.336)
	Wknown	1.645*** (0.262)
	Tod	-12.530*** (1.387)
	seasonal	-11.750*** (1.138)
Share1		-0.279 (0.291)
Share2		0.000 (0.286)

***p < 0.01, ** p < 0.05, * p < 0.10, standard errors in parentheses.

Table 5 shows that the weights for each variable vary across the classes, although there are also variations within classes. Choice2 has larger mean coefficients for tod and seasonal, while choice2 has larger coefficients for local and wknown and the negative sign of the coefficient for contract. What this table shows is that there are approximately three groups which weight alternative-specific variables differently.

VI. Discussion

The MXL has many advantages over other choice models. It enables us to incorporate the possibility of preference heterogeneity among individuals, to relax the IIA assumption and to make the incorporation of correlated preferences across time possible. This model has many potential venues of application in political science. For example, it nicely matches the spatial model that assumes random individual weights that were distributed independently of preferences, which can be modeled via MXL. As Mebane et al. illustrate, MXL latent class model can be applied to model “issue public”(Mebane et al. 2014). According to them, issue public that Converse describes means group of voters who vote on the basis of a single or small set of issues and that this issue or set varies among individuals or groups. The proposition of issue public implies different fixed or mean values for the weights for different groups of voters, which can be exactly modeled via MXL latent class models.

The existing statistical methods such as multinomial logit or conditional logit has some limitations to deal with this kind of preference heterogeneity. Although the models of voting behavior existed for decades, little is understood about the heterogeneity in the impact of

issue positions, demographic variables, and other factors on vote choices. Both MNP and MXL can be specified to explore random preference variation. However, MNP is limited in the number of random coefficients that can be estimated, and these random coefficients should be normally distributed. MXL can include any number of random coefficients, and these coefficients can follow any distributions. In this sense, MXL could provide useful tools by which we can probe many unanswered questions in voting behavior in particular, natural and social science in general.

References

- Agresti, Alan. 2012. *Categorical Data Analysis*. 3rd ed. New York: John Wiley & Sons.
- Bhat, Chandra R. 2001. "Quasi-random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model." *Transportation Research Part B: Methodological* 35(7): 677-693.
- Croissant, Yves. n.d. "Estimation of Multinomial Logit in R: The mlogit packages."
- Cameron, A. Colin and Pravin K. Trivedi. 2010. *Microeconometrics Using Stata*. Rev. ed. College Station, TX: Stata Press.
- Glasgow, Garrett. 2001. "Mixed Logit Models for Multiparty Elections." *Political Analysis* 9(2): 116-136.
- Greene, H. William H. and David Hensher. 2003. "A Latent Model for Discrete Choice Analysis: Contrasts with Mixed Logit." *Transportation Research Part B: Methodological* 37(8): 681-698
- Hilbe, Joseph. H. 2009. *Logistic Regression Models*. Boca Raton, FL: CRC Press.
- Hole, Arne Risa. 2007. "Fitting Mixed Logit Models using Maximum Simulated Likelihood." *The Stata Journal* 7(3): 388-401.
- Huber, Joel and Kenneth Train. 2001. "On the Similarity of Classical and Bayesian Estimates of Individual Mean Partworths." *Marketing Letters* 12(3): 259-269.
- Jaeger, T. Florian. "Categorical Data Analysis: Away from ANOVAs(transform or not) and towards logit mixed models." *Journal of Memory and Language* 59(4): 434-446.

- Long, J. Scott and Jeremy Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata*. 3rd ed. College Station, TX: Stata Press.
- McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." in *Frontiers in Econometrics*, ed. Paul Zarembka. New York: Academic Press.
- McFadden, Daniel and Kenneth Train. 2000. "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* 15(5): 447-470.
- Mebane, Walter, Jr., John E. Jackson, and Jonathan Wall. 2014. "Preference Heterogeneities in Models of Electoral Behavior." Paper presented at the 2014 Annual Meeting of the Midwest Political Science Association.
- Powers, Daniel A. and Yu Xie. *Statistical Methods for Categorical Data Analysis*. 2nd ed. San Diego: Academic Press.
- Revelt, David and Kenneth Train. 1998. "Mixed Logit with Repeated Choices: Households choices of Appliances Efficiency Level." *Review of Economics and Statistics* 80(4): 647-657.
- Train, Kenneth. 1988. "Recreation Demands Models with Taste Variation." *Land Economics* 74(2): 230-239.
- Train, Kenneth. 1999. "Halton Sequences for Mixed Logit." Unpublished manuscript.
- Train, Kenneth. 2009. *Discrete Choice Methods with Simulation*. 2nd ed. New York: Cambridge University Press.
- Train, Kenneth. 2016. "Mixed Logit with a Flexible Mixing Distribution." *The Journal of Choice Modelling* 19: 40-53.

혼합로짓(mixed logit)이란 무엇인가?

이병재*

논문요약

사회과학에서 널리 쓰이는 범주형 데이터, 특히 범주형 종속변수의 분석에 있어 로짓모델은 지난 수 십년간 매우 널리 사용되어왔다. 특히, 다항로짓(multinomial logit)과 조건부로짓(conditional logit)은 종속변수가 3항이상의 명목형 범주로 이루어져 있는 경우 매우 유용한 것이 사실이다. 하지만 이 모델들은 널리 알려진 한계가 있다. 첫째, 개인들 간의 선호도의 동질성의 가정, 둘째, 무관한 대안으로 부터의 독립성(independence of irrelevant alternatives)의 가정, 셋째, 시간별, 개인별 무상관성(uncorrelatedness across time and individual)의 원칙의 가정이 그것이다. 시뮬레이션에 기반한 기법이 확산되기 시작한 이래 범주형 데이터 연구에서 새롭게 활발히 개발, 적용중인 방법중 하나인 혼합로짓(mixed logit)은 개인간 불균질한 선호도(heterogeneous preferences), 교체패턴(substitution patterns)의 자유로운 설정, 시간 및 개인간 상관성 등을 범주형 모델링 과정에 포함시킬 수 있는 가능성을 제시함으로써 전통적으로 사용되던 범주형 데이터 분석의 한계의 극복 뿐만아니라 보다 심화된 데이터 분석을 제시 할 수 있는 방법을 제시한다. 본 논문의 목적은 혼합 로짓을 소개하는 것이다.

주제어: 혼합로짓, 다항로짓, 조건부로짓

투고일: 2017.04.04. 심사일: 2017.04.04. 게재확정일: 2017.04.11.

* 연세대학교