

생태학적 추론(ecological inference, EI) 모형을 활용한 사전투표제 도입에 따른 국회의원선거의 투표참여방식 변화 연구*

박인서** | 연세대학교

백영민*** | 연세대학교

| 국문초록 |

사전투표제 도입이 투표참여율에 미치는 효과에 대해서는 연구자마다 ‘분산효과’와 ‘유인효과’로 의견이 엇갈린다. 본 연구에서는 새로운 이론적·방법론적 접근을 통해 사전투표제의 도입효과를 점검하고자 하였다. 설문표본 데이터를 기반으로 유권자의 투표방식선택을 분석한 대부분의 국내외 선행연구들과 달리, 본 연구에서는 선거구역 단위의 모집단 투표참여 데이터를 활용해 개별유권자의 투표참여방식 변화를 추정하는 ‘생태학적 추론(ecological inference, EI)’ 모형(King et al. 2004)을 이용하였다. 구체적으로 ‘투표참여방식 전이확률(probability of vote choice transition)’이라는 새로운 이론적 개념을 소개하고, 사전투표제 도입 이후 두 차례 실시된 국회의원 선거에 나타난 투표참여방식 변화를 살펴보았다. 아울러 EI 모형으로 추정한 투표참여방식 전이확률 데이터와 동(洞)별 특성 데이터를 융합하여, 투표방식변화가 동단위 선거구역 특성과 어떤 관계를 갖는지 알아보았다. 연구 결과, 분산효과

주제어 | 사전투표제, 투표참여방식, 생태학적 추론(EI), 분산효과, 유인효과

* 이 논문은 2016년도 정부재원(교육부 인문사회연구역량강화사업비)으로 한국연구재단의 지원(NRF-2016S1A3A2925033)을 받아 연구되었습니다.

** 통계데이터사이언스학과 석사과정

*** 사회과학대학 언론홍보영상학부 부교수

및 유인효과 측면에서 사전투표제 도입의 단기적·장기적 효과를 확인하였으며, 투표참여방식 변화에 유권자의 연령분포, 선거구역 규모 외에도 가구구성이 중요한 역할을 한다는 것을 밝혀 낼 수 있었다. 본 연구에서 제안한 투표참여방식 전이확률 개념과 TI 모형은 향후 선거연구에서 유권자들의 투표행위를 설명하는 데 폭넓게 활용될 수 있을 것으로 기대된다.

I. 서론

현대 대의민주주의 체제는 주기적 선거를 통해 정치적 정당성을 확보한다. 이를 위해 대의민주주의 국가들에서는 유권자들의 투표참여를 보장하고 독려하기 위해 다양한 정책들을 실시해왔다(이상신 2014; Gronke et al. 2008). 투표참여를 비용(cost)과 효용(benefit)의 관점으로 파악하는 고전적 투표참여이론에 따르면, 투표참여 소요비용은 ‘투표참여방법의 복잡성’, ‘투표장에 대한 접근성’, ‘투표참여 시간의 제한성’에 따라 달라진다. 적어도 이론적으로는 투표참여방법이 쉽고, 투표장에 쉽게 접근할 수 있으며, 더 많은 투표참여시간이 부여될수록 투표참여 소요비용이 낮아져 투표참여율이 증가한다(Aldrich 1993). 2013년에 도입된 사전투표제는 유권자에게 법정 선거일 이전에도 투표할 수 있는 추가적 기회를 부여하여 더 많은 투표참여시간을 부여한 편의투표(convenience voting) 제도다(Gronke et al. 2008). 특히 한국의 사전투표제는 ‘주민등록제’와 ‘효율적 전산시스템’을 기반으로 유권자들이 보다 쉽게 투표장에 접근할 수 있도록 하기 때문에 투표참여율을 끌어올릴 것이라는 기대가 대두되었다(김찬송 외 2016; 이상신 2014).

그러나 사전투표제로 인해 기대만큼 투표참여가 활성화되었는지에 대해서는 학자들마다 의견이 엇갈린다. 사전투표제 도입이 투표참여율에 끼친 영향에 대한 국내의 연구들은 크게 ‘분산효과’와 ‘유인효과’로 구분할 수 있다. ‘분산효과’를 주장하는 학자들(가상준 2017, 2018; 강신구 2016; Barreto et al. 2006; Berinsky 2005; Garnett 2019; Gronke and Toffey 2007, 2008)은 사전투표제와 같은 편의투표제가 도입되면 기존의 선거참여자들이 사전선거와 당일선거로 분산될 뿐이며, 과거에 투표하지 않았던 투표기권자들을 새로 끌어들이지 못한다고 주장한다. 이들에 따르면 유권자는 투표를 할 것인지 여부를 먼저 결정한 후에 투표방식을

결정하기 때문에, 사전투표제는 투표에 참여할 생각이 없는 유권자를 투표장에 끌어들이 수 없다고 주장한다. 반면 ‘유인효과’를 주장하는 학자들(김찬송 외 2016; 윤기쁨 · 엄기홍 2016; 이상신 2014)은 유권자의 개인적 · 사회적 특성에 따라 투표참여비용이 각각 다르다고 가정한다. 따라서 사전투표제 도입으로 투표참여 소요비용이 낮아지면, 기존의 선거참여자들은 물론 투표참여비용을 높게 여겼던 과거 투표불참자들 중 일부는 추가로 투표에 참여할 수 있기 때문에 사회 전체적으로는 투표참여율이 제고되는 효과가 나타난다고 주장한다.

사전투표제 도입이 투표참여율에 미치는 효과에 대한 논란에도 불구하고, 국내 선행연구들은 설문조사를 통해 얻은 제한된 유권자 표본을 토대로 사전투표제 도입이후 선거에서 나타난 유권자의 투표방식선택을 설명하고 예측한다는 공통점을 갖는다. 설문조사는 제한된 연구자원으로 모집단을 대표하는 연구표본을 얻을 수 있는 효율적 · 효과적인 사회과학 연구방법이지만 여러 한계점에서 자유롭지 않다. 첫째, 연구표본을 얻는 과정에서 투표참여자가 과도하게 표집되며, 무엇보다 투표참여여부에 대한 응답의 타당성에는 문제가 있다고 알려져 있다 (Ansolabehere and Hersh 2012; Karp and Brockington 2005). 즉 설문조사 기법으로 사전투표제 도입후 투표불참자의 투표참여방식 선택을 파악하기 어렵다는 방법론적 제약이 있다. 둘째, ‘사전투표제 도입효과’를 파악하기에 소규모 설문표본은 통계적 검증력(statistical power)이 낮을 수 있다(Cohen 2013). 즉 유인효과를 주장하는 연구자들의 말처럼 사전투표제 도입이 특정 인구집단에게서만 제한적으로 나타난다면, 1,000명 규모인 소규모 설문표본으로 사전투표제 도입에 따라 제한된 인구집단에게서 나타나는 유인효과를 파악하기 쉽지 않을 수 있다. 셋째, 어떤 사회제도든 정책효과를 발휘하기 위해서는 제도가 알려지고 정착되는데 상당한 시간이 소요되지만, 설문조사 방법을 통해서서는 장기간에 걸친 제도적 효과를 추적하기 어렵다. 사전투표제 도입효과를 살펴본 거의 대부분의 국내 선행연구들은 특정 시점에 실시된 단일선거를 배경으로 수집한 설문표본 데이터를 분석하는 것이 보통이다. 끝으로 넷째, 대부분의 설문조사 기법들은 ‘1인1표’의 대의민주주의 투표원칙을 반영하기 위해 개별유권자를 표집단위로 설정한다. 그러나 초창기 투표연구(Larzarsfeld et al. 1944)에서도 잘 드러나듯 투표행위는 사회적 행위이지만, 대부분의 국내외 선행연구들의 경우 설문표본을 수집할 때, 가족과 같은

사회집단 혹은 지역단위의 특성 등을 고려하지 않는 것이 보통이다.

본 연구에서는 기존 연구들과는 다른 각도에서 사전투표제 도입효과를 점검하였다. 선행연구들과 구분되는 본 연구의 특징들은 다음과 같다. 첫째, 본 연구에서는 선거구역 단위의 모집단 집합데이터(aggregate data)를 기반으로 개별유권자의 투표참여방식 변화를 추정하는 ‘생태학적 추론(ecological inference, EI)’ 모형(Adolph et al. 2003; King 1997; King et al. 2004)을 이용하였다. EI에 대한 보다 구체적인 설명은 ‘이론적 논의’에서 소개하였다. 둘째, 본 연구에서는 단일선거에서 나타난 투표참여방식이 아닌, 이전 선거와 현재 선거의 투표참여방식의 변화를 종속변수로 살펴보았다. 개념적으로 분산효과는 기존 투표참여자의 투표참여방식 다변화를, 유인효과는 기존 투표불참자의 투표참여로의 변화를 의미한다. 다시 말해 두 효과 모두 투표참여방식의 변화(분산효과의 경우 ‘당일투표→사전투표’, 유인효과의 경우 ‘기권→사전투표’)를 다루지만, 대부분의 국내 선행연구는 특정된 단일 시점의 투표참여방식을 설명하는 데 국한되어 있다. 셋째, 사전투표제가 도입된 후 실시된 총 2차례의 국회의원 선거를 살펴봄으로써 사전투표제 도입에 따른 장기효과에 주목하였다. 넷째, 서울시 동(洞)단위 공개데이터와 동별 투표참여방식 변화 데이터를 연결시키는 방식으로 선거구역별 특징을 통해 사전투표제 도입의 사회적 의미를 탐색하였다.

II. 이론적 논의

1. 사전투표제의 도입과 투표참여방식 변화

2013년 사전투표제가 도입된 이후 거의 대부분의 선거에서 투표참여율이 상승했다. 지방선거의 경우, 2010년의 경우 투표참여율이 54.5%였으나, 사전투표제 도입이후인 2014년에는 56.8%로, 2018년에는 60.2%로 투표참여율이 증가하였다. 대통령 선거의 경우, 2012년의 투표참여율이 75.8%인 반면, 사전투표제가 도입된 후 투표참여율이 77.2%로 증가하였다. 본 연구에서 살펴볼 국회의원 선거의 경우도, 2012년 국회의원 선거 투표참여율이 54.2%에 비해 사전투표제 도입이후에는 2016년 58%, 2020년 66.2%로 투표참여율이 상승하였다.

이러한 투표참여율 증가가 사전투표제의 도입으로 인해 나타났는지를 둘러싼 연구자들의 견해는 ‘분산효과’와 ‘유인효과’로 엇갈린다. 먼저 분산효과를 주장하는 국내연구자들은 사전투표제 도입시기인 2013년 이후의 선거참여율 증가를 사전투표제 도입효과로 볼 수 없다고 주장한다. 일반적으로 사회적 갈등이 커지면 선거에 대한 관심이 증가하며 선거참여율이 높아지는 경향이 있다(Abramowitz 2010). 예를 들어 2014년 지방선거의 경우 ‘세월호 사건’으로 인해 박근혜 정권에 대한 비판의 수위가 높아지던 시점이었고, 2016년 국회의원 선거의 경우 당시 여당인 새누리당 공천파동, 야당인 새정치민주연합의 계파갈등으로 인해 선거에 대한 관심이 증대된 시점이었으며, 2017년 대통령 선거 역시 박근혜 전대통령 탄핵과 촛불시위로 대표되는 정치사회적 변화로 인해 투표율이 증가한 것이라는 주장이다. 즉 분산효과를 지지하는 연구에서는 2013년도 이후의 투표참여율 증가는 ‘선거상황의 변화’가 주 원인이며, 사전투표제 도입의 효과는 기존투표자들의 투표방식 다원화에 머문다고 주장한다(가상준 2017, 2018; 강신구 2016).

반면 유인효과를 주장하는 연구자들은 2013년 이후의 선거참여율 증가의 상당부분은 사전투표제 도입에 따른 효과일 수 있다고 주장한다. 이들은 사전투표제 도입으로 인한 분산효과를 부정하지는 않는다. 사전투표에 참여한 유권자들 중 대다수는 사전투표제가 없었다라도 투표에 참여했을 유권자이지만, 일부 사전투표 유권자들은 업무상의 출장이나 생업, 가족단위의 여행·행사 등의 이유로 공식 선거일에 투표하기 어려웠던 사람들을 추가로 투표에 참여한 유권자라고 주장한다(김찬송 외 2016; 윤기쁨·엄기홍 2016; 이상신 2014). 즉 유인효과를 주장하는 연구자들은 사전투표제 도입으로 인해 기존 선거에서는 투표에 참여하지 않았거나 못했던 유권자들이 새로이 사전투표에 참여할 수 있게 되었다고 주장한다.

우리나라의 사전투표제와 성격이 동일하지는 않지만, 이러한 논란은 부재자 투표(absentee voting), 우편투표(vote-by-mail) 등의 편의투표제 도입효과를 살펴본 해외연구들에서도 마찬가지다. 편의투표제 도입 후 투표참여율이 감소했다는 매우 드문 사례(예를 들어, Burden et al. 2014)를 제외한다면, 대의민주주의 투표제도가 정착된 북미와 서유럽 국가들에서는 편의투표제 도입 이후 투표참여율은 작지만 유의미하게 증가한다는 결과(약 2~4%)가 누적적으로 확인되었다(Garnett 2019; Gronke et al. 2008; Gronke and Toffey 2007, 2008; Southwell and

Burchett 2000; Stein and Vonnahme 1998). 그러나 한국의 사전투표제 도입 효과를 둘러싼 논란과 마찬가지로 과연 편의투표제 도입이후의 투표참여율 증가가 과연 선거제도 개편에 따른 효과인지, 아니면 의미가 없을 정도로 미미한 효과인지에 대한 연구자들의 평가는 다르다. 편의투표제 도입이 투표참여율에 끼친 효과에 대한 국내외 선거연구자의 전반적인 평가는 ‘분산효과’에 가깝다. 즉 사전투표와 같은 편의투표방식을 택한 유권자는 기권자보다 투표소에 직접 방문하여 투표한 유권자와 유사하며, 편의투표제 도입이후의 미미하게 증가한 투표참여율은 열성적 선거참여자(이들테면, 연령이 높고 정당일체감이 강하며 정치관심도가 높은 유권자)들에게 추가적인 선거기회가 보장되면서 나타났을 뿐, 전통적으로 선거에 참여하지 않았던 유권자들의 참여를 유도하여 나타난 것이 아니라는 것이다(가상준 2017, 2018; 강신구 2016; Barreto et al. 2006; Berinsky 2005; Garnett 2019; Gronke and Toffey 2007, 2008).

사전투표제와 같은 편의투표제 도입이 투표참여율에 미친 효과가 지극히 제한적이라는 국내외 선행연구들에도 불구하고, 본 연구에서는 선행연구들의 방법론적 한계점들에 주목하였다. 국내외 선행연구들은 다음과 같은 방식으로 사전투표제와 같은 편의투표제 도입효과 분석결과를 도출한다. 먼저 설문표본 데이터를 이용해 유권자들을 표집한 후, 응답자의 투표참여방식을 ‘기권’, ‘당일투표’, ‘사전투표(국외의 경우 편의투표)’의 세 범주로 구분한다. 이후 투표참여에 영향을 미치는 다양한 인구통계학적, 사회적, 심리적 변수들을 독립변수로, 3개 범주로 구성된 투표참여방식을 종속변수로 하는 다항로지스틱 회귀모형 추정결과를 기반으로 ‘사전투표’ 선택집단이 ‘기권’ 선택집단보다 ‘당일투표’ 선택집단에 가깝다는 분석결과를 제시한다. 그러나 설문표본 데이터를 기반으로 한 다항로지스틱 회귀모형 분석결과는 다음과 같은 방법론적 한계점들에서 자유롭지 못하다.

첫째, 설문표본이 유권자 모집단을 제대로 대표하지 못할 가능성이 작지 않으며, 설문응답의 타당성도 완전히 확보되기 어렵다. 설문표본을 확보하는 과정에서 투표참여자가 투표불참자에 비해 더 과도하게 표집되는 경우가 거의 대부분이며(가상준 2018; 강신구 2016; 김찬송 외 2016; Ansolabehere and Hersh 2012), 무엇보다 자기응답식 설문조사를 통해 얻은 투표참여여부는 타당성이 낮은 것이 보통이다(Karp and Brockington 2005). 투표참여여부 측정의 문제는 ‘유인효과’

테스트에 문제를 일으킬 수 있다. 왜냐하면 사전투표제 도입으로 늘어난 투표참여자는 첫째, 기존 투표 불참자이며, 둘째, 사회적 주류에서 벗어난 유권자이기 때문이다(김찬송 외 2016). 여러 방법론 문헌들에서 누적적으로 보고되듯, 이런 응답자들은 설문조사 과정에서 잘 표집되지 않는 것이 보통이다(Ansola-behere and Hersh 2012; Karp and Brockington 2005). 사전투표제 도입으로 인한 투표참여율 증진효과가 미미하다는 연구결과는 설문표본 데이터 수집과정에서의 문제점을 반영했을 수 있다.

둘째, ‘사전투표제 도입’이라는 정책효과를 파악하기에 제한된 규모의 설문표본 데이터는 충분한 통계적 검증력(Cohen 2013)을 확보하지 못할 가능성이 높다. 유인효과 연구자들이 인정하듯, 사전투표제 도입효과는 전체 유권자들 중 ‘일부집단’에서만 나타날 가능성이 높다. 선행연구(김찬송 외 2016; 이상신 2014)에 따르면, 사전투표를 한 유권자들 중 유인효과가 나타나는 집단은 업무상 잦은 출장과 근무시간의 한계 등으로 당일투표가 어려운 사람들이었다. 그러나 대부분의 국내 선행연구에서 활용하는 설문표본 데이터 규모는 1,000명 안팎으로, 유인효과를 기대할 수 있는 응답자 집단을 충분히 확보하기 어렵다. 실제로 김찬송 외(2016, 6쪽)의 연구결과에 따르면, 유인효과를 보인 응답자 비율은 5.7%에 그친 반면, 분산효과를 보인 응답자는 18.6%로 나타나 그 비율이 약 1:3으로 나타났다. 이런 상황에서 유권자를 ‘기권’, ‘당일투표’, ‘사전투표’의 3개 범주로 구분한 후 추정한 다항로지스틱 회귀분석 결과는 당연히 ‘분산효과’를 지지하기 쉽다. 사전투표를 선택한 유권자 집단 내부의 이질성(heterogeneity)을 구분하고, 사전투표제 도입에 민감하게 반응하는 소규모 인구집단을 파악하기 위해서는 대규모의 설문표본이 필수적이다.

셋째, 어떤 사회제도든 제도개혁 효과는 단기효과에 머무르지 않는다. 본질적으로 투표는 시민적 습관(habit)이 반영된 정치적 행동이며(Berinsky 2005; Campbell et al. 1980; Larzarsfeld et al. 1944; Stein and Vonnahme 2010), 따라서 유권자들이 새로운 투표제도를 이해하고 익숙해질 때까지 상당한 시간이 소요된다. 그러나 단발성 횡단(cross-sectional) 설문조사 기법을 통해서만 장기간에 걸친 제도개혁의 효과를 추적하기 어렵다. 사전투표제 도입효과를 살펴본 거의 모든 국내 선행연구들(가상준 2017, 2018; 강신구 2016; 김찬송 외 2016; 윤기

뽐·엄기홍 2016)은 특정 시점에 실시된 단일선거를 배경으로 수집한 설문표본 데이터를 분석하는 것이 보통이다. 만약 사전투표제 도입효과가 크기는 작지만 장기간의 누적된 효과라면, 특정선거에서 실시한 일회성 설문조사로는 투표참여율 증가효과를 발견하기 쉽지 않다.

넷째, 설문표본 데이터에 기반한 대부분의 연구들의 분석단위는 ‘개별유권자’다. ‘1인1표’의 대의민주주의 투표원칙을 떠올릴 때, 유권자 개인을 표집단위로 설정하는 것은 자연스럽다. 그러나 초창기 투표연구(Larzarsfeld et al, 1944)에서도 잘 드러나듯 투표행위는 사회적 집단행위(collective behavior) 성격을 강하게 띠고 있다. 일반적으로 가구구성원이 많으면 많을수록 투표참여 가능성이 증가하며(Verba and Nie 1972), 가구구성원들은 단체로 투표장을 찾는 것이 보통이다. 특히 ‘유권자의 가구구성’은 우리나라의 사전투표제를 이해하는데 매우 중요하다. 실제로 2013년 선거관리위원회 여론조사 결과, 사전투표를 선택한 유권자가 꼽은 첫 번째 이유는 선거당일 출장·여행계획이었고(이상신 2014), 공식홍보자료집에는 사전투표 참여후 선거일 가족여행을 즐기라는 내용이 포함되었다(nec.go.kr). 가족과의 시간을 보내는 것을 강조하는 최근의 문화적 경향을 고려할 때, 가구구성원과 유권자의 투표행위의 관계를 검토해볼 가치가 충분하지만, 아쉽게도 선행연구들은 인구통계학적 특성이나 정당일체감, 정치관심, 정치성향 등의 정치심리학적 변수에만 집중하고 있다.

선행연구들의 한계점들을 극복하고 새로운 관점에서 사전투표제 도입후 투표참여방식 변화를 살펴보기 위해 본 연구에서는 선행연구들에서는 시도하지 않았던 새로운 이론적·방법론적 접근을 시도하였다.

첫째, ‘투표참여방식 전이확률(probability of vote choice transition)’이라는 새로운 개념을 통해, 특정 선거에서 유권자가 선택한 투표참여방식(기권, 당일투표, 사전투표)이 아닌, 서로 다른 두 시점에서 진행된 선거에서 나타난 투표참여방식 변화를 살펴보았다. 본 연구에서는 투표참여방식 전이확률을 통해 사전투표제 도입이후의 ‘분산효과(당일투표→사전투표)’와 ‘유인효과(기권→사전투표)’를 구분하여 살펴볼 수 있었다. 투표참여방식 전이확률은 다음 섹션에서 보다 구체적으로 소개하였다.

둘째, ‘생태학적 추론(EI) 모형’을 이용하여 중앙선거관리위원회에서 발표하는

‘동’단위 모집단 투표참여 데이터를 기반으로 선거구역별 유권자의 투표참여방식 전이확률을 추정하였다. 즉 모집단 데이터를 이용한 EI 모형을 기반으로 본 연구에서는 소규모의 표본 데이터에서 나타날 수 있는 문제점들을 완화하였으며, 아울러 사전투표제 도입이후 2020년 현재까지의 투표참여방식 전이확률의 변화 패턴을 장기적으로 추적할 수 있었다. EI 모형에 대한 보다 구체적인 설명은 다음 섹션에서 보다 구체적으로 소개하였다.

셋째, EI 모형으로 추정한 투표참여방식 전이확률 데이터와 동별 특성 데이터를 융합(data-fusion)하여, 분산효과와 유인효과가 어떤 선거구역에서 두드러지게 나타나는지 살펴보았다. 이를 통해 개별 유권자를 분석단위로 하는 선행연구들의 연구결과들이 집단수준, 즉 분석단위가 선거구역인 경우에도 발견되는지 살펴보았으며, 특히 선행연구들이 주목하지 않았던 가구구성과 투표참여방식의 관계에 대해서도 실증적 분석을 시도하였다. 선거구역, 즉 ‘동’을 분석단위로 데이터를 융합하는 과정과 이 과정에서 선거구역 간에 나타나는 지역적 자기상관(spatial autocorrelation)을 공변량으로 투입하여 통제한 확장된(extended) EI 모형인 EI-W 모형에 대한 구체적 설명 역시 다음 섹션에 소개되었다.

2. 투표참여방식 전이확률

본 연구에서는 사전투표제 도입이 투표참여율에 미치는 효과가 분산효과에 머무르는지, 아니면 유인효과를 나타내는지 판가름하기 위해 서로 다른 두 시점의 선거에서 유권자의 투표참여방식이 어떻게 바뀌고 있는지 살펴보았다. 분산효과를 주장하는 논리대로라면, 기존 선거에서 기권했던 유권자가 사전투표할 확률은 무시할 수 있을 정도로 매우 낮게 나타날 것이다. 반면 유인효과를 주장하는 논리를 따르면 기존 선거에서 기권했던 유권자가 사전투표방식을 택할 확률이 두드러지게 높게 나타날 것이다. 즉 사전투표제 도입이 투표참여율 변화에 미치는 효과가 분산효과인지 아니면 유인효과인지를 평가하기 위해서는 과거 선거의 투표참여방식과 현재 선거의 투표참여방식을 교차시킨 ‘투표참여방식 전이확률’을 살펴보아야 한다. 보다 구체적으로 사전투표제 도입 직후의 두 선거에서 나타난 투표참여방식 전이확률은 <표 1>에서 제시된 바와 같이 총 6가지다.

〈표 1〉 사전투표제 도입 직후 투표참여방식 전이확률

		사전투표제 도입이전 선거($t-1$)	
		기권	투표
사전투표제 도입직후 선거 (t)	기권	$P(\text{기권}_t \text{기권}_{t-1})$	$P(\text{기권}_t \text{투표}_{t-1})$
	당일투표	$P(\text{당일}_t \text{기권}_{t-1})$	$P(\text{당일}_t \text{투표}_{t-1})$
	사전투표	$P(\text{사전}_t \text{기권}_{t-1})$	$P(\text{사전}_t \text{투표}_{t-1})$
	확률총합	1.00	1.00

〈표 1〉의 투표참여방식 전이확률을 기준으로 할 때, 사전투표제 도입이 투표참여율에 미치는 효과를 둘러싼 논란의 핵심은 $P(\text{사전}_t | \text{기권}_{t-1})$, 즉 ‘기권→사전투표’로 투표방식을 바꿀 확률이다. 즉 분산효과를 주장하는 연구자들은 $P(\text{사전}_t | \text{투표}_{t-1})$ 에 주목하면서 $P(\text{사전}_t | \text{기권}_{t-1})$ 이 0에 가까울 것이라고 주장하는 반면, 유인효과를 주장하는 연구자들은 $P(\text{사전}_t | \text{기권}_{t-1})$ 이 0보다 크며 이 값이 투표참여율 증가에 기여한다고 주장한다.

또한 본 연구에서는 제도변화에 따른 장기적인 투표참여방식 변화 역시 살펴보았다. 사전투표제 도입이후의 투표참여방식 전이확률은 〈표 2〉에서 제시된 바와 같이 총 9가지로 다소 복잡해진다.

〈표 2〉 사전투표제 도입 후 2회차 이후의 투표참여방식 전이확률

		사전투표제 도입직후 선거(t)		
		기권	당일투표	사전투표
사전투표제 도입 후 2회차 선거 ($t+1$)	기권	$P(\text{기권}_{t+1} \text{기권}_t)$	$P(\text{기권}_{t+1} \text{당일}_t)$	$P(\text{기권}_{t+1} \text{사전}_t)$
	당일투표	$P(\text{당일}_{t+1} \text{기권}_t)$	$P(\text{당일}_{t+1} \text{당일}_t)$	$P(\text{당일}_{t+1} \text{사전}_t)$
	사전투표	$P(\text{사전}_{t+1} \text{기권}_t)$	$P(\text{사전}_{t+1} \text{당일}_t)$	$P(\text{사전}_{t+1} \text{사전}_t)$
	확률총합	1.00	1.00	1.00

〈표 1〉과 마찬가지로 〈표 2〉에서도 $P(\text{사전}_{t+1} | \text{기권}_t)$ 의 크기가 분산효과와 유인효과를 가름하는 핵심 측정치다. 더 나아가 〈표 2〉에 제시된 투표참여방식 전이확률을 통해 사전투표제의 장기 효과를 추정할 수도 있다. 예를 들어 사전투

표방식을 고수할 가능성, 즉 $P(\text{사전}_{t+1}|\text{사전}_t)$ 의 값이 1에 가깝다면, 사전투표가 유권자의 선거습관으로 새롭게 정착되고 있다고 해석할 수 있다. 또한 $P(\text{당일}_{t+1}|\text{사전}_t)$ 과 $P(\text{사전}_{t+1}|\text{당일}_t)$ 을 비교해 보면 투표를 결심한 유권자에게 사전투표방식이 어떤 의미를 갖는지를 유추해 볼 수 있다. 이를테면 $P(\text{사전}_{t+1}|\text{당일}_t) > P(\text{당일}_{t+1}|\text{사전}_t)$ 인 결과가 나타날 경우 투표를 결심한 유권자들이 전반적으로 사전투표방식을 점점 더 선호한다는 것을 의미할 것이다. 또한 $P(\text{기권}_{t+1}|\text{사전}_t)$ 와 $P(\text{기권}_{t+1}|\text{당일}_t)$ 을 비교해 보면 사전투표를 선택한 유권자의 선거참여 의지를 비교해 볼 수 있다. 만약 $P(\text{기권}_{t+1}|\text{사전}_t)$ 가 $P(\text{기권}_{t+1}|\text{당일}_t)$ 보다 낮다면, 사전투표에 참여했던 유권자가 전반적으로 투표 참여 의지가 높다고 해석할 수 있다.

본 연구에서는 투표참여방식 전이확률 개념을 중심으로 사전투표제 도입이후 유권자의 투표참여방식이 어떤 변화를 겪고 있는지 살펴보기 위해 아래와 같은 ‘연구문제1’을 제시하였다.

연구문제1: 사전투표제 도입 이후 유권자의 투표참여방식은 어떻게 달라지고 있는가? 특히 ‘분산효과’ 및 ‘유인효과’와 관련 기존 선거 기권자가 사전투표에 참여할 가능성(기권→사전투표)과 기존 선거 투표자가 사전투표할 가능성(투표→사전투표)은 사전투표제 도입 이후에 어떻게 변하고 있는가?

‘연구문제1’은 투표참여방식의 변화 추이를 다루고 있다. 그렇다면 투표참여방식 변화에 영향을 미치는 요인은 무엇일까? 유권자의 사전투표(혹은 편의투표) 참여를 다루는 국내외 선행연구에서는 유권자의 인구통계학적 속성들(성별, 나이, 인종, 교육수준, 소득 등)과 정치심리적 변수들(정당일체감, 이념성향, 정치관심, 정치효능감 등)이 사전투표방식 선택 가능성에 미치는 효과에 집중하고 있다. 국내외 선행연구결과에 따르면 사전투표(혹은 편의투표)를 선택하는 유권자는 전반적으로 나이가 많으며, 특정 정당에 대한 정당일체감이 강하고, 정치지식이나 정치관심도가 높은 편이다(가상준 2017, 2018; 강신구 2016; Berinsky et al. 2001; Garnett 2019; Gronke and Toffey 2008; Stein and Vonnahme 2010). 즉 사전투표 참여자는 전통적인 열성적 투표자의 모습과 상당히 유사하다.

그러나 앞서 언급하였듯, 소규모의 횡단 설문조사 데이터를 기반으로 특정 단일 시점에서 유권자가 선택한 투표참여방식에 주목하는 선행연구로는 ‘기권→사전투표’ 투표참여방식을 바꾼 유권자 집단과 ‘투표→사전투표’로 투표참여방식을 바꾼 유권자 집단을 구분하기 어렵다. 반면 본 연구에서 주목하고 있는 투표참여방식 전이확률의 경우 어떤 유권자가 사전투표방식으로 투표참여방식을 바꾸었는지에 대한 변화에 초점을 맞추었다. 다음 섹션에서 보다 자세히 소개하겠지만, 본 연구의 분석단위는 선거구역, 보다 구체적으로는 서울시의 ‘동’이다. 본 연구에서는 공공에 개방된 동별 지역 특성에 따라 투표참여방식 전이확률이 어떻게 바뀌는지 살펴보았다. 행정안전부와 서울시의 공공 데이터를 기반으로 동의 규모(인구수), 동의 인구유동성(인구유출입 상황), 동별 거주민의 세대분포(20~30대 유권자 비율, 60대 이상 유권자 비율),¹⁾ 저소득층 비율(기초생활수급자 비율), 동별 거주민의 세대구성(3인이상 가구비율) 등에 따라 투표참여방식 전이확률을 살펴보았다. 선거구역별 특성변수에 대한 보다 자세한 설명은 ‘III. 연구방법’에 소개하였다. 선거구역 특성과 투표참여방식 변화의 관계를 살펴보기 위해 아래와 같이 ‘연구문제2’를 설정하였다.

연구문제2: 어떤 특징을 갖는 선거구역에서 사전투표 방식으로 투표방식을 변경하는 유권자 비중이 증가하는가? 구체적으로 기존 선거 기권자가 사전투표에 참여할 가능성(‘기권→사전투표’) 및 기존 선거 투표참여자가 사전투표할 가능성(‘투표→사전투표’)은 어떤 선거구역에서 특히 높게 나타나는가?

3. 생태학적 추론(EI) 모형

EI 모형은 집단 단위로 집계된 데이터(aggregate data)로부터 개인수준의 행동을 추론하는 데이터 분석기법(King 1997; King et al. 2004)으로, 2000년대 이후 해외의 선거연구에서 널리 사용되고 있다(이를테면, Imai and Khana 2016;

1) 본 연구에서는 40~50대 유권자를 기준집단으로 설정하였으며, 20~30대 유권자를 ‘저연령 유권자’로, 60대 이상 유권자를 ‘고연령 유권자’로 상정하였다.

Klima et al. 2019; King et al. 2004).

선거구역별 투표참여율에 대한 모집단 데이터로 확보할 수 있으며, 이는 개별 유권자의 투표행위를 집계한 데이터에 해당한다. 이 과정에서 집단-개인 수준 간 분석단위의 차이가 발생하며, 이로 인해 선거구역 단위로 집계된 선거결과로부터 개별유권자의 투표선택을 설명할 때 오류가 발생할 수 있다. 특히 집단수준의 인과관계가 개인수준의 인과관계와 일치하지 않을 수 있다는 ‘생태학적 오류 (ecological fallacy)’는 사회과학 연구방법에서도 매우 널리 알려져 있다 (Robinson 1950). 그러나 많은 경우 현실적·제도적 한계로 인해 개인수준이 아니라 집단수준 데이터에만 접근할 수 있고, 연구자들은 집단 단위로 집계된 데이터를 기반으로 개인행동을 추론하는, 즉 EI 모형을 사용해야만 하는 상황에 종종 놓이게 된다.

EI 모형은 두 가지 접근, 즉 Duncan and Davis(1953)의 결정론적 접근과 Goodman (1953)의 통계학적 접근을 축으로 발전해왔다. 본 연구에서는 ‘기권’, ‘당일투표’, ‘사전투표’의 세 가지 투표참여방식을 다루지만, EI 모형에 대한 설명의 편의를 위해 ‘투표’와 ‘기권’이라는 두 가지 투표참여방식만을 고려해 보자. 두 시점 간 투표참여방식 전이확률은 아래의 2×2 교차표로 정리할 수 있다.

〈표 3〉 선거구역 i 에서의 두 시점 간 투표참여방식 전이확률

		시점 t 투표방식		
		투표	기권	
시점 $t + 1$ 투표방식	투표	$P(\text{투표}_{t+1} \text{투표}_t)$	$P(\text{투표}_{t+1} \text{기권}_t)$	Y_i
	기권	$1 - P(\text{투표}_{t+1} \text{투표}_t)$	$1 - P(\text{투표}_{t+1} \text{기권}_t)$	$1 - Y_i$
		X_i	$1 - X_i$	

알림1. 선거구역 i 에 대해 X_i 는 시점 t 의 투표율을, Y_i 는 시점 $t + 1$ 의 투표율을 의미한다.

〈표 3〉에 따르면 시점 $t, t + 1$ 에 치러진 두 선거에서 투표참여방식 변화는 총 4가지 경우의 수를 갖는다. EI 모형은 집단수준 데이터인 투표율 X_i 와 Y_i 만이 알려져 있을 때, 이전-현재 선거의 투표참여방식 전이확률을 추정한다. 전체 유권자를 투표 여부에 따라 두 범주로 구분하므로, Y_i 는 또한 다음의 공식으로 표현할

수 있다.

$$Y_i = P(\text{투표}_{t+1} | \text{투표}_t) \times X_i + P(\text{투표}_{t+1} | \text{기권}_t) \times (1 - X_i)$$

$P(\text{투표}_{t+1} | \text{투표}_t)$, $P(\text{투표}_{t+1} | \text{기권}_t)$ 2개 모수를 추정하는 방식에 따라 두 접근방식이 구별된다. 먼저 Duncan and Davis(1953)의 결정론적 접근에서는 각 모수가 가질 수 있는 상한과 하한을 이용해 경계(bound)를 계산한다. 반면, Goodman(1953)의 통계학적 접근에서는 X_i 를 독립변수로 Y_i 를 종속변수로 설정한 후 생태학적 회귀(ecological regression) 분석을 실시한다. 즉 관측값인 (X_i, Y_i) 와 결측값인 모수들의 관계에 대해 전자는 결정론적인 관점을, 후자는 통계적 또는 확률적인 관점을 취한다. 두 접근방식은 약 50년간 상호 보완적으로 발전해오다가, King과 동료들의 연구를 통해 통합된 접근방식으로 발전하였다(King 1997; King et al. 1999; King et al. 2004; Rosen et al. 2001). King(1997)의 EI 모형은 Goodman(1953)의 통계적 접근을 기반으로, 절단된 이변량 정규분포를 가정하는 방식으로 Duncan and Davis(1953)의 경계 정보를 모형추정과정에 반영했다.

King과 동료들의 EI 모형은 다양한 형태로 확장되었다. 특히 이항-베타 위계모형(binomial-beta hierarchical model, King et al. 1999)은 베이지안 관점에서 모수들에 대해 삼단 위계를 가정함으로써 마르코프 연쇄 몬테카를로(Markov Chain Monte Carlo, MCMC)기법을 적용할 수 있게 되었다. 달리 표현하면, 경계 정보가 반영된 분포 가정 아래 2×2 교차표의 모수들을 더욱 효율적으로 추정할 수 있게 되었다. 이는 또한 $R \times C$ 상황(이때 $R, C \geq 2$)에서 다항-디리클레 위계모형(multinomial-Dirichlet hierarchical model, Rosen et al. 2001)으로 확장될 수 있다. 본 연구에서는 다항-디리클레 위계모형을 채택하였다.

EI 모형추정결과가 타당하기 위해서는 세 가정이 필요하다(King 1997). 첫째, 추정확률의 쌍이 동일한 집단에서 나와야 하고, 둘째, 집단과 집단 사이에는 자기상관(autocorrelation)이 없어야 하며, 셋째, 집단수준의 관측치(즉 투표방식별 참여자수)가 EI 모형 추정확률과 독립적이어야 한다. 이러한 세 가정이 충족된 EI 모형을 흔히 ‘기본(basic) EI 모형’이라고 부른다. 그러나 선거연구에서는 선거구

역(집단)간 자기상관관계가 없다고 가정하기 어려우며(Anselin and Cho 2002), 이 상황은 한국선거에서도 마찬가지다. 기본 EI 모형의 가정을 확보하기 어렵기 때문에, 본 연구에서는 ‘확장된(extended) EI 모형’을 채택하였다. 확장된 EI 모형에서는 모수를 설명할 수 있는 공변량을 모형에 포함시켜 가정사항을 완화한다(Klima et al. 2019). 본 연구에서는 동시자기회귀(simultaneous autoregressive, SAR) 모형을 이용하여 지역간 자기상관을 반영한 공변량을 추정·투입한 확장된 EI 모형을 사용하였다.²⁾

EI 모형으로 추정한 추정치(‘투표선택방식 전이확률’)와 집단특성변수(‘선거구역 특성’)의 연관관계를 추정하는 것도 가능하다. 예를 들어 <표 3>에서 $P(\text{투표}_{t+1} | \text{기권}_t)$, 즉 직전 선거에서 기권한 사람이 투표할 확률이 선거구역별 특성에 따라 어떻게 다른 패턴을 띠는지 알아볼 수 있다. 이 경우 EI 모형으로 추정한확률을 얻는 1단계와 추정한확률을 이용해 회귀분석을 실시하는 2단계로 이루어진다. King과 동료들이 제안한 EI-W 모형은 EI 추정 후 2단계에서 가중최소자승(weighted least squares, WLS) 기법을 이용한다(Adolph et al. 2003).³⁾

결론적으로 본 연구에서는 지역간 자기상관을 통제한 확장된 EI 모형으로 투표 참여방식 전이확률을 추정한 후, 이렇게 추정한 전이확률을 종속변수로 선거구역 특성변수들을 독립변수로 투입한 후 WLS를 적용한 EI-W 모형을 사용하였다.

2) 다항-디리클레 위계모형은 또한 3수준 모수의 사전분포(prior distribution)에 영향을 받는 것으로 알려져 있다(Wakefield 2004). 본 연구에서는 King과 동료들(Rosen et al. 2001)의 제안대로 감마 분포(gamma distribution), 구체적으로 $\Gamma(4, 2)$ 를 3수준 모수의 사전분포로 설정하였다. 선행연구(Klima et al. 2019)에 따르면 이것은 일종의 ‘무정보(non-informative) 사전분포’로 활용가능하다. 본 연구에서 사용한 모형의 경우, 2수준에서 공변량을 투입했을 때 3수준 분포가정에 따른 추정결과의 차이가 크지 않았다. 따라서 무정보 사전분포로써 $\Gamma(4, 2)$ 가 적절하다고 판단하였다.

3) King과 동료들의 시뮬레이션 연구(Adolph et al. 2003)에 따르면 EI-W 모형은 다양한 상황에서 아주 미미한 오차만을 내는 것으로 나타났다. 물론 데이터마다 구체적인 모형가정, 가정사항 충족 여부, $R \times C$ 상황으로의 확장가능성 등이 다르기에 해당 논문에서 모형추정 결과 해석에 주의를 요구한 것이 사실이다. 그러나 EI-W 관련 논쟁은 연구자들 사이에 ‘합의’가 이루어졌으며, 본 연구에서는 오차가능성보다 연구주제와의 적합성이 더 크다는 판단 아래 EI-W 모형을 채택하였다.

Ⅲ. 연구방법

1. 데이터 수집범위

본 연구에서는 지역적으로는 서울시, 시간적으로는 사전투표제 도입이전인 2012년 실시된 제19대 국회의원 선거와, 사전투표제 도입이후 실시된 2016년 제20대, 2020년 제21대 국회의원 선거의 ‘동’단위 투표참여데이터와 ‘동’별 선거구역 특성 데이터를 수집하였다. ‘동’은 집단 수준에서 가장 소규모의 동질적 행정단위이며 동일 지역구나 행정구 내에서도 정치적·사회적 특성은 동에 따라 다를 수 있다는 점(손낙구 2010)을 고려하였다.

우선 선거유형을 국회의원 선거에 한정하였다. 선거관련 선행연구에 따르면 대통령 선거와 같이 선거관심이 높은 선거는 지방선거와 같은 선거관심이 낮은 선거에 비해 투표참여율이 높다(Verba and Nie 1972). 선거유형에 따른 투표참여율의 변동성을 통제하기 위해 본 연구에서는 국회의원 선거에서 나타난 투표참여방식만 살펴보았다.

다음으로 선거구역 단위로 서울시의 ‘동’을 선택하였다. 물론 국회의원 선거는 전국단위 선거이기 때문에 전국의 모든 읍면동을 살펴보는 것도 불가능하지는 않다. 그러나 다음과 같은 이유들을 고려하여 본 연구에서는 서울시의 투표참여방식 변화만을 살펴보았다. 첫째, 지방의 경우 인구유출에 따른 행정구역 변화⁴⁾가 상대적으로 심하기 때문에 투표참여방식 변화를 추정하기 어려운 반면, 서울시의 경우 4개 동을 제외하고는 행정구역이 일정하게 유지되었기 때문에 안정적 투표참여방식 변화 추정이 가능하다. 둘째, 서울시의 동의 경우 평균거주민 수가 약 2만 명 가량으로 안정된 EI 추정치를 얻기 쉽다. 셋째, 다른 지역과 비교할 때 서울지역의 동별 선거구역 특성변수를 확보하기 용이했다.

끝으로 시간적으로 사전투표제가 도입되기 직전의 2012년 제19대 국회의원 선거와 도입이후인 2016년 제20대, 2020년 제21대 국회의원 선거들을 살펴보았다.

4) 특히 20대 국회의원 선거의 경우 헌법재판소 판결이후 지역구의 인구편차를 2:1 이내로 조정하는 과정에서 행정구역이 크게 변화하였다.

제19-20대 국회의원 선거의 투표참여방식 변화를 통해 사전투표제 도입 직후에 나타난 투표방식변화를 살펴보았으며, 제20-21대 국회의원 선거의 투표참여방식 변화를 통해 사전투표제 도입에 따른 투표방식의 변화가 어떻게 정착되고 있는지 그리고 두 차례의 투표참여방식 변화 데이터에서 어떤 결과가 일관적으로 관측되는지 살펴보았다.

2. 연구데이터 수집

본 연구의 데이터는 크게 두 가지 유형으로 구분할 수 있다. 첫 번째 데이터는 사전투표제도 등장이후 유권자의 투표참여방식 변화, 즉 투표참여방식 전이확률을 추정하기 위한 선거구역별 투표참여방식 분포자료다. 이를 위해 ‘동’단위 사전투표제 도입 이전 제19대 국회의원 선거결과와 사전투표제 도입 이후 제20대, 제21대 국회의원 선거결과를 수집하였다. 제19-21대 국회의원 선거결과는 ‘중앙선거관리위원회 선거통계시스템(<http://info.nec.go.kr/>)’에서 수집하였다. 해당자료에는 동별 선거인수와 함께 관내 사전투표자수(사전투표제도 등장이후인 제20대, 제21대 국회의원 선거의 경우)와 공식선거일 당일투표수 정보가 담겨 있다. 본 연구에서는 제19대 국회의원 선거의 경우 동별 유권자들의 투표참여방식을 ‘기권’, ‘당일투표’의 2개 범주로, 제20대와 제21대 국회의원 선거의 경우 ‘기권’, ‘당일투표’, ‘사전투표’의 세 범주로 구분하였다. 세 번의 국회의원 선거 시점간 행정시스템 개편으로 인해 ‘동’이 신설되거나 통폐합된 네 개 선거구역⁵⁾의 경우 부득이하게 분석에서 제외하였다.

두 번째 데이터는 ‘동’별 선거구역 특성변수 자료다. 본 연구에서는 ‘행정안전부(<http://27.101.213.4/index.jsp>)’와 ‘서울시(<https://data.seoul.go.kr/>)’의 공개 데이터를 사용하였다. 먼저 행정안전부의 공개 데이터에서는 동별 주민등록 인구통계자료를 이용하여 제19, 20, 21대 선거당시의 동단위 세대별 거주민수 데이터를 추출하였다. 추출된 동별 거주민수는 ‘선거구역 규모’를 추정하는 변수로 사용되

5) 종로구 명륜3가동(2012년 해화동에 통합), 송파구 위례동(2015년 신설), 구로구 항동(2020년 신설), 송파구 둔촌제1동(2020년 재건축으로 인해 둔촌제2동에 통합 운영) 등 4개를 제외하였다.

었으며, 또한 직전 선거대비 현재 선거당시 거주민수의 비율을 이용하여 ‘선거구역 인구유출입’ 변수를 구성하였다.⁶⁾ 아울러 해당 데이터의 세대별 데이터를 이용하여 ‘저연령 유권자 비율’과 ‘고연령 유권자 비율’ 변수도 생성하였다. 본 연구에서는 40~50대 유권자를 기준집단으로 20~30대 유권자를 ‘저연령 유권자’로 60대 이상 유권자를 ‘고연령 유권자’로 정의하였다.

서울시의 공개 데이터에서는 동단위 ‘국민기초생활보장 수급자수’, ‘세대원수별 세대수 통계’의 두 가지 자료들을 이용해 ‘저소득층 비율’, ‘3인이상 다인가구 비율’ 변수들을 구성하였다. 첫째, ‘저소득층 비율’ 변수는 앞서 소개한 동별 거주민수 데이터를 토대로 거주민 10,000명당 국민기초생활보장 수급자수를 계산하는 방식으로 구성하였다. 둘째, ‘3인이상 다인가구 비율’ 변수는 동별 전체 가구수를 기준으로 3인이상 다인가구의 비율을 계산하는 방식으로 구성하였다. 서울시 소재 ‘동’별 선거구역 특성변수들의 자세한 기술통계치는 <표 4>에 제시되어 있다.

<표 4> 제20대, 제21대 국회의원 선거의 서울시 소재 ‘동’별 선거구역 특성변수 기술통계치

	제20대		제21대	
	평균 (M)	표준편차 (SD)	평균 (M)	표준편차 (SD)
선거구역 규모 (상용로그변환)	0.33	0.22	0.32	0.21
선거구역 인구유출입 (상용로그변환)	-0.02	0.09	-0.01	0.09
저연령(20~30대) 유권자 비율	0.31	0.05	0.32	0.06
고연령(60대이상) 유권자 비율	0.19	0.04	0.23	0.04
저소득층 비율	0.03	0.02	0.03	0.03
3인이상 다인가구 비율	0.43	0.12	0.38	0.12

알림. ‘선거구역 규모’와 ‘선거구역 인구유출입’ 변수들은 상용로그를 이용해 변환되었다.

- 6) 예를 들어 제19-20대 선거에서 ‘ i 번째 선거구역 인구유출입’($i = 1, 2, 3, \dots$) 변수는 다음 공식에 따라 산출되었다.

$$\text{제19-20대 선거구역 인구유출입}_i = \frac{\text{제20대 선거시 거주민수}_i}{\text{제19대 선거시 거주민수}_i}$$

즉 ‘제19-20대 선거구역 인구유출입’ 지수가 1보다 클 경우 ‘인구유입’이, 1보다 작을 경우 ‘인구유출’이 일어난 선거구역이다. ‘제20-21대 선거구역 인구유출입’ 변수 역시 유사한 방식으로 계산하였다.

3. 데이터분석방법: 지역간 자기상관을 통제한 확장된 EI-W 모형

앞서 설명하였듯, 본 연구에서는 지역간 상관관계를 통제한 후 EI를 통해 얻은 추정확률이 집단수준 변수, 즉 선거구역의 특성에 따라 어떻게 달라지는지를 테스트하는 확장된 EI 모형, 보다 구체적으로 ‘확장된 EI-W 모형’을 사용하였다(Adolph et al. 2003; King et al. 2004). 확장된 EI-W 모형은 2단계로 진행하였다. 1단계에서는 각 선거구역(여기서는 ‘동’단위)별로 공식집계된 투표결과를 수집한 후, 앞서 소개한 <표 1>과 <표 2>에서 제시한 투표참여방식 전이확률을 추정하였다. 확장된 EI 모형을 추정할 때, 본 연구에서는 동단위의 선거구역들의 지역간 자기상관관계를 통제하였다(Anselin and Cho 2002).⁷⁾ 또한 확장된 EI 모형을 통해 안정된 투표참여방식 전이확률을 추정하기 위해, 본 연구에서는 시물레이션 연구결과(Klima et al. 2016)를 기반으로 총 5,000번의 번인(burn-in), 10번의 씨닝(thinning) 조건에서, 각 선거구역별로 10,000번의 투표참여방식 전이확률값을 추정하였다.⁸⁾ 1단계이후, 제19-20대 총선의 경우, 422개 동별 6개의 투표참여방식 전이확률을 10,000번 추정하여 총 25,320,000개($= 422 \times 6 \times 10,000$)의 투표참여방식 전이확률을 얻었으며, 제20-21대 총선의 경우 총 37,980,000개($= 422 \times 9 \times 10,000$)의 투표참여방식 전이확률을 얻었다.

다음으로 2단계에서는 이렇게 얻은 투표참여방식 전이확률값을 종속변수로, 각 선거구역별 특성변수들을 예측변수로 하는 회귀모형을 추정하였다. 본 연구에서는 EI-W 모형 개발자인 King과 동료들의 제안을 따라 WLS 기법을 이용하여 선거구역별 특성변수가 투표참여방식 전이확률에 미치는 효과를 추정하였다(Adolph et al. 2003).⁹⁾

7) 제20대, 제21대 국회의원 선거의 투표 데이터의 경우 지역간 자기상관 수준을 나타내는 모란 I지수(Moran's I)가 각각 0.459($p < .001$), 0.407($p < .001$)로 나타나 지역간 자기상관관계가 뚜렷하게 관측되었다.

8) MCMC 시물레이션에서 초기값의 영향을 줄이기 위해 초기 5,000번의 추정치는 제거하고(burn-in), 10번 간격으로 포함시키는 방식으로(thinning) 총 10,000번의 전이확률값을 추정한다는 의미다.

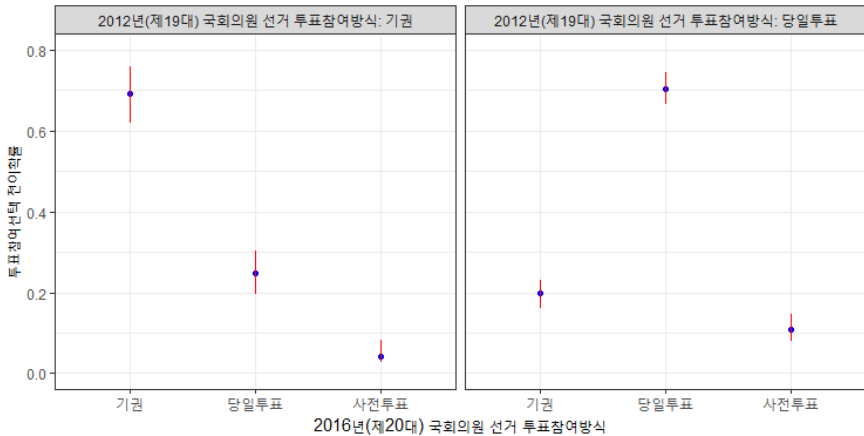
9) 본 연구의 모든 데이터 분석은 R(version 3.6.3)을 기반으로 하였다. 또한 선거구역들의 지역간 자기상관을 통제하기 위해 spdep 패키지(version 1.1-5)와 spatialreg 패키지

IV. 연구결과

1. 제19-20대, 제20-21대 총선 투표참여방식 전이확률 추정결과

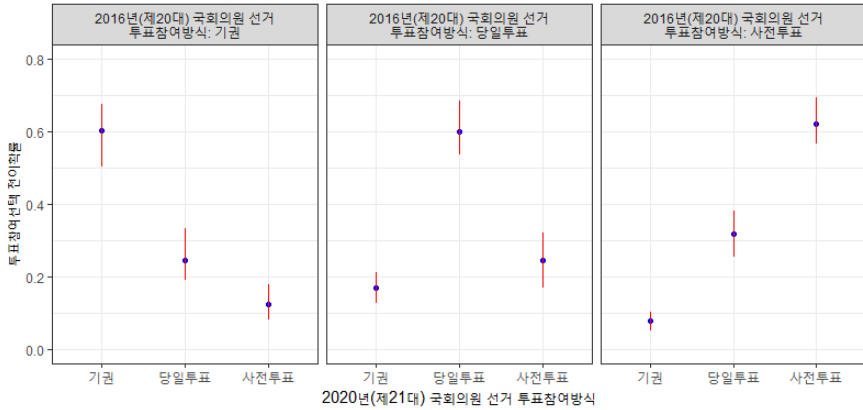
‘연구문제1’에 대한 경험적 해답을 살펴보기 위해 어느 정도 규모로 ‘사전투표’로 투표방식이 전이되는지 살펴보았다. 제19대와 제20대 총선의 동별 투표방식 집계 데이터를 토대로 확장된 EI 모형을 이용하여 제19-20대 총선의 투표참여방식 전이확률 6가지를 추정한 결과는 <그림 1>에, 제20-21대 총선의 투표참여방식 전이확률 9가지 추정결과는 <그림 2>에 제시되어 있다.

<그림 1> 제19-20대 총선 투표참여방식 전이확률 점추정치와 90% 신뢰구간(CI)



(version 1.1-5)를, EI-W 모형을 추정하기 위해 ei 패키지(version 1.3-3)와 eiPack 패키지(version 0.2-1)를 이용하였으며, 데이터 전처리와 추정결과 시각화를 위해 tidyverse 패키지(version 1.3.0)를 이용하였다.

〈그림 2〉 제20-21대 총선 투표참여방식 전이확률 점추정치와 90% 신뢰구간(CI)



먼저 〈그림 1〉의 좌측 패널과 중앙 패널의 패턴은 상식적 결과, 즉 직전 선거에서 기권한 사람은 기권할 확률이 높고, 투표에 참여한 사람은 당일투표에 참여할 확률이 높다는 것을 잘 보여준다. ‘연구문제1’과 직접적으로 관련된 결과는 〈그림 1〉의 우측 패널에서 찾아볼 수 있다. EI 추정결과 제19대 총선(2012년)의 기권자가 ‘사전투표’를 선택할 전이확률은 약 0.043 [90% CI, (0.026, 0.083)]였지만, 투표자가 ‘사전투표’를 선택할 확률은 약 0.109 [90% CI, (0.080, 0.147)]였다. 제19대 국회의원 투표참여율이 약 54%인 것을 감안하면, ‘분산효과’를 보인 유권자수와 ‘유인효과’를 보인 유권자 수의 비율을 약 3:1이며, 이 결과는 선행연구에서 보고된 사전투표에 참여한 두 유권자집단(즉, ‘잠재적 일반투표자’와 ‘순수사전투표자’)의 비율과 거의 동일하다(김찬송 외 2016, 6쪽). 즉 국회의원 선거에서 사전투표제도가 처음 도입되었을 때, ‘분산효과’는 ‘유인효과’보다 두드러지게 나타났다.

다음으로 〈그림 2〉의 주요 발견들은 다음과 같다. 첫째, 〈그림 2〉의 좌측패널에서 잘 나타나듯, ‘당일투표→기권’ 전이확률은 0.170 [90% CI, (0.127, 0.211)]인 반면, ‘사전투표→기권’ 전이확률은 0.078 [90% CI, (0.053, 0.103)]으로 나타났다. 이를 통해 당일투표 참여자에 비해 사전투표 참여자는 투표참여 의지가 더 강하다는 선행연구결과(가상준 2018)를 다시금 확인할 수 있었다.

둘째, 〈그림 2〉의 중앙 패널의 제20-21대 ‘당일투표→사전투표’ 전이확률은 0.245 [90% CI, (0.170, 0.322)]로, 제19-20대 ‘당일투표→사전투표’ 전이확률인 0.109 [90% CI, (0.080, 0.147)] 보다 2배 이상 높게 나타났다. 아울러 〈그림 2〉의 우측 패널의 ‘사전투표→사전투표’ 전이확률은 0.623 [90% CI, (0.565, 0.695)]로 매우 높게 나타나, 과거 선거에서 사전투표를 경험해 본 유권자들은 사전투표를 계속하여 유지하려는 경향이 매우 강한 것을 확인할 수 있었다. 또한 〈그림1〉의 우측패널과 마찬가지로 ‘당일투표→사전투표’ 전이확률(0.245, [90% CI, (0.170, 0.322)])이, ‘기권→사전투표’ 전이확률보다 더 높게 나타났다(0.125, [90% CI, (0.081, 0.180)]). 2016년 제20대 국회의원 선거의 당일투표자 비율이 약 46%이며, 기권자 비율이 약 42%라는 점을 감안하면, 분산효과를 보인 유권자수와 유인효과를 보인 유권자수는 약 2:1이다. 이 비율은 앞서 설명했던 사전투표제 도입 직후 실시된 국회의원 선거에서 얻었던 분산효과 대비 유인효과 유권자수 비율인 3:1 보다 작다. 다시 말해 이 결과는 사전투표제가 도입되고 시간이 흐르면서, 유권자층의 일부가 투표습관을 사전투표방식으로 바꾸면서(가상준 2018; Stein and Vonnahme 2010) 분산효과는 조금씩 감소하고 있다고 해석할 수 있다.

‘연구문제1’ 관련 결과를 요약하면 다음과 같다. 첫째, 국회의원 선거에 사전투표제도가 도입된 이후, 전반적으로 당일투표에서 사전투표로 선택을 바꾸는 ‘분산효과’가 두드러지게 나타나고 있다. 둘째, 기권에서 사전투표로 선택을 바꾸는 ‘유인효과’는 전반적으로 매우 작다. 셋째, 한번 사전투표 방식을 채택한 유권자는 계속하여 사전투표 방식을 유지하려는 경향을 보인다.

2. 투표참여방식 전이확률과 선거구역 특성변수들과의 관계

앞에서는 두 차례 국회의원 선거에서 나타난 투표참여방식 전이확률을 통해 ‘분산효과’와 ‘유인효과’의 전반적 형태를 살펴보았다. 다음으로 ‘연구문제2’에 대한 경험적 해답을 얻기 위해 서울시 ‘동’단위 선거구역 특성에 따라 투표참여방식 전이확률이 어떻게 달라지는지 위한 EI-W 모형을 추정하였다. 앞서 설명하였듯, 제 19-20대의 경우 총 6개의 투표참여방식 전이확률을, 제20-21대의 경우 총 9개의 투표참여 전이확률을 얻을 수 있다. 여기서는 선거구역 특성 변수들과 ‘분산효과’

와 ‘유인효과’와 관련된 투표참여 전이확률의 연관관계를 EI-W 모형으로 추정한 결과만을 제시하였다.¹⁰⁾ 제19-20대, 제20-21대 국회의원 선거에서 나타난 투표참여방식 전이확률이 선거구역 특성변수들과 어떻게 관련되어 있는지 살펴본 EI-W 모형 추정결과는 <표 5>와 같다.

<표 5> 선거구역 특성과 분산효과와 유인효과를 나타내는 투표참여방식 전이확률의 관계 추정결과

	제19-20대 국회의원 선거 투표참여방식 전이확률		제20-21대 국회의원 선거 투표참여방식 전이확률		
	유인효과	분산효과	유인효과	분산효과	
	기권→ 사전투표	투표→ 사전투표	기권→ 사전투표	당일투표→ 사전투표	사전투표→ 사전투표
절편	0.039*** (0.001)	0.117*** (0.001)	0.103*** (0.001)	0.250*** (0.002)	0.756*** (0.001)
선거구역 규모	-0.043*** (0.005)	-0.046*** (0.005)	-0.090*** (0.007)	-0.171*** (0.011)	-0.075*** (0.005)
선거구역 인구유출입	0.024 (0.014)	-0.014 (0.011)	-0.018 (0.023)	-0.109** (0.035)	-0.029 (0.016)
저연령 유권자 비율	0.195*** (0.028)	0.144*** (0.034)	0.109* (0.047)	0.114 (0.077)	0.044 (0.034)
고연령 유권자 비율	0.128*** (0.035)	0.060 (0.044)	0.044 (0.052)	-0.042 (0.085)	-0.004 (0.037)
저소득층 비율	0.020 (0.035)	-0.019 (0.050)	0.024 (0.063)	0.080 (0.105)	0.030 (0.047)
3인이상 다인가구 비율	0.103*** (0.013)	0.061*** (0.016)	0.144*** (0.025)	0.085* (0.041)	0.036 (0.018)

알림1. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

알림2. 서울시 소재 422개 동에서 얻은 선거관리위원회 공식통계치를 기반으로 지역간 자기상관을 통제한 확장된 EI 모형(King et al. 2004) 모형으로 투표참여방식 전이확률을 추정한 후, 이를 WLS 회귀분석으로 추정한(Adolph et al. 2003) 회귀계수와 표준오차를 제시하였다.

알림3. ‘선거구역 규모’는 ‘만명 단위 선거구역 거주민수’를, ‘선거구역 인구유출입’는 ‘제19대 총선 당시 선거구역 거주민수 대비 제20대 총선 당시 선거구역 거주민수의 비율’(1보다 클 경우 인구유입 상황, 1보다 작을 경우 인구유출 상황)을, ‘저연령 유권자 비율’은 ‘20~30대 거주민 비율’을 그리고 ‘고연령 유권자 비율’은 ‘60세 이상 거주민 비율’을, ‘저소득층 비율’은 ‘만명 기준 기초생활수급자 수’를, ‘3인이상 다인가구 비율’은 ‘전체가구수 대비 3인이상 다인가구의

10) 본문에 제시되지 않은 다른 형태의 투표참여 전이확률과 선거구역 특성변수들의 연관관계를 추정한 EI-W 모형 추정결과는 저자들에게 별도 문의하면 얻을 수 있다.

비율'을 의미한다. 또한 절편값 해석을 위해 모든 예측변수들에 대해 평균중심화 변환을 실시한 후 투입하였다.

〈표 5〉의 확장된 EI-W 모형 추정결과를 살펴보자. 첫째, 선거구역 규모가 클수록 '기관→사전투표' 전이확률(제19-20대, $b = -0.043$, $p < .001$; 제20-21대, $b = -0.099$, $p < .001$)과 '투표→사전투표' 전이확률(제19-20대, $b = -0.046$, $p < .001$; 제20-21대 '당일투표→사전투표', $b = -0.171$, $p < .001$, '사전투표→사전투표', $b = -0.075$, $p < .001$)이 일괄적으로 감소하는 것으로 나타났다. 즉 이전 선거에서 투표참여를 했던 기권을 했던 상관없이 선거구역이 크면 클수록 사전투표 투표로 전이될 확률이 낮게 나타났다.

둘째, 선거구역의 인구유출입과 '기관→사전투표' 전이확률의 관계는 제19-20대, 제20-21대 모두 통계적으로 유의미한 관계가 나타나지 않았다. 아울러 선거구역의 인구유출입과 '투표→사전투표' 전이확률은 제20-21대에서의 '당일투표→사전투표' 전이확률($b = -0.109$, $p < .01$)을 뺀 나머지 경우, 통계적으로 유의미한 관계는 나타나지 않았다. 즉 두 차례 선거 모두에서 일관된 경향성은 나타나지 않았다는 점에서 서울시의 동단위 선거구역 인구유출입 규모는 사전투표로의 투표전이와 큰 상관이 없었다.

셋째, 20-30대 젊은 유권자 비율이 높은 선거구역일수록 '기관→사전투표' 전이확률, 즉 유인효과가 높게 나타났다(제19-20대, $b = 0.195$, $p < .001$; 제20-21대, $b = 0.109$, $p < .05$). 이 결과는 젊은 유권자일수록 사전투표제 덕분에 선거에 참여하게된 '순수사전투표자'일 가능성이 높아진다는 선행연구결과(김찬송 외 2016)와 일치한다. 또한 제19-20대 총선의 경우 저연령 유권자 비율이 높은 선거구역일수록 '투표→사전투표' 전이확률이 통계적으로 유의미하게 높게 나타났지만($b = 0.144$, $p < .001$), 이러한 연관관계는 제20-21대 총선에서는 다시 확인되지 않았다.

넷째, 고연령 유권자 비율이 높은 선거구역일수록 '기관→사전투표' 전이확률, 즉 유인효과가 높게 나타났으나($b = 0.128$, $p < .001$), 이 효과는 제19-20대 총선에서만 나타났으며, 제20-21대 총선에서는 나타나지 않아 일관된 경향성을 찾기 어려웠다. 또한 고연령 유권자 비율과 '투표→사전투표' 전이확률은 두 차례

총선 모두에서 통계적으로 뚜렷한 관계를 갖지 않는 것으로 나타났다. 언뜻 보면 이 결과는 개인수준 설문조사 데이터 분석결과를 기반으로 고연령자일수록 사전투표에 더 적극적으로 참여한다는 선행연구결과(가상준 2017, 2018)와 다른 것처럼 보일 수도 있다. 그러나 사실 이 결과는 선행연구결과와 부합된다. 왜냐하면 선행연구결과에서도 사전투표 참여자는 기권자와 비교할 때 평균연령이 높을 뿐 당일선거 참여자와 비교하면 평균연령이 거의 비슷하기 때문이다(예를 들어 가상준 2018, 112쪽).

다섯째, 기초생활 수급자 비율로 측정한 저소득층 비율 역시 두 차례의 국회의원 선거에서 ‘기권→사전투표’ 전이확률 및 ‘투표→사전투표’ 전이확률과 뚜렷한 연관관계를 갖지 않는 것으로 나타났다. 즉 사전투표 제도 도입이 저소득층 유권자의 투표참여를 독려한다고 보기 어렵다.

여섯째, 3인이상 다인가구의 비율이 높을수록 유인효과를 나타내는 ‘기권→사전투표’ 전이확률은 두 차례 선거에서 모두 높아졌으며(제19-20대, $b = 0.103$, $p < .001$: 제20-21대, $b = 0.144$, $p < .001$), 분산효과를 나타내는 ‘투표→사전투표’(제19-20대, $b = 0.061$, $p < .001$)와 ‘당일투표→사전투표’(제20-21대, $b = 0.085$, $p < .05$)의 전이확률도 다소 증가하는 모습을 보였다. 즉 다인가구 비중이 높은 선거구역일수록 유인효과와 분산효과가 모두 증가하는 것으로 나타났으며, 유인효과가 분산효과보다 더 두드러지게 나타났다.

〈표 5〉의 결과에서 나타난 유인효과와 분산효과를 보다 구체적으로 살펴보기 위해 본 연구에서는 ‘분산효과 대비 유인효과 비율’과 선거구역 특성변수의 관계를 추가로 추정해 보았다. ‘분산효과 대비 유인효과 비율’을 이용하면 직전선거 투표자에 비해 직전선거 기권자가 사전투표방식을 선택할 확률이 어느 선거구역에서 두드러지게 나타나는가를 살펴볼 수 있다. 이를 위해 유인효과를 나타내는 ‘기권→사전투표’ 전이확률을 분산효과를 나타내는 ‘투표→사전투표’ 전이확률로 나누고 로그값을 취한 값¹¹⁾을 종속변수로 한 후 EI-W 모형을 추정하였다. 모형추

11) 해당 추정치에 적용한 공식은 아래와 같다(t 는 선거시점, i 는 선거구역을 나타낸다).

$$\text{분산효과 대비 유인효과 비율}_{ti} = \log \frac{P(\text{사전투표}_{ti} \mid \text{기권}_{(t-1)i})}{P(\text{사전투표}_{ti} \mid \text{투표}_{(t-1)i})}$$

‘분산효과 대비 유인효과 비율’은 로지스틱 회귀분석의 ‘로그오즈(log-odds)’나 범주형 데이

정결과는 아래의 〈표 6〉과 같다.

〈표 6〉 선거구역 특성과 직전선거 투표자 대비 기권자의 사전투표 선택확률비의 관계 추정결과

	제19-20대 국회의원 선거	제20-21대 국회의원 선거
절편	-1.176*** (0.015)	-1.079*** (0.007)
선거구역 규모	-0.755*** (0.090)	-0.280*** (0.040)
선거구역 인구유출입	0.794*** (0.204)	0.124 (0.138)
저연령 유권자 비율	3.740*** (0.580)	0.891** (0.294)
고연령 유권자 비율	2.337** (0.761)	0.554 (0.340)
저소득층 비율	0.453 (0.855)	0.170 (0.396)
3인이상 다인가구 비율	2.135*** (0.284)	1.255*** (0.157)

알림1. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

알림2. 모수추정치에 대한 설명은 〈표 5〉의 알림2를 참조하라.

알림3. 선거구역 특성변수의 의미에 대한 보다 자세한 설명은 〈표 5〉의 알림3을 참조하라.

알림4. 비교를 위하여 제20-21대 국회의원 선거에서의 유인효과 대비 분산효과 비율 추정치를 산출할 때 제20대 국회의원 선거에서 당일투표자와 사전투표자를 ‘투표자’로 통합하였다.

〈표 6〉에서 주목할 결과는 크게 네 가지다. 첫째, ‘절편’값에서 잘 드러나듯, 두 차례 국회의원 선거에서 모두 유인효과는 분산효과보다 그 크기가 작았다(제19-20대, $b_0 = -1.176$; 제20-21대, $b_0 = -1.079$). 둘째, 선거구역의 규모가 클수록 ‘유인효과’보다는 ‘분산효과’가 더 두드러지게 나타났다(제19-20대, $b = -0.755$, $p < .001$; 제20-21대, $b = -0.280$, $p < .001$). 셋째, 20~30대의 젊은 유권자들이 많은 선거구역일수록 분산효과보다 유인효과가 더 두드러지게 나타났다(제19-20대, $b = 3.740$, $p < .001$; 제20-21대, $b = 0.891$, $p < .01$). 이 결과는 설문

터 분석 모형에서 흔히 등장하는 ‘로그상대위험도(log relative risk)’와 개념적으로 동일하다. 즉 분산효과를 나타내는 전이확률과 유인효과를 나타내는 전이확률이 동일할 경우 해당 추정치의 값은 0이 된다(즉, $0 = \log(1)$)

조사 데이터 분석을 기반으로 ‘유인효과’를 주장한 김찬송 외(2016)의 연구결과와 유사하다. 넷째, 3인 이상 다인가구 비율이 높을수록 ‘분산효과’ 보다는 ‘유인효과’가 매우 두드러지게 나타났다(제19-20대, $b = 2.135$, $p < .001$: 제20-21대, $b = 1.255$, $p < .001$).

‘연구문제2’와 관련 사전투표제 도입효과에 대한 분석결과를 정리하면 다음과 같다. 첫째, 사전투표방식을 선택한 유권자들에게서 분산효과가 유인효과보다 더 두드러지게 나타났다. 둘째, 선거구역 규모가 클수록 분산효과와 유인효과가 더 작게 나타났으며, 이 두 효과 중 유인효과가 더욱 크게 감소하였다. 셋째, 젊은 유권자들이 많은 선거구역의 경우 유인효과가 두드러지게 나타났다. 넷째, 3인 이상의 다인가구 비율이 높은 선거구역일수록 유인효과가 매우 두드러지게 나타났다.

V. 논의 및 결론

사전투표제 도입과 투표참여율 변화의 관계를 다룬 국내의 선행연구들은 횡단 설문조사 데이터를 기반으로 유권자의 투표방식선택에 집중하였다. 투표선택방식 전이확률이라는 새로운 개념과 확장된 EI-W 모형이라는 새로운 분석기법을 시도한 본 연구의 주요 연구결과와 사전투표제 도입이 투표참여율 및 참여방식 변화에 끼친 효과를 살펴보면 다음과 같다.

첫째, 사전투표제 도입이후 분산효과(‘당일투표→사전투표’)가 유인효과(‘기권→사전투표’) 보다 더 두드러지게 나타났다. 또한 사전투표방식을 선택했던 유권자들은 이후의 투표에서도 사전투표방식을 강하게 고수하는 것으로 나타났다. 사전투표제 도입 후 첫 번째로 실시된 20대 국회의원 선거의 경우, 분산효과를 보인 유권자수와 유인효과를 보인 유권자수는 약 3:1의 비율을 보였으며, 두 번째로 실시된 21대 국회의원 선거의 경우 그 비율은 약 2:1로 감소하였다. 이 결과는 사전투표제 도입으로 국내외 선행연구들에서 누적적으로 발견된 분산효과는 물론 몇몇 연구들에서만 나타났던 유인효과 역시 유발했다는 것을 잘 보여주고 있다.

이와 관련하여 많은 선행연구들이 유인효과를 발견하지 못한 이유를 이론과 방법론 두 측면에서 생각해 볼 수 있다. 먼저 이론적으로는 사전투표를 선택한 유권

자 집단의 이질성을 고려하지 못한 채 ‘사전투표자’를 단일한 집단으로 개념화했다는 점이다. 그러나 본 연구에서 시도한 ‘투표참여방식 전이확률’에서는 ‘사전투표자’를 기존 투표참여자와 기권자로 구분하였다. ‘사전투표자’가 ‘당일투표자’와 비슷한 특성을 가진다는 선행연구들의 결과는 분산효과가 유인효과를 압도할 정도로 강하다는 것을 보여줄 뿐, 유인효과 존재를 부정하는 결과로 해석될 수 없다. 또한 데이터의 규모라는 점에서 1,000명 내외의 소규모 설문표본 데이터로는 유권자 중 일부집단에서 나타나는 유인효과를 발견하기 어렵다. 본 연구결과에 따르면 유인효과는 제20대 국회의원 선거의 경우 전체 유권자의 약 2%, 제21대 국회의원 선거의 경우 약 5%라는 일부 인구집단에게서 나타났다. 설문조사 데이터는 제한된 연구비로 대표성이 확보된 연구표본을 효율적으로 얻을 수 있지만, 인구의 2~5% 정도에 불과한 인구집단에게서 나타나는 사전투표제 도입효과를 발견하기에는 표본 규모가 작다. 이런 점들을 고려할 때, 향후 연구에서는 투표방식선택이 아닌 선택의 변화를 반영할 수 있는 설문문항을 추가로 고려하며, 아울러 유인효과를 보일 것으로 예상되는 인구집단을 추가표집하는 것을 고려해볼 필요가 있다.

둘째, 사전투표제 도입이후 두 차례 국회의원 선거에서 선거구역 특성에 따라 분산효과와 유인효과 크기가 다르게 나타났다. 선거구역 규모가 작을수록, 20~30대 젊은 유권자 비율이 높을수록, 3인 이상 다인가구 비율이 높을수록 분산효과와 유인효과가 모두 크게 나타났으며(〈표 5〉 참조), 특히 20~30대 젊은 유권자 비율과 다인가구 비율이 높을수록 분산효과보다 유인효과가 더 두드러지게 크게 나타났다(〈표 6〉 참조). 또한 선거구역 특성변수가 투표방식선택 변화에 미치는 효과의 경우, 사전투표제 도입후 두 번째로 실시된 국회의원 선거보다는 첫 번째로 실시된 국회의원 선거에서 보다 뚜렷하게 나타났다.

그렇다면 유권자 연령분포, 다인가구 비율은 왜 유인효과 및 분산효과 크기와 연관성을 띠는 것일까? 먼저 ‘유권자 연령분포’와 유인효과 관계는 선행연구(김찬송 외 2016)에서 이미 소개한 바 있다. 유권자의 연령이 낮을수록 투표참여율이 낮은 것은 거의 대부분의 대의민주주의 국가에서 관측되는 현상이다. 젊은 유권자의 낮은 투표참여 원인들 중 하나는 이들 세대의 높은 ‘유동성(mobility)’이다. 20~30대 젊은 유권자들은 직업이나 학업의 이유로 타지에서 생활하고 있거

나 선거일에 출장일정이 잡혀 있을 가능성이 높다. 다시 말해 사전투표제는 이러한 20~30대 젊은 유권자에게 투표참여의 기회를 넓혀주었으며, 그 결과 저연령 유권자 비율이 높은 선거구역에서 유인효과를 나타내는 ‘기권→사전투표’ 전이확률이 높게 나타났다고 볼 수 있다.

또한 3인이상 다인가구 비율이 높은 선거구역일수록 분산효과와 함께 특히 유인효과가 크게 증가하는 것으로 나타났다. 가족단위로 투표소를 찾는 것은 매우 혼한 일임에도 초창기 선거연구를 제외한다면, ‘가구구성’이 투표참여 및 투표방식 선택에 미치는 효과에 대한 선행연구를 찾기는 쉽지 않다. 우리나라의 경우 선거일이 공휴일로 지정되어 있으며, 특히 국회의원 선거의 경우 계절적으로 여행인구가 많은 봄철에 선거가 실시된다. 실제로 여행에 대한 관심증가로 선거당일 국내나 해외로 여행을 떠나는 인구는 지속적으로 증가하는 추세이며, 무엇보다 최근 ‘워라벨’로 상징되듯 가족과 같이 보내는 시간이 강조되고 있다. 이런 점에서 본 연구결과는 가족단위 유권자가 어떻게 사전투표제를 활용하는지를 보여주는 흥미로운 결과다. 1인 가구의 증가와 같이 가족단위가 변하는 한국사회상황에서 유권자의 투표행동과 가구구성의 관계는 향후 연구에서 중요하게 고려할 필요가 있다.

‘투표선택방식 전이확률’과 ‘확장된 EI-W 모형’이라는 새로운 이론적·방법론적 접근에도 불구하고, 본 연구 역시도 다음과 같은 한계점들을 갖는다. 첫째, 집단수준에서 얻은 데이터를 이용하여 개인수준의 행동을 간접적으로 추정하는 EI 모형 역시 ‘생태학적 오류’에서 완전히 자유롭지 않다. 본 연구에서 채택한 EI-W 모형(Adolph et al. 2003; King 1997; King et al. 2004)이 생태학적 오류의 가능성을 대폭 감소시키는 것은 사실이지만, EI 모형의 가정 충족 여부에 따라 생태학적 오류의 가능성이 완전히 배제되지 않았을 수도 있다(예를 들어 Freedman et al. 1998).

둘째, 본 연구에서는 선행연구(Klima et al. 2019)와 마찬가지로 선거구역의 인구유동성이 투표참여방식 전이확률 추정에 영향을 미치지 않는다고 가정하였다. 그러나 엄밀히 말해 모든 선거구역의 시간에 따른 인구구성은 결코 동일하게 유지되지 않는다. 물론 4년이라는 상대적으로 짧은 국회의원 선거주기를 고려할 때 유권자 구성이 크게 달라질 가능성은 희박하며 유출입되는 인구구성이 매우 상이하게 바뀌지는 않을지 모른다. 그러나 본 연구가 시간에 따른 투표참여방식 변화

를 완벽하게 추론하지 못한 것 또한 사실이다.

셋째, 모집단 데이터를 사용했다는 장점에도 불구하고, 공개된 행정정보 데이터만을 사용하였다는 한계점을 갖는다. 즉 투표참여방식에 영향을 미치는 여러 정치심리학적 변수들, 이를테면 정당일체감이나 정치관심, 정치지식 등이 투표참여방식 전이확률에 미치는 효과를 살펴볼 수 없는 아쉬움이 있다. 향후 연구에서는 투표참여방식 전이확률 개념을 반영할 수 있는 설문조사 문항을 통해 정치심리학적 변수가 투표참여방식 변화에 어떤 역할을 하는지 살펴볼 필요가 있다.

몇 가지 한계점들에도 불구하고, 본 연구는 새로운 이론적 개념과 국내에서는 지금껏 시도되지 않은 TI 모형을 기반으로 사전투표제 도입효과를 살펴보았다. 이를 통해 선행연구에서 제기된 분산효과 및 유인효과와 규모를 추정하였으며, 아울러 선행연구에서 주목하지 못했던 투표참여방식에 가구구성이 어떻게 왜 중요한지도 새로 밝혀낼 수 있었다. 본 연구에서 시도한 투표참여방식 전이확률 개념과 TI 모형은 사전투표제 이후의 투표참여방식은 물론 시민들의 다른 정치적 행동을 설명하는 데도 유용하게 사용될 수 있길 기대한다.

참·고·문·헌

- 가상준. 2017. “2016년 국회의원선거에서 사전투표제 영향 분석.” 『동서연구』 29-2, 91-116.
- . 2018. “사전투표 유권자의 특징 변화.” 『한국정당학회보』 17-4, 99-120.
- 강신구. 2016. “사전투표제와 투표율: 제20대 국회의원선거 유권자 조사 자료분석.” 『한국정치연구』 25-3, 225-250.
- 김찬송·유재승·이현우. 2016. “사전투표제 세부분석: 20대 총선과 순수사전투표제.” 『21세기정치학회보』 26-4, 1-23.
- 손낙구. 2010. 『대한민국 정치 사회 지도: 수도권편』. 서울: 후마니타스.
- 윤기쁨·엄기홍. 2016. “사전투표제가 새로운 유권자의 투표참여를 동원하였는가?: 제6회 지방선거에 대한 경험적 분석.” 『미래정치연구』 6-2, 31-53.
- 이상신. 2014. “제6회 지방선거와 사전투표.” 『선거연구』 5, 53-82.
- Abramowitz, Alan I. 2010. *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy*. Yale University Press.
- Adolph, Christopher, Gary King, Michael C. Herron, and Kenneth W. Shotts. 2003. “A Consensus on Second-Stage Analyses in Ecological Inference Models.” *Political Analysis* 11-1: 86-94.
- Aldrich, John H. 1993. “Rational Choice and Turnout.” *American Journal of Political Science* 37-1: 246-278.
- Anselin, Luc, and Wendy K. Tam Cho. 2002. “Spatial Effects and Ecological Inference.” *Political Analysis* 10-3: 276-297.
- Ansolabehere, Stephen, and Eitan Hersh. 2012. “Validation: What Big Data Reveal about Survey Misreporting and the Real Electorate.” *Political Analysis* 20-4: 437-459.
- Barreto, Matt A., Matthew J. Streb, Mara Marks, and Fernando Guerra. 2006. “Do Absentee Voters Differ from Polling Place Voters?: New Evidence from California.” *Public Opinion Quarterly* 70-2: 224-234.
- Berinsky, Adam J. 2005. “The Perverse Consequences of Electoral Reform in the

- United States.” *American Politics Research* 33-4: 471-491.
- _____, Nancy Burns, and Michael W. Traugott. 2001. “Who Votes by Mail? A Dynamic Model of the Individual-Level Consequences of Voting-by-Mail Systems.” *Public Opinion Quarterly* 65-2: 178-197.
- Burden, Barry C., David T. Canon, Kenneth R. Mayer, and Donald P. Moynihan. 2014. “Election Laws, Mobilization, and Turnout: The Unanticipated Consequences of Election Reform.” *American Journal of Political Science* 58-1: 95-109.
- Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1980. *The American Voter*. University of Chicago Press.
- Cohen, Jacob. 2013. *Statistical Power Analysis for the Behavioral Sciences*. Academic press.
- Dahl, Robert A., and Tufte, Edward R. 1973. *Size and Democracy*. Stanford University Press.
- Duncan, Otis Dudley, and Beverly Davis. 1953. “An Alternative to Ecological Correlation.” *American Sociological Review* 18-6: 665-666.
- Frandsen, Gaardsted Annie. 2002. “Size and Electoral Participation in Local Elections.” *Environment and planning C: Government and Policy* 20-6: 853-869.
- Freedman, David A., Klein, Stefan P., Ostland, Michael, and Roberts, Michael R. (1998). “On ‘Solutions’ to the Ecological Inference Problem.” *Journal of the American Statistical Association* 93-444: 1518-1522.
- Garnett, Holly Ann. 2019. “Early Voting: Comparing Canada, Finland, Germany, and Switzerland.” *Election Law Journal: Rules, Politics, and Policy* 18-2: 116-131.
- Goodman, Leo A. 1953. “Ecological Regressions and Behavior of Individuals.” *American Sociological Review* 18-6: 663-664.
- Gronke, Paul, Eva Galanes-Rosenbaum, Peter A. Miller, and Daniel Toffey. 2008. “Convenience Voting.” *Annual Review of Political Science* 11: 437-455.
- _____, and Daniel Krantz Toffey. 2007. “Early Voting and Turnout.” *PS: Political Science and Politics* 40-4: 639-645.
- _____, and Daniel Krantz Toffey. 2008. “The Psychological and Institutional Determinants of Early Voting.” *Journal of Social Issues* 64-3: 503-524.

- Imai, Kosuke, and Kabir Khanna. 2016. "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records," *Political Analysis* 24-2: 263-272.
- Karp, Jeffrey A., and David Brockington. 2005. "Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries," *The Journal of Politics* 67-3: 825-840.
- King, Gary. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton University Press.
- _____, Ori Rosen, and Martin A. Tanner. 1999. "Binomial-Beta Hierarchical Models for Ecological Inference," *Sociological Methods and Research* 28-1: 61-90.
- _____, Ori Rosen, and Martin A. Tanner. 2004. *Ecological Inference: New Methodological Strategies*, Cambridge University Press.
- Klima, André, Thomas Schlesinger, Paul W. Thurner, and Helmut Küchenhoff. 2019. "Combining Aggregate Data and Exit Polls for the Estimation of Voter Transitions," *Sociological Methods and Research* 48-2: 296-325.
- Lazarsfeld, Paul F., Berelson, Bernard., and Gaudet, Hazel. 1944. *The People's Choice: How the Voter Makes up His Mind in a Presidential Campaign*, Columbia University Press.
- Oliver, J. Eric. 2000. "City Size and Civic Involvement in Metropolitan America," *American Political Science Review* 94-2: 361-373.
- Robinson, William S. 1950. "Ecological Correlations and the Behavior of Individuals," *American Sociological Review* 15-3: 337-341.
- Rosen, Ori, Wenxin Jiang, Gary King, and Martin A. Tanner. 2001. "Bayesian and Frequentist Inference for Ecological Inference: The $R \times C$ Case," *Statistica Neerlandica* 55-2: 134-156.
- Southwell, Priscilla L., and Justin I. Burchett. 2000. "The Effect of All-Mail Elections on Voter Turnout," *American Politics Quarterly* 28-1: 72-79.
- Stein, Robert M., and Greg Vonnahme. 2010. "Early, Absentee, and Mail-in Voting." J. E. Leighley, eds. In *The Oxford Handbook of American Elections and Political Behavior* (pp. 182-199). Oxford University Press.

Verba, Sidney., and Nie, Norma H. 1972. Participation in America. New York: Harper.

Wakefield, Jon. 2004. "Ecological inference for 2×2 tables". Journal of the Royal Statistical Society: Series A (Statistics in Society) 167: 385-425.

투고일: 2020년 8월 14일 심사일: 2020년 9월 8일 게재확정일: 2020년 9월 17일
--

Abstract

Early Voting and the Citizens' Change in Vote Choice in Korean Legislative Elections: An Ecological Inference (EI) Approach

In Seo Park | Yonsei University

Young Min Baek | Yonsei University

Does early voting reform actually increase voter turnout? Many studies have produced inconsistent findings and yet there is no consensus. This paper takes a new theoretical and methodological approach to examine the effect of early voting reforms. While most of the previous studies have modeled voting choices using individual-level survey data, we used precinct-level election results to estimate the 'probability of vote choice transition' based on ecological inference (EI, King et al. 2004). Using EI model, we focus on the changes in citizens' voting behaviors between three consecutive legislative elections from 2012 to 2020 in South Korea. Furthermore, by integrating the estimated probabilities data and precinct-level demographic characteristics data, we investigate the relationship between precinct characteristics and changes in citizens' vote choice across legislative elections. We found both the short- and long-term effects of early voting reforms on voting turnout reported in previous studies; additionally, highlighted on the role of 'family-size,' along with the size and age distribution of precincts. Theoretical and practical implications are also discussed.

KeyWords | Early voting, Vote choice, Ecological inference (EI), re-allocation of existing voters, additional recruit of nonvoters

