

사전/사후 비교 및 전통적인 이중차분법을 활용한 정책효과 추정에 대한 비판적 접근: 안전속도 5030 정책을 중심으로*

고길곤**

변영주***

홍민준****

〈目 次〉

I. 서론

II. 선행연구 및 이론적 검토

III. 연구설계

IV. 분석결과

V. 결론

〈요 약〉

본 논문은 근거기반정책의 핵심 요소인 정책효과 추정이 방법론적 선택에 따라 얼마나 상이한 결과로 도출될 수 있는지를 ‘안전속도 5030’ 정책 사례를 통해 분석하였다. 이 사례는 정책 시행 시점이 비교적 명확하고, 개입 전후의 교통사고 부상자 및 사망자 수 변화가 정량적으로 관측 가능하다는 점에서 정책효과를 계량적으로 검토하기에 적합하다. 정책 도입 초기에는 사전·사후 분석 및 이중차분 분석을 통해 정책 효과가 상당한 것으로 보고된 바 있으나, 본 연구는 서울시 전체 및 자치구를 분석 단위로 설정하여 다양한 방법론적 설정을 적용한 결과, 정책효과 추정치가 매우 제한적이거나 통계적으로 유의하지 않게 나타날 수 있음을 확인하였다. 이러한 결과는 근거기반정책에서 직관적으로 설득력 있어 보이는 수치에만 의존하여 정책 효과를 해석하는 접근이 분석상의 한계를 내포할 수 있음을 시사한다. 나아가, 자료의 특성과 정책 맥락에 부합하는 분석 방법을 신중히 선택함으로써 근거의 타당성과 해석의 신뢰성을 제고할 필요성을 보여준다. 아울러, 사전·사후 비교 및 이중차분법 역시 보편적으로 적용 가능한 정형화된 도구라기보다는, 분석 대상과 환경에 따라 유연하게 조정·보완되어야 할 분석 전략임을 강조한다.

【주제어: 근거기반정책, 사전/사후 비교, 이중차분법, 준실험설계, 안전속도 5030】

* 본 논문은 2025년 한국행정학회 하계공동학술대회 및 국제학술대회의 발표논문을 학술적으로 발전한 것입니다.

** 제1저자, 서울대학교 행정대학원 교수(kilkon@snu.ac.kr)

*** 교신저자, 서울대학교 행정대학원 박사과정(zoooo21@snu.ac.kr)

**** 공동저자, 서울대학교 글로벌행정발전연구소 객원연구원(hmj2815@snu.ac.kr)

논문접수일(2025.9.22), 수정일(2025.12.8), 게재확정일(2025.12.16)

I. 서론

한국 정부는 도시부 전체 도로의 제한속도를 하향 조정하는 ‘안전속도 5030’ 정책을 시행하였으며, 이는 교통사고 발생 시 사망자 감소와 보행자 안전 향상을 주요 목표로 한다(도로교통공단, 2021/04/19). 본 정책은 2017년 부산 영도구, 2019년 서울 4대문 지역에서 시범 운영 후 점차 확대되었으며, 2021년 4월부터 전국적으로 시행되었다(경찰청·행정안전부·국토교통부, 2021/04/19). 도로교통공단(2021/04/19)은 시범 운영 도시 13곳에서 교통사고 사망자가 102명에서 62명으로 약 39%, 중상자는 3,165명에서 2,697명으로 약 15% 감소하였다고 보고하였다(경찰청·국토교통부, 2021). 또한 전면 시행 후 적용 지역의 보행자 사망자가 전년도 동기간 대비 16.7% 감소하였다고 발표하였다. 한편, 시행 전후 평균 통행속도에는 큰 변화가 없다는 연구 결과도 제시되었지만(예: 김희경 외, 2019; 한상진·엄기중·송원방, 2023), 사전/사후 비교분석(예: 손승오 외, 2019; 김하음 외, 2021), 이중차분법(예: 김기만·배관표, 2024)을 활용한 연구에서는 안전속도 5030 정책이 교통사고 감소에 유의미한 효과가 있는 것으로 나타났다. 즉, 시범 운영을 통해 정책효과가 확인되었고, 전면 시행 이후에도 그 효과가 나타나고 있다는 실증적 근거가 제시되었다. 따라서, 안전속도 5030 정책은 근거기반의 합리적 정책결정의 사례라고 볼 수도 있을 것이다(김기만·배관표, 2024).

본 연구는 안전속도 5030 정책을 사례로 활용하여 ‘근거기반정책(Evidence-Based Policy)’에서 핵심이 되는 근거의 형성과 해석이 어떠한 분석 방법론을 적용하느냐에 따라 달라질 수 있음을 보여주는 것을 목적으로 한다. 이를 위해 먼저, 가장 보편적이고 단순한 사전/사후 비교와 인과효과를 측정하는데 유용하다고 알려진 이중차분법(Difference-in-Differences, DiD) 분석을 수행하였다. 이후 자기회귀오차 모형(Autoregressive Error Model, ARE), 단절시계열 모형(Interrupted Time Series, ITS), 통제집단합성법(Synthetic Control Method, SCM) 등 다양한 분석 방법을 적용하고 이들과의 결과를 비교함으로써, 동일한 자료를 사용하더라도 선택된 분석 모형에 따라 도출되는 정책효과의 크기와 방향이 달라질 수 있음을 실증적으로 제시하고자 한다.

가령 이중차분법은 준실험적 정책평가 기법 중 하나로 널리 활용되고 있지만, 전통적인 이중차분법은 지역 간 효과의 이질성, 특정 집단에 대한 차별적 영향, 정책의 장/단기적 효과, 추세 변화와 같은 측면까지 포착하는 데에는 한계가 있다. 다시 말해, 전통적인 이중차분법으로 평균적인 효과는 파악할 수 있으나, 정책효과의 이질성과 시간적 동태성을 충분히 설명하기는 어렵다. 그러나 손호성·이재훈(2018)이 국내 행정학 및 정책학 분야 주요 학

회보를 검토한 결과, 일부 연구에서는 매칭 기법을 병행하고 있으나, 상당수 연구는 여전히 처리집단과 통제집단의 정책 시행 전후 평균 차이를 비교하는 전형적인 이중차분법 추정 방식에 의존하고 있는 것으로 나타났다.

기존에 널리 활용되는 방법론들의 유용성과 중요성을 부정할 수는 없지만, 이러한 접근에 과도하게 의존할 경우, 정책효과 분석이 평균적 효과의 추정에 머무르게 되어, 다차원적으로 평가하는 것이 어렵다는 한계가 있다. 따라서 정책평가의 다목적성과 다차원성을 고려할 때, 연구자가 무엇에 주목하느냐에 따라 평가 목적과 분석 질문을 명확히 설정한 뒤, 이에 부합하는 방법론을 유연하게 적용하는 다방법론적 접근이 요구된다(고길곤, 2017). 그렇지 않다면, 복잡한 정책 문제를 지나치게 단순화하여 부적절한 정책 수단을 선택하는 오류에 빠질 수 있다(고길곤 외, 2020).

이러한 문제 인식하에, 본 연구는 안전속도 5030 정책을 사례로 다양한 분석 모형을 적용하여, 분석 단위 및 모형 선택에 따라 정책효과 추정 결과가 어떻게 달라질 수 있는지, 그리고 각 모형이 어떤 연구 질문에 답할 수 있는지를 살펴보고자 한다. 이를 통해 정책평가에서 분석 모형의 선택은 단순한 기술적 결정이 아니라, 정책효과에 대한 해석과 판단에 실질적인 영향을 미칠 수 있음을 강조하고자 한다. 나아가, 정책평가의 질적 수준 및 근거의 질적 수준을 제고하기 위해서는 정확한 방법론 선택이 필수적임을 제시한다. 결국, 동일한 자료를 사용하더라도 선택한 분석 설계에 따라 정책에 대한 판단 자체가 달라질 수 있으며, 이는 ‘근거’가 고정된 객관적 사실이 아니라 분석 설계와 방법론적 선택을 통해 구성되는 것임을 보여주는 데 본 논문의 의의가 있다.

II. 선행연구 및 이론적 검토

1. 안전속도 5030 정책

‘안전속도 5030’ 정책은 보행자 등 교통약자의 안전 확보와 교통사고의 심각도 완화를 목적으로, 도시부 도로의 제한속도를 전반적으로 하향 조정하는 국가 차원의交通安全 정책이다. 정부는 2016년 4월 「교통인프라 구축 기본계획」을 수립하면서 정책을 공식화하였으며, 같은 해 9월 경찰청, 국토교통부, 행정안전부, 한국교통안전공단 등 민·관·학이 참여하는 ‘5030 협의회’를 출범시켜 세부 기준과 추진 체계를 마련하였다. 이후 2017년 부산 영도구, 2019년 서울 4대문 지역을 포함한 여러 지역에서 시범 운영을 실시하였고, 그 결과와 해

외 사례·연구 결과 등을 반영하여 2019년 4월 17일 「도로교통법 시행규칙」 제19조를 개정하였다. 이에 따라 도시부의 주거·상업·공업지역 일반도로는 시속 50km 이하로, 주택가 이면도로·스쿨존 등 보행자 보호 우선 도로는 시속 30km 이하로 제한속도를 하향할 수 있는 법적 근거가 마련되었으며, 2020년 12월 20일부터 서울시 전역, 2021년 4월 17일부터 전국적으로 전면 확대되었다(경찰청·행정안전부·국토교통부, 2021).

시범 운영 및 전면 시행 이후, 안전속도 5030 정책이 적용된 지역에서는 교통사고 사망자 수가 감소한 것으로 나타났다. 경찰청(2021)에 따르면, 시범 운영 지역인 부산 영도구와 서울 4대문 안에서 각각 보행자 교통사고 사망자 수가 37.5%, 중상자 수가 30% 감소하여, 정책 도입 이후 일관된 사망 및 부상 감소 효과가 관측되었다. 특히, 2019년 11월부터 정책을 전면 시행한 부산의 경우, 2020년 보행자 교통사고 사망자 수가 전년 대비 33.8% 감소하였으며, 이는 해당 정책이 보행자 교통안전 확보에 일정 수준 기여를 했을 가능성을 시사한다. 또한 경찰청·국토교통부(2021)는 정책 전면 시행 이후 적용 지역 내 전체 교통사고 사망자 수가 전년 동기간 대비 12.6%(317명 → 277명), 보행자 사망자 수는 16.7%(167명 → 139명) 감소하였다고 보고하였다. 이는 비적용 지역과 비교할 때 각각 2.7배 및 4.5배 높은 감소율이다. 반면 제한속도 하향에도 불구하고 통행속도의 평균은 33.1km/h로, 전년 동기간 대비 약 1.0km/h 수준의 미미한 감소에 그친 것으로 나타났으며, 이는 차량 흐름에 큰 영향을 미치지 않았음을 시사한다. 한편, 전면 시행 이후 정책 조정의 필요성이 제기되면서, 2022년 대통령직인수위원회의 요청에 따라, 같은 해 말까지 전국 109개 구간(246.95km)의 제한속도를 60km/h로 상향 조정하는 등 유연한 적용이 이루어졌다(한국일보, 2022/03/27; 대전일보, 2023/04/04).

이처럼 안전속도 5030은 광범위하게 시행된 교통안전 정책이며 정책 시행 이후 의미 있는 변화가 관측되었음에도, 그 실효성과 정책효과에 대한 논의는 여전히 지속되고 있다. 특히, 해당 정책의 영향과 안전성 개선 효과를 실증적으로 분석한 기존 연구는 여전히 제한적이며, 분석 결과 또한 일관된 결론에 도달하지 못하고 있다. 대부분의 선행연구는 정책 시행 전후의 평균 통행속도를 비교하는 방식에 집중되어 있으며, 이들 연구에서는 속도 변화가 미미하거나(김희경 외, 2019; 한상진·엄기중·송원방, 2023), 시민들의 부정적인 인식이 확인되기도 하였다(김희경 외, 2019). 사고 발생에 대한 분석에서는 도시부 도로에서 제한속도 하향이 사고 감소에 긍정적인 영향을 미쳤고(손승오 외, 2019), 주행속도가 높을수록 사고 발생 확률이 유의미하게 증가한다는 분석도 제시되었다(김하음 외, 2021). 이중차분법을 활용한 김기만·배관표(2024)의 연구에서는 정책 시행 이후 사고 건수, 부상자 수, 사망자 수가 모두 유의미하게 감소한 것으로 나타났다. 이와 같은 결과들은 대체로 정책효과를 지지

하는 경향을 보이지만, 정책효과에 대한 신뢰할 수 있는 결론과 합의에 도달하였다고 보기는 어렵다. 다양한 평가 대상의 설정과 분석 방법의 활용이 선행되어야 하지만, 기존 연구들에서는 그 범위가 제한적으로 적용되었기 때문이다.

2. 근거기반정책에서의 근거 생산

근거기반정책(Evidence-Based Policy, EBP)은 정책의 결정과 집행 과정에서 과학적 연구, 체계적인 데이터 분석, 실증적 근거를 활용하여 정책의 효과성과 효율성을 제고하고자 하는 접근이다(Pawson, 2002). 이는 정책 결정이 직관이나 정치적 논리에 좌우되기보다는, 정책 개입의 사전·사후 효과를 체계적으로 평가하고 그 결과를 공개함으로써 정책의 신뢰성과 책임성을 높이는 데 기여한다(Angrist & Pischke, 2009). 특히 계량적 접근은 정책의 인과적 효과를 객관적이고 검증 가능한 방식으로 측정할 수 있게 함으로써, 정책의 정당성과 타당성을 뒷받침하는 핵심 논거로 작용한다(Head, 2010). 따라서 유사 정책의 효과성 평가 결과나 시범사업을 통해 도출된 실증적 성과는 정책 채택과 확산을 정당화하는 근거로 활용되며, 근거의 질은 정책 실행의 설득력과 정당성을 결정하는 중요한 요소이다(Davies & Nutley, 2000).

이러한 점에서 안전속도 5030 정책은 대표적인 근거기반정책의 사례로 평가된다. 이 정책은 국내외 교통사고 통계, 차량 속도와 사고 심각도 간의 상관관계, 보행자 사망률 등 다양한 실증적 데이터(예: 국토교통부·한국교통연구원, 2017)를 기반으로 설계 및 도입되었으며, 특히 시속 50km/h 이하로 속도를 제한할 경우 보행자 사망률이 현저히 감소한다는 연구 결과가 핵심적인 근거로 작용하였다. 즉, 정책 시행에 앞서 다년간의 시범 운영과 지역별 순차적 도입을 통해 실증적 근거가 축적되었고, 이를 바탕으로 정책이 확산되었다. 그러나 전면 시행 이후 일정 기간이 지나자 정책 조정의 필요성에 대한 논의가 본격화되었다. 구체적으로 2022년 대통령직인수위원회가 제한속도 정책의 탄력적 적용을 요청하면서(한국일보, 2022/03/27), 경찰청이 정책 매뉴얼을 개정한 결과, 같은 해 말에는 전국 18개 시도 109개 구간(246.95km)에서 제한속도를 60km/h로 상향 조정하는 등, 기본 체계를 유지하면서 유연한 적용이 가능하도록 정책을 조정하였다(대전일보, 2023/04/04).

그러나 이러한 정책 조정 과정에서 근거의 질에 대한 체계적인 검토는 이루어지지 않은 것으로 보인다. 미국의 근거기반정책결정위원회(commission on evidence-based policymaking)는 ‘근거’를 “통계분석을 통해 도출된 검증 가능한 수준의 정보”로 정의하고 있으며, 고길곤 외(2019) 역시 “통계적 활동을 통해 얻은 분석 가능한 정보”로 규정하였다. 그러나 통계적 분석을 통해 생성된 정책평가 결과나 실증적 성과가 인과관계를 의미하는 것

은 아니지만, 이러한 결과가 인과적 효과로 오인되는 경향이 존재한다. 예컨대, 안전속도 5030 정책 시행 이후 오히려 보행자 사망률이 증가하는 현상이 관찰될 경우, 단순히 “속도 제한 → 보행자 안전 증가”라는 선형적 인과구조를 전제한 해석은 적절하지 않을 수 있다. 인과관계를 규명하기 위해서는 특정 개입이 이루어졌을 때의 결과와 동시에 개입이 없었을 경우의 결과에 대한 정보가 모두 필요하기 때문이다. 그러나 동일한 단위에 대해 두 조건을 동시에 관찰하는 것은 현실적으로 불가능하기 때문에, 본질적인 한계가 존재한다. 이는 인과추론에서 발생하는 주요 문제로, 반사실적 조건(counterfactual condition)을 구성하고 비교하는 데 있어 방법론적 정교함이 요구되는 이유이기도 하다. 이러한 점에서 실증적 데이터에 기반한 정책 설계가 중요하지만, 항상 타당한 결정을 보장하는 것은 아니다. 결국, 근거기반정책의 실효성을 높이기 위해서는 단순히 실증적 정보의 존재 여부를 넘어서, 무엇을 근거로 삼을 것인지 그리고 그 근거가 얼마나 적절한지를 검토해야 한다, 그리고 다방법론적 접근을 통해 분석 결과를 병렬적으로 비교하고 검증하는 과정이 수반되어야 한다.

3. 정책효과 추정 방법

안전속도 5030 정책의 효과를 추정한 기존 연구들은 대체로 사전/사후 평균 비교 또는 이중차분법(Difference-in-Differences, DiD)과 같은 준실험적 접근을 중심으로 이루어져 왔다. 이러한 방법들은 분석의 간결성과 해석의 용이성 측면에서 장점이 있으나, 시간에 따른 구조적 변화, 집단 간 이질성, 외생적 요인의 개입 가능성 등 다양한 한계를 내포하고 있다. 따라서 본 절에서는 기존 분석 방법의 추정 방식과 전제를 비판적으로 검토하고, 보다 정교하고 유연한 정책효과 추정을 가능하게 하는 대안적 분석 방법들을 함께 탐색하고자 한다.

1) 사전/사후 비교 방법

단일집단의 시계열 자료를 활용하여 정책효과를 분석하는 가장 단순한 방법은 정책 개입 사전/사후 평균을 비교하는 t-검정(t-test)이다. t-검정을 적용하기에 앞서, 등분산성 여부를 검토하고, 그 결과에 따라 적절한 검정 방식을 선택해야 한다. 만일 등분산 가정이 충족되지 않았음에도 공통 분산(pooled variance)을 사용하여 검정통계를 계산하는 Student의 t-검정을 사용하는 경우, 실제보다 작게 계산된 표준오차로 인해 제1종 오류(Type I error)의 발생 확률이 증가할 수 있다. 따라서 이 경우에는 이분산성을 전제로 각 집단의 분산을 개별적으로 추정하고, 보다 보수적인 자유도를 적용하는 Welch의 t-검정이 적절하다.

t-검정을 활용한 단순 사전/사후 비교 방법은 정책 개입 전후의 평균 차이를 비교하는 데

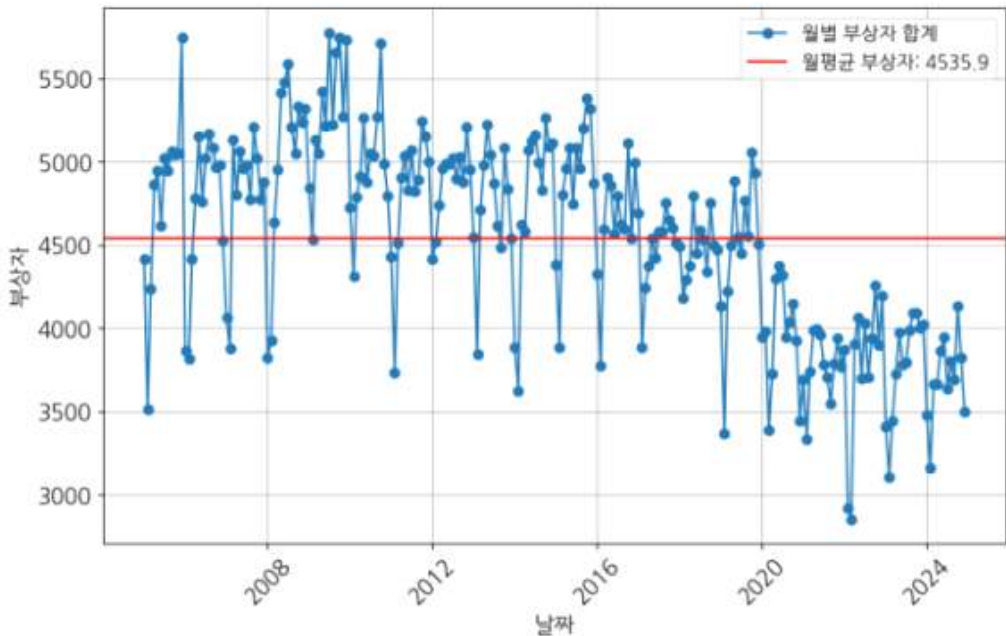
있어 가장 기본적이고 직관적인 방법이며, 실제 일부 선행연구에서는 정책 개입 전후 성과 변수의 파이를 정책효과로 간주하기도 한다(손승오 외, 2019; 김하음 외, 2021). 그러나 그 결과를 일반화하거나 인과적으로 해석하는 데에는 분명한 한계가 따른다. 특히, 분석 단위 간 구조적 차이나 초기 조건의 영향을 통제하지 못하기 때문에, 관측된 변화가 정책에 의한 것인지 외생적 요인에 의한 것인지 구분하기 어렵다. 이러한 한계를 보완하기 위해, 고정효과(fixed effects)를 통제한 사전/사후 분석이 대안으로 활용될 수 있다. 예컨대, 지자체 단위의 분석을 수행할 때는 인구 규모, 교통 인프라 수준 등 개별 지역에 고유하게 존재하는 불변의 특성을 통제함으로써, 정책 시행이라는 시간적 사건의 순효과를 보다 정밀하게 추정할 수 있다. 즉, 단순 전후 비교로는 제거되지 않는 지자체 간 이질성을 고정효과를 통해 통제함으로써, 보다 신뢰도 높은 정책효과 추정이 가능해진다.

다음으로 단순 사전/사후 비교 방식은 시계열 상의 추세를 고려하지 못하므로 정책효과를 과대평가하게 된다는 문제점이 존재한다. 예를 들어, <그림 1>에서 확인할 수 있듯이 서울시의 교통사고 부상자 수는 정책 도입 시기인 2021년 이전부터 전반적인 하향 추세를 보이고 있다. 이 경우, 정책 시행 이후 관측된 감소분이 실제로는 정책효과가 아니라 기존의 추세에 따른 자연 감소일 가능성이 크기 때문에, 단순한 평균 차이만으로 정책효과를 추정하는 것은 잘못된 해석을 초래할 수 있다. 또한 해당 시계열 자료는 월별 단위로 구성되어 있어 계절성(seasonality)과 오차항의 자기상관(autocorrelation)이 존재할 가능성이 크며, 특히 경기의 수축과 확장 등 순환적 요인(cyclical factors)이 장기적인 변동성을 유발할 수 있다. 따라서 추세가 있는 시계열 자료는 정상성(stationarity) 가정¹⁾을 충족하지 못하여, 시간의 흐름 그 자체가 종속변수에 영향을 미치는 허위상관 문제가 발생할 수 있다. 또한 시점에 따른 변화는 단순한 수준(level) 변화뿐만 아니라 추세(slope)의 변동을 수반할 수 있다. 이러한 변동은 상수가 아닌 시간에 따라 불확실하게 움직이는 확률적 과정(stochastic process)으로 간주할 수 있으며, 정태적 분석 틀로는 적절히 포착하기 어렵다.²⁾

1) 정상성 조건을 충족하지 않는 경우, 오차가 발산하는 문제가 발생한다.

2) 이러한 점을 고려하여, 이후의 단절시계열모형 분석은 자기회귀오차모형과 결합된 형태로 수행하였다.

〈그림 1〉 서울시 전체의 월별 교통사고 부상자 수 추세



자료: 서울 열린데이터 광장자료

이러한 한계를 보완하기 위하여, 시계열 자료의 구조적 특성을 반영하는 자기회귀오차 모형(Autoregressive Error Model, ARE), 단일집단 단절시계열 모형(Interrupted Time Series, ITS) 등을 활용할 수 있다.

먼저, 자기회귀오차 모형은 회귀식의 오차항이 시간에 따라 자기상관 구조를 가진다는 점을 반영함으로써, 시계열 자료의 통계적 특성을 보다 정밀하게 고려할 수 있는 분석 기법이다. 특히 시계열 자료에서 흔히 발생하는 오차항 간의 상관성을 모형 내에 내재화함으로써, 회귀계수의 표준오차가 과소추정 되거나 유의확률이 왜곡되는 문제를 완화할 수 있다. 또한 이 모형은 시간 추세나 계절성 등 정책과 무관한 외생적 변수를 설명변수로 포함하므로 정책 개입의 순효과를 식별하는 데에 효과적이다. 예를 들어, 자동차 안전기술의 발전이나 계절적 교통량 변화와 같은 시간 경과에 따른 외생 효과는 추세 변수를 통해 통제함으로써, 실제 정책효과와 시간 효과를 구분할 수 있도록 한다. 자기회귀오차 모형의 구조는 다음과 같다.

$$Y_t = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot Shift_t + \sum_s \gamma_s \cdot D_{st} + u_t$$

$$u_t = \epsilon_t - \phi_1 \cdot u_{t-1} - \phi_2 \cdot u_{t-2} - \dots - \phi_p \cdot u_{t-p}$$

단, Y_t : 시점 t 에서의 결과 변수 (예: 교통사 부상자 수), t : 시간,
 $Shift_t$: 개입 여부(개입 전 = 0, 개입 후 = 1), D_{st} : s 번째 지역 더미 변수,

u_t : 오차항(자기회귀 구조 포함),

ϕ_k : 오차항의 자기회귀 계수($k=1, \dots, p$), $\epsilon_t \sim i.i.d.$

한편, 단절시계열 모형은 단순한 평균 차이를 넘어 시간적 변화의 구조까지 포착한다는 점에서 유용하다. 구체적으로 개입 이전의 시간 추세를 기준선으로 설정하고, 정책 시행 이후의 수준(level) 변화와 추세(slope) 변화가 기존 경로에서 유의미하게 벗어났는지를 평가할 수 있다. 이때 수준 변화는 개입 시점에서의 개입으로 인한 결과변수의 차이, 즉, 즉각적인 변화를 의미하며, 추세 변화는 개입 이전 시점의 시계열 추세선 기울기 크기와 개입 이후의 추세선 기울기 간의 차이를 의미한다. 기본적인 단절시계열 모형의 구조는 다음과 같이 표현된다.

$$Y_t = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot Shift_t + \beta_3 \cdot (t - T_{Shift}) \cdot Shift_t + \epsilon_t$$

단, Y_t : 시점 t 에서의 결과 변수, T_{Shift} : 개입이 일어난 시점,

$Shift_t$: 개입 여부(개입 전 = 0, 개입 후 = 1), ϵ_t : 오차항

2) 이중차분법

정책의 효과를 받은 집단과 받지 않은 집단을 구분할 수 있는 경우, 이중차분법(difference-in-differences, DiD)은 효과적인 정책효과 추정 방법이 될 수 있다. 이중차분법은 관측되지 않은 집단 수준의 고정된 교란요인(confounding factor)을 제거하기 위한 준 실험적 식별 전략이다(Card & Krueger, 1993, Athey & Imbens, 2006). 이러한 반사실적(counterfactual) 추세의 모사는 무작위 배정이 어려운 현실의 정책평가에서 DiD를 유용한 대안으로 만든다. 일반적인 DiD 회귀모형은 다음과 같이 표현된다.

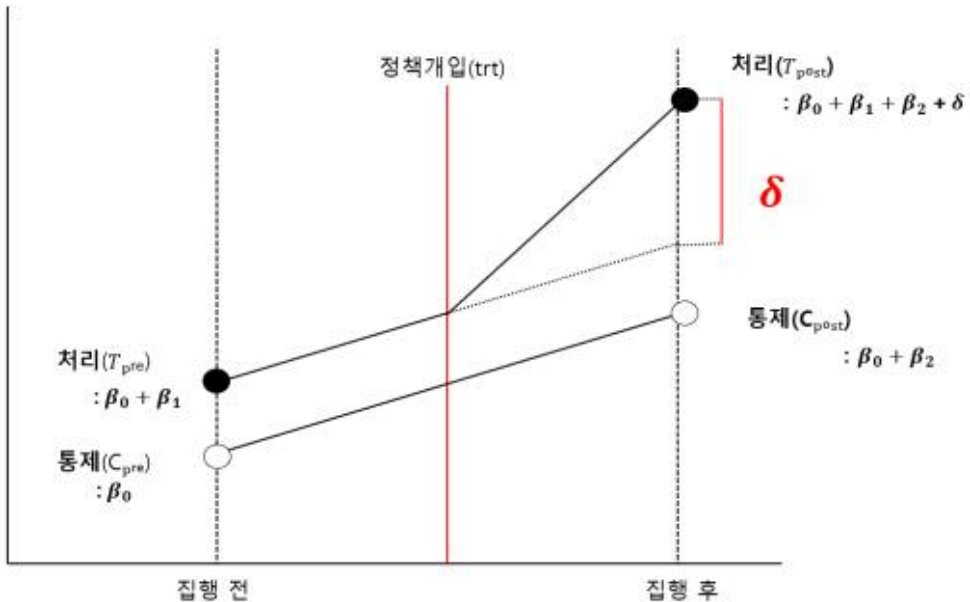
$$Y_i = \beta_0 + \beta_1 \cdot Trt_i + \beta_2 \cdot Shift_t + \delta \cdot (Shift_t \times Trt_i) + \epsilon_i$$

단, Y_i : 관측치 i 의 결과 변수, Trt_i : 처리집단여부(통제집단 = 0, 처리집단 = 1),

$Shift_t$: 개입 여부(개입 전 = 0, 개입 후 = 1), ϵ_t : 오차항

δ 는 ‘처리집단의 개입 전후 변화량($\delta + \beta_2$)’과 ‘통제집단의 개입 전후 변화량(β_2)’의 차이를 나타내며, 이는 정책효과 크기를 추정하는 데 사용된다. 즉, DiD는 개입 이전에 이미 존재하는 두 집단 간의 차이와 추세에 의한 변화를 함께 통제함으로써 보다 정확한 개입 효과를 추정하고자 하는 방법이다. 최근 등장하는 다양한 확장 모형 역시, 위와 같은 DiD의 기본 구조에 기초하고 있다(Goodman-Bacon, 2021).

〈그림 2〉 이중차분법 원리



출처: 저자 작성

주: 인과효과는 실제 관측된 사후 처리집단의 평균값과 반사실 결과의 차이($T_{post} - T_{counterfactual}$)를 의미하지만, DiD 추정치는 $T_{post} - [T_{pre} + (C_{post} - C_{pre})]$ 로 계산됨

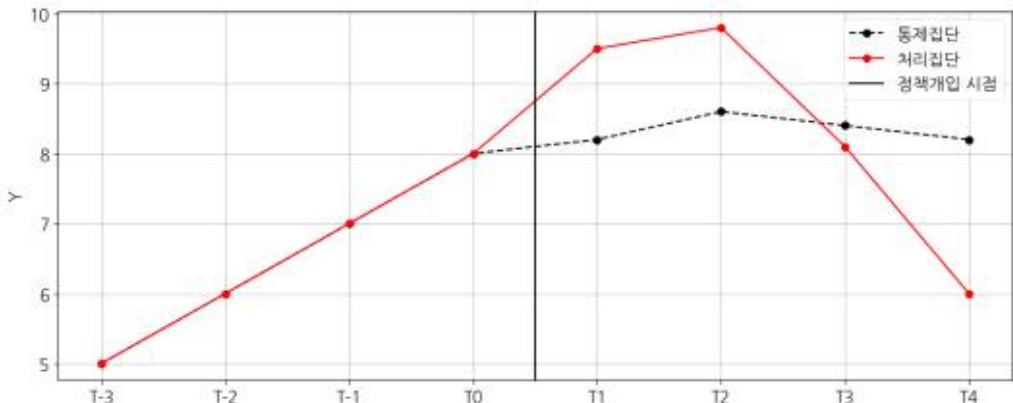
그러나 DiD 접근에는 본질적인 제약이 따른다. 우선 핵심적인 전제 중 하나인 평행 추세 가정(parallel trends assumption)은 정책 개입 이후(post-treatment)의 반사실 경로가 실증적으로 관측될 수 없다는 점에서 본질적으로 검증 불가능한 가정이다. 따라서 사전

(pre-treatment) 시점에서 처리집단과 통제집단 간의 추세가 유사하다고 해서, 정책 개입이 없었을 경우에도 동일한 추세가 유지될 것이라고 단정할 수 없다(Roth et al., 2023). 예를 들어, 사춘기 이전까지는 소년과 소녀의 평균 키가 유사한 추세를 보이지만, 약 13세 이후에는 급격한 차이가 나타났다고 해서 사춘기가 소년들의 키 변화에 인과적 영향을 미쳤다고 해석할 수 없다(Kahn-Lang & Lang, 2020). 이처럼, 사전 추세의 유사성이 평행 추세 가정을 완전히 입증하는 것이 아니며, 정책 이후 시점에 외생적인 혼란 요인이 존재하지 않는다는 추가적인 가정 또한 함께 요구된다.

또한 DiD는 분석에 활용되는 데이터 구조에 따라 다른 형태의 제약을 수반한다. 결합 횡단면 자료(independently pooled cross-section data)를 사용하는 경우, 각 시점마다 서로 다른 개체가 표본으로 추출되므로 개체별 고정효과(fixed effects)를 통제할 수 없고, 이에 따라 관측되지 않는 이질성이 정책효과 추정에 편향을 야기할 수 있다. 반면 동일 개체를 반복 측정하는 구조라 하더라도, 시점 간 오차항이 상관될 가능성이 높기 때문에 시계열 상의 공분산 구조를 적절히 반영하지 않으면 표준오차가 과소추정되어 유의성 판단에 오류가 발생할 수 있다. 이처럼 자료 구조에 따른 통계적 전제 위반은 추정 결과의 신뢰도에 중대한 영향을 미칠 수 있다.

뿐만 아니라 단순한 DiD는 시점 간 평균값의 수준 차이를 기반으로 효과를 추정하기 때문에, 기울기(추세)의 변화, 비선형적 효과, 또는 분산 변화 등 시간 구조의 복잡성을 포착하지 못한다. 예컨대 <그림 3>과 같이 정책 시행 전에는 처리집단과 통제집단이 동일한 추세를 보였더라도, 이후 처리집단의 값이 일시적으로 상승한 뒤 급격히 감소하는 등 비선형적이고 시차가 있는 반응이 나타날 경우, DiD는 이러한 동적 효과를 왜곡하거나 간과할 수 있다.

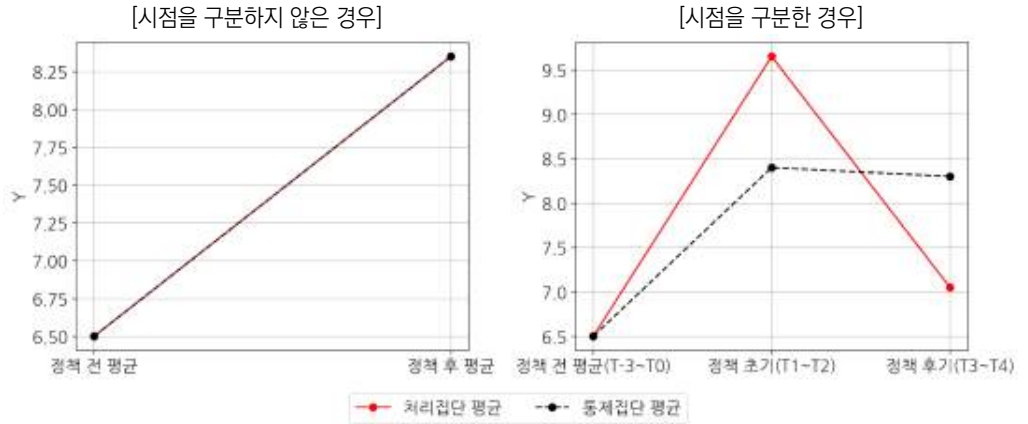
<그림 3> 시점별 이질적 처치효과 사례



출처: 저자 작성

실제로 <그림 3>의 예시에서 집단별 사전, 사후 평균값을 구하면, <그림 4>의 좌측과 같이 집단 간 차이가 나타나지 않는다. 2×2 DiD 분석이 정책 시행 전후의 집단별 평균값만을 비교하기 때문에, 정책효과가 시점에 따라 달라지는 양상을 반영하지 못하기 때문이다. 이 경우, 정책효과가 과소추정되거나 통계적으로 유의하지 않다는 결론에 이를 수 있다. 반면, 우측과 같이 정책 시행 직후와 이후 구간을 시계열적으로 구분하여 분석하면, 정책효과가 특정 시점에 집중되거나 점진적으로 누적되는 등 동적 변화 양상을 보다 정밀하게 관찰할 수 있다. 나아가 DiD는 정책 도입 이후 집단 내 분산이 확대되거나 하위 집단별로 상이한 효과가 나타나는 경우를 포착할 수 없다. 예를 들어, 동일한 처리집단 내에서도 성별, 연령, 소득 수준에 따라 정책효과가 다르게 나타날 수 있으나, DiD는 이를 이질적 처리효과(heterogeneous treatment effects)로 분리하여 식별하지 못한다. 이러한 맥락에서, 정책효과를 단일한 평균 효과 추정치로 측정하는 접근은 집단 내부의 구조적 변화를 간과할 위험이 있다.

〈그림 4〉 정책 개입 이후 시점 구분 여부에 따른 결과 차이 사례



출처: 저자 작성

이상의 한계로 인하여, 최근에는 전통적인 이중차분법에서 벗어나, 두 시점 이상을 포함하거나 집단별로 상이한 정책 도입 시기를 고려하는 등의 다양한 확장 모형이 제안되고 있다(Callaway & Sant'Anna, 2021; Roth et al., 2023). 이와 더불어, 통제집단을 활용한 단절 시계열 모형(Controlled interrupted time series, CITS) 및 통제집단합성법(Synthetic Control Method, SCM)과 같은 대안적 인과효과 분석 방법론 역시 점차 폭넓게 활용되고 있다(Abadie et al., 2010; Lopez Bernal et al., 2018).

대안적 방법론 중 통제집단을 활용한 단절시계열 모형은 단일집단 시계열 분석의 한계를 보완하여 외생적 충격이나 시간적 공변량의 영향을 보다 정확히 통제할 수 있는 방법이다. 수준 변화와 추세 변화를 함께 분석할 수 있는 단일집단 단절시계열 모형에서 더 나아가, 처리집단과 통제집단의 단절시계열을 비교하여 외생요인을 제거하고 보다 명확하게 정책효과를 식별할 수 있다. 또한 기존의 DiD가 단지 두 시점(정책 전·후)의 집단 평균 차이만을 비교하는 방식이라면, 통제집단을 활용한 단절시계열 모형은 장기적인 시간 흐름에 따른 결과변수의 진화 경로를 고려함으로써 정밀한 정책효과 분석이 가능하다. 따라서 정책효과가 일정 시점에 갑작스럽게 발생하지 않고 점진적·누적으로 나타나는 상황에서는, 단순 평균 차이에 집중하는 DiD보다 더 현실적인 분석 도구로 기능한다. 통제집단을 활용한 단절시계열 모형식은 다음과 같이 표현된다.³⁾

$$Y_t = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot Trt_t + \beta_3 \cdot (Trt_t \times t) + \beta_4 \cdot Shift_t \\ + \beta_5 \cdot (t - T_i) \cdot Shift_t + \beta_6 \cdot (Shift_t \times Trt_i) + \beta_7 \cdot (Shift_t \times t \times Trt_i) + \epsilon_t$$

단, Y_i : 관측치 i 의 결과 변수, Trt_i : 처리집단여부(통제집단 = 0, 처리집단 = 1),
 $Shift_t$: 개입 여부(개입 전 = 0, 개입 후 = 1), ϵ_t : 오차항

통제집단합성법은 단일 처리 단위에 대한 정책효과를 분석할 때, 대조 후보군의 사전시점 결과변수의 궤적을 가중 결합한 ‘합성 통제집단’을 통해 반사실 경로를 추정하는 방법론이다. 이때 합성 통제집단은 복수의 대조 후보군들을 적절한 가중치로 결합하여, 정책 개입 이전(pre-treatment) 시점에서 처리 단위와 유사한 경향을 갖도록 구성된다. 통제집단합성법은 이처럼 사전 기간의 경로 유사성에 기반해 비교 가능한 통제집단을 만들고, 정책 시행 이후 시계열 상의 변화 차이를 분석함으로써 개입 효과를 추정한다. 따라서 이 방법은 평행추세 가정이 충족되지 않거나, 정책 처리가 단일 혹은 소수의 단위에만 적용된 경우, 통제집단 간 이질성이 큰 경우, 또는 정책효과가 선형적이지 않거나 시간이 지남에 따라 누적되는 경우 등에서 전통적인 DiD 접근의 현실적 한계를 보완하는 대안이 될 수 있다. 통제집단합성법은 다음과 같은 수식으로 표현된다.

3) Pinto(2022)의 수식을 참고하여 작성하였다.

$$\hat{\alpha}_{1t} = Y_{1t}^{treated} - \sum W_j \cdot Y_{jt}^{control}$$

단, $\hat{\alpha}_{1t}$: 시점 t 에서의 처리 단위의 정책효과 추정값,

$Y_{1t}^{treated}$: 처리 단위의 실제 관측값,

$\sum W_j \cdot Y_{jt}^{control}$: 합성 통제집단의 예측값,

W_j : 대조후보군 단위 j 에 부여된 가중치(subject to $W_j \geq 0$, $\sum W_j = 1$)

이상의 논의를 종합하면 DiD 분석은 처치 여부에 따른 평균 차이를 기반으로 정책효과를 추정하는 유용한 수단이지만, 그 자체로 인과관계를 입증하는 논리적 근거로 기능하기에는 충분하지 않다. 특히 ‘인과관계’ 또는 ‘인과효과’라는 표현은, 자칫 변수들 사이에 고정적이고 보편적인 인과 메커니즘이 존재한다는 인상을 줄 수 있어 오해를 유발할 소지가 있다. 이는 DiD 분석 결과를 곧바로 정책효과로 간주하는 해석상의 오류로 이어진다. 또한, 정책은 단일한 개입을 통해 복수의 경로와 시차를 거쳐 다양한 결과 변수에 영향을 미치는 복합적인 메커니즘에 기반하기 때문에, 단일한 인과효과 추정값만으로 정책의 전체 효과를 단정 짓는 것은 타당하지 않다. 따라서 분석에 앞서 정책이 초래한 변화가 '어떤 종류의 차이'인지(예컨대 수준의 변화인지, 추세의 변화인지, 혹은 분산 구조의 변화인지)를 명확히 설정하고, 이에 적합한 식별 전략과 분석 방법론을 선택하는 것이 정책평가 연구의 타당성을 높이는 데 필수적이다.

III. 연구설계

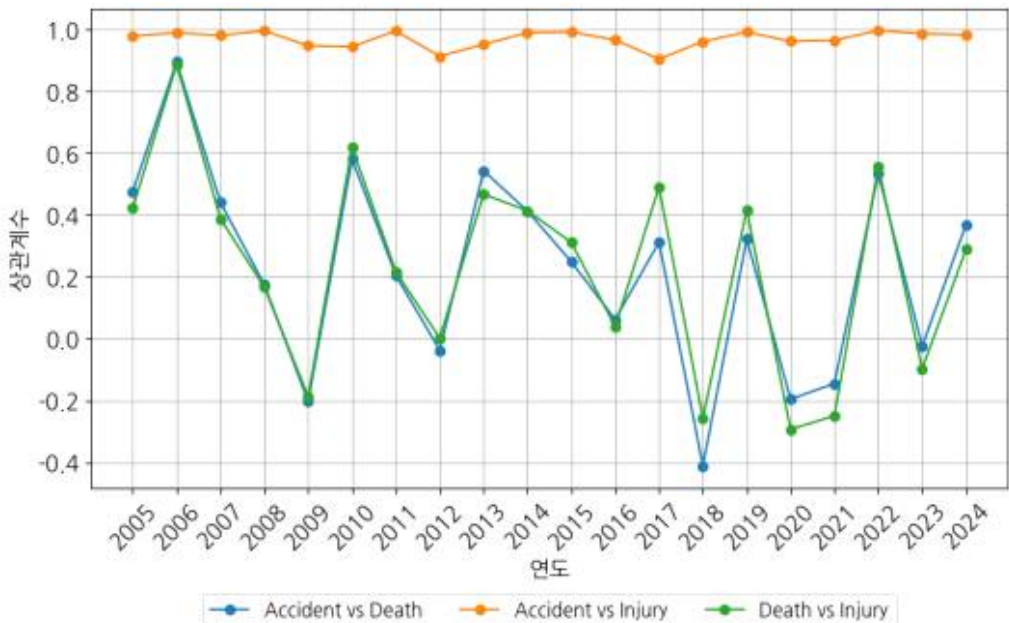
1. 분석자료 및 목표변수 설정

본 논문은 서울특별시가 ‘서울 열린데이터 광장’을 통해 제공하는 25개 자치구별 교통사고 부상자 수에 대한 2005년부터 2024년까지의 월별 자료를 활용하였다. 분석의 시작 연도를 2005년으로 설정한 이유는, 그 이전까지의 교통사고 자료가 사고 발생 지역 단위가 아닌 지방경찰청 단위로 수집되어 비교가능성이 떨어지기 때문이다. 또한 본 연구는 분석 대상 지역을 서울특별시로 한정하였다. 물론 서울시 4대문 내 지역의 시범사업 효과를 분석할 때, 안전속도 5030 정책이 도입되지 않은 비수도권 지자체를 통제집단으로 선정할 수도 있다. 그러나, 비수도권 지역 지자체의 경우 급속한 고령화와 인구 감소 등으로 인해 교통환경

의 변화 양상이 수도권과 현저히 다르다. 따라서 이러한 이질적 요인을 분석 모형에서 충분히 통제하지 못할 경우, 정책효과 추정의 신뢰성과 타당성이 저하될 수 있다고 판단하여 서울시를 분석 대상으로 한정하였다.

다음으로, 정책평가에서 어떤 지표를 정책의 목표변수로 설정할 것인지는 매우 중요한 문제이며 실제로 정책의 성과를 무엇으로 정의할지는 논쟁의 대상이 된다. 특히 본 연구에서 다루는 안전속도 5030과 같은 속도 제한 정책의 경우, 사고, 사망, 부상 중 어느 변수에 상대적으로 초점을 맞추고 있는지가 명확하지 않다. 이에 정부 역시 해당 정책의 효과를 판단할 때 사망과 부상 감소 효과를 함께 고려하였다. 속도 제한이 사망 사고를 줄이는 데 효과적이라는 주장도 가능하지만, 전체 사고 건수나 부상자 수를 줄이는 데 더 효과가 있다는 주장 또한 가능하다. <그림 5>는 사고 건수, 부상자 수, 사망자 수 간의 연도별 상관관계를 보여준다. 사고 건수와 부상자 수는 모든 연도에서 0.903~0.995의 매우 높은 상관을 보인 반면, 사고 건수와 사망자 수, 부상자 수와 사망자 수 간의 상관은 일관되게 높지 않았다. 이러한 특성을 고려하여 본 연구는 상대적으로 발생 빈도가 높고 우연적 변동 가능성이 낮은 부상자 수를 기본 정책목표 변수로 사용하였다.

<그림 5> 연도별 사고, 사망, 부상 간 상관관계 변화



자료: 서울 열린데이터 광장

2. 분석방법

본 연구는 안전속도 5030 정책이 2019년 서울 4대문 안 지역(중구 및 종로구)에서 우선적으로 시행되었고, 이후 2020년 12월 서울시 전역에 일괄 적용되었다는 점을 고려하여(서울특별시 도시교통실, 2022), 크게 두 단계에 걸친 분석을 수행하였다.⁴⁾ 첫 번째 단계에서는 정책이 서울 전체에 적용된 이후의 영향을 파악하기 위하여, 단일집단을 대상으로 한 분석을 진행하였다. 우선, 처리집단의 정책 시행 전후를 비교하는 사전/사후 차이 분석을 수행하였다. 단일집단의 시계열 자료를 분석하는 가장 단순한 방법은 사전/사후 평균비교를 하는 t-검정(t-test)이지만, 이는 지자체 간 구조적 차이나 초기 조건의 영향을 통제하지 못하기 때문에 정책효과에 대한 일반화나 인과적 해석에 한계가 존재한다. 이를 보완하기 위해 지자체의 고정효과(fixed effects)를 통제한 사전/사후 분석을 수행하여 지자체마다 고유하게 존재하는 불변의 특성들(인구, 교통 인프라 등)을 통제하였다. 또한, 정태적 분석 틀로는 적절히 포착하기 어려운 시계열 자료의 구조적 특성을 고려하기 위해 자기회귀오차 모형(Autoregressive Error Model), 단일집단 단절시계열 모형(Interrupted Time Series, ITS)을 추가적으로 활용하여 방법론별 정책효과 추정 결과를 비교하였다. 이때 자기회귀오차 모형에서는 정책 개입 여부에 따라 분산이 달라지는 조건부 이분산 구조를 모형화한 경우와 그렇지 않은 경우의 결과를 함께 제시하였다. 또한 단절시계열 모형은 시계열 자료의 특성을 고려하여 자기회귀오차를 포함한 형태로 구성하였다.

다음으로 서울시는 안전속도 5030 정책을 서울 전역에 시행하기 이전, 2018년 6월 서울의 중심 도로인 종로(세종로사거리~홍인지문교차로 구간)의 통행속도를 시범적으로 하향 조정하였으며(서울특별시 보행친화기획관, 2018), 2019년 4월에는 4대문 안 전체로 정책을 확대 적용하였다(치안정책연구소 외, 2022). 이러한 단계적 시행 시점을 활용하여, 본 연구는 2020년 12월 전면 시행 이전 정책 대상 지역이었던 중구와 종로구를 처리집단으로, 그 외 지역인 서울 내 다른 자치구를 통제집단으로 설정하여 두 번째 단계의 분석을 수행하였다. 먼저 처리집단과 통제집단의 초기 차이와 추세에 따라 발생한 효과를 함께 통제하여 정책효과를 추정하는 이중차분법을 활용하였다. 이중차분법은 정책 시행 전후의 변화가 정책의 영향을 받지 않은 통제집단에서도 동일하게 나타날 것이라는 ‘평행추세 가정’에 기반하여, 처리집단과 통제집단 간 변화의 차이를 비교하는 분석 방법이다. 그러나 전통적인 DiD

4) ‘안전속도 5030’ 정책은 2021년 4월 17일 전국 확대 시행 이전에, 2020년 12월 21일부터 서울시 전역에 선제적으로 적용되었다. 그러나 일부 정부 기관의 보도자료에는 전국 확대 일만 기재되어 있어, 시범사업 종로 및 본격적 개입 시점을 2021년 4월로만 오인하는 경우가 존재한다. 이로 인해 정책효과 분석 시 개입 시점을 잘못 설정할 가능성이 있으며, 분석 결과 또한 달라질 수 있다.

방식은 시점 간 평균값의 수준 차이를 기반으로 효과를 추정하기 때문에, 기울기(추세)의 변화, 비선형적 효과, 또는 분산 변화 등 시간 구조의 복잡성을 포착하지 못한다. 본 연구는 이를 보완하기 위해 통제집단을 활용한 단절시계열 모형(Controlled interrupted time series, CITS)과 통제집단합성법(Synthetic Control Method, SCM)을 사용하여 추가적인 분석을 수행한 후, DiD 방법을 적용하였을 때 도출된 정책효과와의 차이를 비교하였다.

한편, 정책 도입 직후에는 일시적으로 뚜렷한 효과가 나타나더라도, 시간이 지남에 따라 그 효과가 점차 약화되어 장기적으로는 유의미한 변화가 사라진다면, 정책의 지속가능성 측면에서 중요한 한계로 작용할 수 있다. 본 연구는 전이함수(transfer function) 모형을 활용하여 정책 개입 시점 이후 효과의 변화 양상을 추가적으로 살펴보았다.

〈표 1〉 분석 모형 및 방법의 요약

순번	분석 방법	분석 단위	내용
1단계 : 서울시 대상 전면 시행 효과 분석 (처리기간: 2021년 1월 - 2024년 12월)			
1	단순 사전/사후 비교	서울시 / 자치구	평균 차이에 대한 t-검정
2	고정효과 사전/사후 비교	자치구	지자체 고정효과 통제 후 평균 차이에 대한 t-검정
3	자기회귀오차 모형	서울시 / 자치구	시점 간 오차의 자기상관, 지자체별 고유 특성, 시간 추세효과 통제
4	단일집단 단절시계열 모형	서울시 / 자치구	개입 전후의 시간 흐름에 따른 결과변수의 변화 분석
2단계 : 서울 4대문 지역 내 시범 사업 효과 분석 (처리기간: 2019년 4월 - 2020년 11월)			
5	이중차분법	(처리)중구, 종로구 (통제)그 외	처리집단과 통제집단의 초기 차이와 추세에 따라 발생한 효과를 함께 통제
6	통제집단이 있는 단절시계열 모형	서울시 내 자치구	개입 전후 시점의 추세 변화를 고려하여 처리집단과 통제집단 간 차이 비교
7	통제집단합성법	(처리)중구, 종로구 (대조후보군)그 외 서울시 내 자치구	가상의 통제집단 구성 후, 집단 간 차이 비교
추가 분석			
8	전이함수 모형	서울시 / 자치구	정책 개입이 대상 지역에 미치는 영향의 시간 지연(lag) 및 누적효과를 모형에 포함하여 정책효과와 변화 양상 분석

IV. 분석결과

1. 전면 시행 효과 분석

부상자 수로 측정한 정책효과는 서울시 전체와 개별 자치구 수준에서 구분할 수 있다. 우선 분석 단위를 서울시 전체로 설정하여 수행한 정책 전·후 비교분석 결과는 <표 2>와 같다. 단순 사전/사후 평균 비교를 위한 t-검정 결과, 정책 전면 시행 이후 서울시의 월평균 교통사고 부상자 수는 약 974.8명 감소하였으며, 이 차이는 통계적으로 유의미한 수준($p < .001$)으로 나타났다. 분석 기간 중 월평균 부상자 수가 약 4,535.9명이었던 점을 감안할 때, 이는 정책 시행 이후 부상자 수가 상당히 감소하였음을 시사한다. 단순 평균 비교 방식은 분석결과를 직관적으로 해석하는 데에는 용이하나, 서울시 교통사고는 시간에 따른 추세가 존재하기 때문에, 외생적 요인의 영향을 충분히 통제할 수 없다는 한계가 존재한다. 즉, 관측된 변화는 정책의 효과가 아니라 동일한 시기에 발생한 다른 요인의 영향이나 시계열의 추세와 같은 구조적 변화 자체에 기인한 결과일 수 있다.

사전/사후 비교 방식과 달리, 자기상관(autocorrelation) 및 추세를 모형에 반영한 결과 정책효과가 통계적으로 유의미하지 않은 것으로 나타났다. 먼저 자기상관을 반영하기 위하여 자기회귀오차 모형(ARE)을 적용한 분석 결과, 조건부 이분산 구조를 고려한 경우와 고려하지 않은 경우 모두에서 정책 전면 시행 이후 서울시의 월평균 교통사고 부상자 수 변화에 대한 추정치는 약 -139.04명으로 나타났다. 이 추정치는 사전/사후 비교 방식으로 도출된 정책효과 크기와 비교할 때 현저히 작은 수준일 뿐만 아니라, 통계적으로도 유의하지 않았다. 단일집단 단절시계열(ITS) 분석 결과에서도 유사하게 정책 시행 이후 월평균 약 -137.61명 수준의 변화가 있었던 것으로 추정되었으나, 마찬가지로 통계적으로 유의하지 않았다. 또한, 시간 추세변수의 회귀계수는 자기회귀오차 모형에서 -4.300($p < .01$), 단일집단 단절시계열 분석에서 -4.237($p < .05$)로 나타났으며, 이는 부상자 수가 시간의 흐름에 따라 점진적으로 감소하는 경향을 보였음을 의미한다. 이러한 결과는 단순 사전/사후 비교 방식의 추정치에 정책효과뿐만 아니라 시간 경과에 따른 자연 감소 효과가 함께 포함되어 있었을 가능성을 의미한다. 한편, 단일집단 단절시계열 분석은 정책 시행 이후 추세의 변화까지 살펴볼 수 있다는 점에서 유용하지만, 본 연구의 분석에서는 서울시 단위에서는 정책에 의한 추세 변화가 통계적으로 유의미하지 않은 것으로 나타났다.

〈표 2〉 전면 시행 부상자 수 감소 효과 분석결과: 서울시

분석방법	단순 사전/ 사후비교	자기회귀 오차모형 (최대시차=12, 이분산성 미고려)	자기회귀 오차모형 (최대시차=12, 이분산성 고려)	단일집단 단절시계열 분석 (최대시차=12)
정책효과	-974.800*** (-17.25)	-139.044 (-1.00)	-139.044 (-1.42)	-137.605 (-1.47)
시간 추세		-4.300** (-2.72)	-4.300** (-2.72)	-4.237* (-2.32)
추세 변화				-0.493 (-0.10)
Intercept		4981.00*** (21.30)	4981.00*** (20.99)	4980.00*** (20.88)
Folded F				
F Value	2.45**			
R-Square		0.766	0.766	0.766
Durbin-Watson		1.828	1.828	1.828

*p<.05, **p<.01, ***p<.001

주: 괄호 안의 값은 t-value를 의미함

다음으로 〈표 3〉은 분석 단위를 자치구로 설정하여 분석한 결과를 보여준다. 이는 정책 시행이라는 개입이 서울시 각 자치구에 전반적으로 어떠한 영향을 미쳤는지를 평가하기 위함이다. 먼저 자치구별 단순 사전/사후 평균 비교를 살펴보면, 정책 시행 이후 월평균 교통 사고 부상자 수는 약 38.99명 감소한 것으로 나타났다($p<.001$). 그러나 이 경우 자치구별 이질적 특성(heterogeneity)을 통제하지 못했다는 한계가 존재하므로, 지자체의 고정효과를 통제한 상태에서 사전/사후 평균을 비교하는 분석을 추가로 수행하였다. 분석 결과, 정책 시행 이후의 평균 차이는 -38.99명으로, 자치구 고정효과를 통제하지 않은 경우와 동일한 값을 보였다. 그러나 t-value는 -15.23에서 -37.45로 크게 증가하였다. 이는 고정효과를 통제함으로써 자치구 간 구조적 차이를 제거하고, 정책효과를 보다 정밀하고 일관되게 추정할 수 있었음을 의미한다.

분석 단위를 자치구로 설정하고 자기회귀오차 모형을 활용한 결과, 정책 시행 이후 자치구별 월평균 부상자 수는 약 6.32명 감소한 것으로 나타났다($p<.01$). 또한 단일집단 단절시계열 분석에서는 약 6.37명 감소로 나타났으며($p<.01$), 시계열적 요소를 고려하지 않은 단순 사전/사후 평균 비교 결과(-38.99명)와는 상당한 차이를 보였다. 한편, 시간의 흐름에 따른 구조적 변화를 통제하기 위해 포함된 시간 추세변수도 두 모형 모두에서 통계적으로 유의한 음(-)의 계수를 나타냈다. 이는 정책과 무관하게 시간 경과에 따라 교통사고 부상자 수

가 감소하는 추세가 존재하며, 단순 사전·사후 평균 비교에서 나타난 정책효과가 시간적 자연감소분을 함께 포함했을 가능성이 있음을 의미한다. 즉, 분석 단위를 자치구로 설정한 경우에서도 서울시 전체를 대상으로 분석하였을 때와 마찬가지로 시계열 구조를 고려한 분석의 정책효과 추정치가 단순 평균 비교보다 작게 나타났다. 이는 비통제 분석에서 과대추정된 효과가 포함되었음을 시사한다.⁵⁾ 참고로 6.293명의 감소는 정책 도입 이전의 자치구 평균보다 3.3% 감소한 수준에 불과하다.

〈표 3〉 전면 시행 부상자 수 감소 효과 분석결과: 자치구

분석방법	단순 사전/ 사후비교	고정효과 사전/사후 비교	자기회귀 오차모형 (최대시차=12, 이분산성 미고려)	자기회귀 오차모형 (최대시차=12, 이분산성 고려)	단일집단 단절시계열 분석 (최대시차=12)
정책효과	-38.991*** (-15.23)	-38.991*** (-37.45)	-6.321** (-2.21)	-6.321** (-2.97)	-6.369** (-2.99)
시간 추세			-0.158*** (-9.10)	-0.158*** (-7.80)	-0.159*** (-5.45)
추세 변화					0.009 (0.08)
Intercept			206.860 (0.00)	206.860*** (21.20)	212.264*** (21.37)
Folded F					
F Value		1191.18***			
R-Square		0.833	0.888	0.888	0.888
Durbin-Watson			2.004	2.004	2.004

*p<.05, **p<.01, ***p<.001

주: 괄호 안의 값은 t-value를 의미함

2. 시범사업 효과 분석

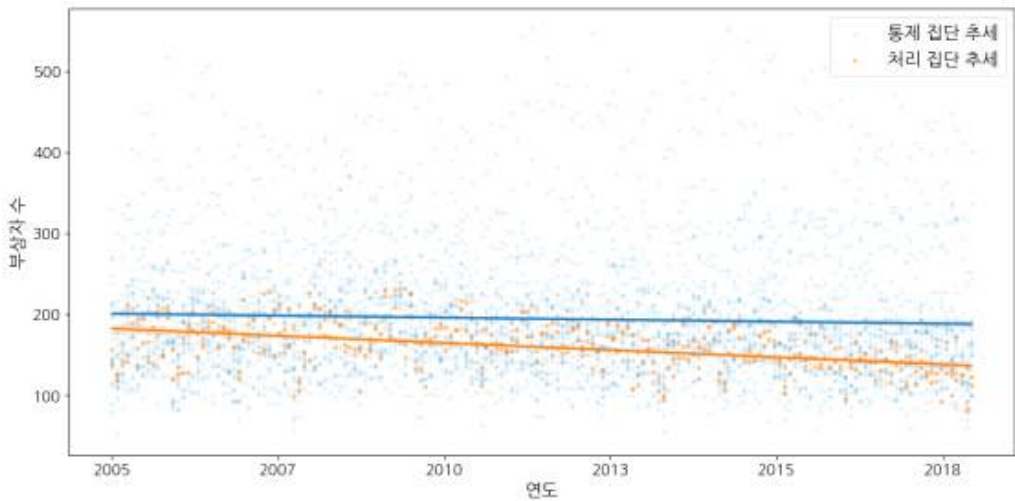
서울시는 안전속도 5030 정책의 전면 시행에 앞서, 2019년 4월 서울 4대문 안 지역을 대상으로 시범사업을 실시하였다. 이 시범사업의 효과를 분석하는 것은, 해당 정책을 서울시 전역에 확대 적용할지 여부를 판단하는 데 있어 중요한 근거가 되므로 그 검토가 필수적이다. 본 연구는 시범사업 대상 지역인 서울시 중구, 종로구를 처리집단으로 설정하고, 그 외

5) 전이함수모형을 활용하여 전면 시행효과(최대 시차=12)를 분석한 결과, 전면시행으로 총 1,037명(p-value=0.025)의 부상자 수가 감소한 것으로 나타났다. 반면 개별 시차 계수를 살펴보면 통계적으로 유의미하지 않았다.

서울시 자치구를 통제집단으로 설정하여 이중차분법, 통제집단이 있는 단절시계열 모형, 통제집단합성법을 활용한 분석을 수행하였다.

먼저 이중차분법(Difference-in-Differences)을 적용하기 위해, 정책 개입 이전에 처리 집단과 통제집단이 유사한 추세를 보였는지 여부, 즉 평행 추세 가정(parallel trends assumption)의 성립 여부를 검토하였다. 정책 시행 이전 기간을 기준으로 집단별 시간 추세에 대한 회귀분석을 실시한 결과, <표 4>와 같이 처리집단과 통제집단 모두 시간의 흐름에 따라 통계적으로 유의한 음의 추세를 보이는 것으로 나타났다. 추가로, 처리 여부와 시간변수 간의 상호작용항을 포함한 회귀모형을 활용한 분석에서도 해당 상호작용항이 통계적으로 유의하지 않아, 정책 개입 이전 양 집단 간의 추세가 유사하다는 점이 확인되었다.

<그림 6> 평행추세 가정 검토 결과



<표 4> 평행추세 가정 검토 결과

	Coef	Std err	Pr> t
개입 이전 통제집단	-0.075	0.025	0.0028
개입 이전 처리집단	-0.268	0.028	<.0001
	Type III SS	Mean Square	Pr > F
처리 여부 × 시간	88719.0159	521.8766	1.000

평행 추세 가정이 충족되었음을 확인한 후 이중차분법을 활용한 분석을 수행한 결과, 정책 개입 이후 처리집단의 부상자 수는 통제집단에 비해 월평균 약 22.02명 추가적으로 감소

했다($p<.1$). 즉, 전통적인 이중차분법을 활용한 분석 결과는 속도 제한 정책 시행이 교통사고 부상자 수 감소에 일정 부분 기여했을 가능성을 보여준다. 그리고 통제집단의 부상자 수는 월평균 약 17.23명 감소한 것으로 나타났다($p<.001$).

그러나 통제집단을 포함한 단절시계열 분석과 통제집단합성법을 활용한 분석 결과는 이중차분법에 기반한 분석 결과와는 다소 상이한 양상을 보였다. 먼저 <표 5>에 제시된 통제집단을 포함한 단절시계열 분석 결과에 따르면, 정책 시행 이후 두 집단 간 수준 변화 및 추세 변화의 차이가 모두 통계적으로 유의미하지 않게 나타나, 이중차분법 분석 결과와 달리 정책효과가 나타났다고 보기는 어려웠다. 정책 개입 이후 통제집단의 부상자 수는 시간의 흐름에 따라 매월 1.34명씩 추가적으로 감소하는 추세로 전환된 것으로 나타났지만($p<.001$), 정책 시행 이후 처리집단의 수준 변화와 추세 변화는 통제집단과 비교하여 유의미한 차이를 보이지 않았다. 이러한 결과는 정책효과를 평가할 때 단순한 평균 수준 변화에만 의존하기 보다는, 시간 추세 및 구조적 변화와 같은 시계열적 요소를 함께 고려할 필요가 있으며, 분석 방법에 따라 해석 결과가 달라질 수 있음을 시사한다.

<표 5> 시범사업 부상자 수 감소 효과 분석 결과: 이중차분법 & 통제집단이 있는 단절시계열 분석

	이중차분법	통제집단이 있는 단절시계열 분석
처리집단 여부×정책 시행 여부 (개입 이후 집단 간 수준 변화 차이)	-22.024+ (-1.67)	176.983 (0.83)
처리집단 여부 (개입 이전 집단 간 차이)	-34.903*** (-8.19)	-22.761 (-0.83)
정책 시행 여부 (개입 이후 통제집단 수준 변화)	-17.234*** (-4.63)	-22.635 (-1.15)
처리집단 여부×정책 시행 여부×시간 (개입 이후 집단 간 추세 변화 차이)		-0.768 (-0.63)
시간 (개입 이전 통제집단 추세)		-0.061* (-2.21)
정책 시행 여부×개입 시점 기준 시간 (개입 이후 통제집단 추세 변화)		-1.342*** (-4.54)
처리집단 여부×시간 (이전 추세 차이)		-0.159 (-1.26)
Intercept (개입 이전 통제집단 부상자 수)	194.286*** (161.28)	200.199*** (10.72)
F Value	38.78***	-
R-Square	0.024	0.845
Durbin-Watson	-	1.984

+ $p<.1$ * $p<.05$, ** $p<.01$, *** $p<.001$

주 1: 분석 단위는 자치구이며, 괄호 안의 값은 t-value를 의미함

주 2: '시간'은 전체 시간 흐름을 의미하며, '개입 시점 기준 시간'은 정책 시행 시점 이후 시간 흐름을 의미함

〈표 6〉은 중구·종로구와 정책 시행 이전 교통사고 패턴이 가장 유사한 자치구(은평구, 용산구, 양천구, 광진구, 강동구)를 선정하여 새로운 통제집단을 구성한 후 분석한 결과이다.⁶⁾ 이러한 절차는 처리집단과 통제집단 간 고정효과 차이 및 잠재적 이질성이 분석 결과에 미치는 영향을 최소화하기 위한 것이다. 유사 통제집단을 활용한 추가 분석 결과, 정책효과와 통계적 유의성이 더 이상 나타나지 않는 것으로 확인되었다. 이는 정책효과가 특정 통제집단의 선택에 민감하게 반응할 수 있음을 보여준다. 반면, 통제집단이 있는 단절시계열 분석은 기존 분석과 동일하게 수준 변화와 추세 변화 모두에서 유의한 정책효과를 확인할 수 없었다. 이러한 결과는 정책효과 추정이 비교집단 구성과 분석 방법의 선택에 따라 크게 달라질 수 있음을 시사한다.

〈표 6〉 유사 통제집단을 활용한 시범사업 부상자 수 감소 효과 분석 결과: 이중차분법 & 통제집단이 있는 단절시계열 분석

	이중차분법	통제집단이 있는 단절시계열 분석
처리집단 여부×정책 시행 여부 (개입 이후 집단 간 수준 변화 차이)	-7.198 (-1.19)	64.111 (0.36)
처리집단 여부 (개입 이전 집단 간 차이)	-4.397* (-2.26)	6.555 (0.37)
정책 시행 여부 (개입 이후 통제집단 수준 변화)	-32.060*** (-9.96)	8.646 (0.98)
처리집단 여부×정책 시행 여부×시간 (개입 이후 집단 간 추세 변화 차이)		-0.403 (-0.41)
시간 (개입 이전 통제집단 추세)		-0.196*** (-7.64)
정책 시행 여부×개입 시점 기준 시간 (개입 이후 통제집단 추세 변화)		-1.569** (-3.24)
처리집단 여부×시간 (이전 추세 차이)		-0.089 (-1.44)
Intercept (개입 이전 통제집단 부상자 수)	163.780*** (157.18)	179.056*** (51.27)
F Value	55.46***	-
R-Square	0.111	0.362
Durbin-Watson	-	2.006

+p<.1 *p<.05, **p<.01, ***p<.001

주 1: 분석 단위는 자치구이며, 괄호 안의 값은 t-value를 의미함

주 2: '시간'은 전체 시간 흐름을 의미하며, '개입 시점 기준 시간'은 정책 시행 시점 이후 시간 흐름을 의미함

6) 유사 통제집단은 정책 시행 이전 기간(2005-2019년)의 사전 평균 사고지표(교통사고 건수, 부상자 수, 사망자 수)를 기준으로 중구·종로구와의 유클리드 거리를 산출하여 선정하였다. 세 지표에 대한 사전 평균 값을 표준화한 후 두 시범지역과의 거리(distance)를 계산하였으며, 가장 가까운 상위 5개 자치구(은평구, 용산구, 양천구, 광진구, 강동구)를 최종 통제집단으로 구성하였다.

또한 <표 7>에 제시된 통제집단합성법 분석 결과⁷⁾, 중구와 종로구의 평균 정책효과는 각각 -7.93명, -11.32명으로 추정되었다. 그러나 시범사업이 운영된 2019년 4월부터 2020년 11월까지의 월별 효과를 살펴본 결과, 2020년 9월 중구에서 부상자 수가 41.39명 감소한 경우를 제외하면 교통사고 부상자 수의 감소 효과는 통계적으로 유의하지 않은 것으로 나타났다.⁸⁾

〈표 7〉 시범사업 부상자 수 감소 효과 분석 결과: 통제집단합성법⁹⁾

		중구		종로구	
		2019년	2020년	2019년	2020년
월별 효과	1월		-4.59 [-47.11:35.46]		-19.67 [-64.87:26.60]
	2월		-9.68 [-51.39:31.03]		3.37 [-41.46:49.29]
	3월		-33.21 [-75.06:7.24]		-29.73 [-74.61:15.83]
	4월	24.33 [-17.83:64.50]	-31.80 [-73.83:8.43]	-8.36 [-54.11:37.05]	-12.56 [-57.59:33.00]
	5월	-9.78 [-52.45:29.58]	6.10 [-36.18:46.33]	16.98 [-27.89:62.50]	-22.99 [-68.16:22.98]
	6월	-1.67 [-44.22:38.11]	-14.84 [-56.69:25.53]	-18.96 [-64.10:26.56]	-28.90 [-75.25:15.87]
	7월	2.05 [-39.21:43.08]	-39.57 [-81.84:0.44]	-15.21 [-60.23:30.16]	1.92 [-42.86:47.37]
	8월	-19.14 [-60.85:21.64]	-22.41 [-64.89:17.77]	28.15 [-17.41:73.19]	-34.10 [-78.51:12.02]
	9월	-16.17 [-56.92:25.39]	-41.39 [-82.95:-0.04]	2.91 [-42.20:48.75]	-33.00 [-77.50:13.02]
	10월	-21.68 [-63.50:18.86]	-0.85 [-41.98:39.54]	-7.71 [-53.64:37.86]	-36.62 [-81.69:8.59]

7) 통제집단합성법을 통해 추정된 정책효과에 대한 통계적 유의미성을 살펴보는 방법은 다양하지만, 본 연구에서는 Jackknife+ 방법을 활용하여 예측 신뢰구간(predictive confidence interval)을 추정하였다. Jackknife+는 처치 이전의 각 시점 데이터를 하나씩 제외한 뒤, 통제집단합성법을 적합하여 사후값을 예측하고, 실제 관측치와 예측치의 차이를 활용하여 hold-out error를 계산한다. 이 절차는 처치 이전 모든 시점에 반복적으로 적용되며, 이렇게 얻은 hold-out error를 바탕으로 예측 신뢰구간을 추정한다. Jackknife+ 방법에 대한 구체적인 내용은 Barber et al. (2021)을 참고하면 된다. 다만 통제집단합성법 추정치의 통계적 유의미성 판단 기준에 대한 논의가 활발히 이루어지고 있는 만큼, 향후 연구에서는 다양한 검정 방법에 대한 체계적 논의가 필요하다.

8) 종로구와 통제집단 평균 간의 차이, 중구와 통제집단 평균 간의 차이를 각각 종속변수로 설정하고, 전이 함수모형을 활용하여 시범사업의 효과를 분석한 결과, 두 모형 모두에서 통계적으로 유의미한 정책효과는 확인되지 않았다.

9) 통제집단합성법 분석 결과에 대한 시각화는 별첨 1에 첨부하였다.

11월	24.59 [-16.79:65.39]	-1.69 [-43.11:39.14]	-16.08 [-61.26:29.61]	-12.18 [-57.87:33.27]
12월	52.84 [11.60:93.81]		16.39 [-28.66:62.09]	
정책효과	TT: -158.57	ATT:-7.93	TT: -226.35	ATT:-11.32
모형정보	RMSE: 17.59	RMSPE: 0.12	RMSE: 19.35	RMSPE: 0.13

주 1: 괄호 안의 값은 95% 신뢰구간을 의미함

주 2: TT(Total Treatment effect), ATT(Average Treatment effect on the Treated)를 의미함

3. 소결

이상의 분석 결과는 정책효과의 추정이 적용하는 분석 방법에 따라 매우 상이한 결론에 도달할 수 있음을 보여준다. 분석 결과를 요약하면, 정책 전면 시행 효과를 단순 사전/사후 평균 비교 방식으로 추정할 경우, 서울시 전체 단위에서는 월평균 약 974.8명, 자치구 단위에서는 약 38.99명의 교통사고 부상자 수가 감소한 것으로 나타났다. 그러나 시계열적 요소를 반영한 자기회귀오차 모형 및 단절시계열 분석을 적용한 결과, 서울시 전체에서는 각각 -138.48명, -137.61명, 자치구 단위에서는 각각 -6.29명, -6.37명으로 추정되어, 단순 평균 비교에 비해 정책효과 추정치가 전반적으로 크게 감소하는 경향이 확인되었다. 특히 서울시 전체를 대상으로 한 자기회귀오차 모형과 단절시계열 분석 모두에서 추정치는 통계적으로 유의하지 않게 나타났으며, 이는 시간 경과에 따른 자연적 감소 추세나 외생적 요인을 통제하지 않는 경우 정책효과가 과대 추정될 수 있음을 시사한다.

전통적인 이중차분법 분석에서는 정책 시행이 교통사고 부상자 수 감소에 일정 부분 기여했을 가능성이 제기되었으며, 시범사업 시행 이후 대상 자치구인 중구와 종로구에서 월평균 부상자 수가 다른 자치구에 비해 약 22.04명 추가적으로 감소한 것으로 나타났다($p<.1$). 그러나 시계열 구조를 보다 정밀하게 반영하는 단절시계열 분석이나 반사실 시나리오를 기반으로 한 통제집단합성법 분석에서는 통계적으로 유의한 정책효과가 관측되지 않았다. 또한 단절시계열 분석의 경우에는 정책 시행 이후 추세 변화가 통계적으로 유의미하지 않게 나타났다. 더 나아가 비교집단의 이질성을 최소화하기 위해 중구·종로구와 사고 패턴이 유사한 자치구로 통제집단을 재구성하여 추가로 실시한 이중차분법 분석에서도 정책효과는 더 이상 유의미하게 나타나지 않았다.

이러한 결과는 정책평가의 근거를 단순한 집단 간 평균 차이로 해석할 경우, 실제 효과를 과대 추정할 수 있음 의미한다. 또한, 정책효과의 일관성과 강건성을 확보하기 위해 다양한 분석 방법을 교차 검토하는 방식이 필수적임을 시사한다. 특히 특정 시점의 평균 효과에만

초점을 두는 경우, 정책의 장기적 영향이나 실제 작동 메커니즘을 충분히 설명하지 못할 수 있다. 나아가 단일 분석 모형에 대한 의존은 복잡한 정책 환경을 과도하게 단순화시킬 수 있으며, 정책효과에 대한 판단과 해석 역시 분석 설계의 기술적 선택에 따라 달라질 수 있음을 보여준다.

V. 결론

본 연구는 교통안전 정책인 안전속도 5030을 사례로, 사전·사후 비교 검정, 이중차분법, 단절시계열 분석, 통제집단합성법 등 다양한 분석 기법을 적용하여 분석 단위와 방법론 선택이 정책 추진의 근거를 어떻게 변화시킬 수 있는지를 실증적으로 검토하였다. 그 결과 동일한 데이터를 기반으로 하더라도 분석 목적, 모형의 구조, 시점 설정 등에 따라 정책효과 추정치의 크기와 방향, 통계적 유의성이 상당한 차이를 보였으며, 이는 정책효과 추정이 분석 설계에 매우 민감하다는 점을 보여준다. 특히 본 연구는 통계분석을 통해 도출되는 수치가 절대적 진실을 의미하지 않으며, 분석 방법의 선택, 데이터의 구성, 해석의 관점 등에 따라 얼마든지 달라질 수 있음을 실증적으로 확인했다는 데 의의가 있다.

실제로 안전속도 5030 정책의 경우 초기에는 단순 사전·사후 비교에 기반한 효과가 강조되며 정책 도입의 정당화 근거로 활용되었다. 그러나 이후 추가 데이터와 다양한 분석방법을 적용하자 정책효과와 크기와 방향이 상당히 달라졌고, 정책 수정 가능성에 대한 새로운 해석도 제기되었다. 이는 전통적인 이중차분법과 같은 평균효과 중심의 방법론이 이질적 효과, 시간적 동태성, 특정 집단에 대한 차별적 영향을 충분하게 포착하지 못할 수 있다는 점을 보여준다. 결국 연구자가 무엇을 ‘효과’로 정의하며 어떤 정책적 질문을 던지는지에 따라 적절한 분석 전략은 달라질 수밖에 없다.

이러한 논의는 방법론의 변화에 따라 결과가 달라질 수 있다는 점이 자칫 근거 회의주의나 방법론적 상대주의로 비약될 수 있는 위험도 있음을 보여준다. 그러나 근거기반정책(Evidence-Based Policy, EBP)의 규범적 핵심은 다양한 연구 설계에서 제시되는 근거를 그 질적 수준에 따라 체계적으로 평가하고, 가정의 명시성, 편향 위험, 자료의 한계, 분석 과정의 투명성, 반복 검증 가능성 등 절차적 기준을 통해 객관성을 확보하는 데 있다. EBP가 말하는 “근거의 질”은 단일한 수치의 정확성에 있는 것이 아니라, 이러한 절차적 기준을 충족하는 근거가 무엇인지 평가하는 과정 그 자체에 있다. 따라서 서로 다른 방법론이 상이한 추정치를 제시하더라도 이는 “근거 무용론”으로 이어지는 것이 아니라, 어떤 조건에서 정책효과

가 안정적으로 나타나고 어떤 조건에서 달라지는지를 파악하게 하는 중요한 정보가 된다.

또한 EBP는 단일 분석 모형의 결과를 절대화하지 않고 상이한 연구 설계들이 제시하는 근거의 일관성과 방향성을 비교하는 다중 방법론적 접근(triangulation)을 요구한다. 본 연구에서 확인된 모형 간 차이는 바로 이러한 접근의 필요성을 보여주는 사례다. 결국 정책효과 평가에서 중요한 것은 특정 수치 자체가 아니라, 분석의 전 과정이 얼마나 투명하고 재현 가능하며, 가정의 현실 적합성을 확보하고 있는지 하는 절차적 객관성이다. 이러한 점에서 본 연구의 의의는 정책효과 평가에서 다양한 방법론의 적용이 필요함을 보여주는 데 있을 뿐만 아니라, 더 나아가 방법론적 절차의 엄격함을 통해 질 높은 근거를 생산하는 것이 정책 판단의 정당성과 타당성을 확보하는 핵심 조건임을 제시했다는 점에 있다.

한편, 본 연구는 자치구 단위로 집계된 자료를 활용하였기 때문에 정책의 실제 적용 범위와 국지적 이질성을 완전히 반영하기에는 한계가 있었다. 예컨대 4대문 지역 전면 시행은 특정 도심 구간을 중심으로 이루어졌으나, 행정구역 전체가 동일하게 적용 대상이 된 것은 아니며, 일부 구간의 조기 시범 운영 역시 정확한 시행 시점을 식별하기 어려웠다. 이러한 제약은 정책효과에 대한 정밀한 근거 생산에 구조적 한계가 될 수 있다. 그럼에도 서울시의 공간적 연계성과 교통 흐름 특성을 고려할 때 자치구 분석은 의의가 있다. 향후 보다 정밀한 공간 단위 자료가 확보된다면 정책효과의 정교한 추정과 근거 생성의 질적 향상에 크게 기여할 수 있을 것이다.

나아가 향후 연구에서는 본 연구가 제기한 방법론적 민감성과 근거 구성 가능성이 실제 정책 과정에서 어떻게 관리·통제될 수 있는지를 규범적·제도적 관점에서 더 탐구할 필요가 있다. 특히 다양한 분석 설계에서 생성된 근거를 질적 수준에 따라 계층적으로 평가하는 기준, 편향 위험 평가 절차, 분석 과정의 투명성 확보, 다중 방법론 적용 전략 등은 근거의 자의적 취사선택을 방지하고 정책결정의 정당성을 강화하는 핵심적인 연구과제로 남아 있다. 더 나아가 정책 과정에서 이해관계자들이 근거를 어떻게 선택·배제하며, 이러한 근거 사용 방식이 정책의 효과성과 정당성에 어떤 영향을 미치는지를 탐구하는 연구 역시 필요하다. 이는 근거기반정책 생태계의 제도화를 위해 학문적·실천적으로 중요한 방향이 될 것이다.

참고문헌

- 고길근. (2017). 자료기반 평가에서 질문주도 평가의 필요성과 효과크기 중심 평가의 한계. 『정책 분석평가학회보』, 27(2): 201-229.

- 고길곤·김경동·이민아. (2020). 질문주도 EBP 모형 개념과 적용에 대한 연구: 미세먼지 문제의 분석 사례를 중심으로. 「한국정책학회보」, 29(3): 1-27.
- 국토교통부·한국교통연구원. (2017). 도시부 제한속도 감속(5030)에 따른 교통영향 연구 최종보고서. 2017-11.
- 김기만·배관표. (2024). 도로속도제한의 규제효과 분석: 민식이법과 안전속도5030의 비교. 「규제연구」, 33(2): 49-81.
- 김하음·손승오·박준영·김기용. (2021). 도심부 제한속도변화에 따른 직접 및 간접 교통안전 효과 분석. 대한교통학회 학술대회지, 93-94.
- 김회경·김민정·정대환·김아름. (2019). 영도구 5030 속도하향 시범사업 효과분석. 대한교통학회 학술대회지, 82-87.
- 손승오·박준영·김민우·최병호. (2019). 도시부 도로 제한속도 하향 사업 안전성 효과평가. 「대한교통학회지」, 37(6): 514-524.
- 손호성·이재훈. (2018). 행정학·정책학 연구에서의 이중차분 추정기법의 활용과 쟁점. 「현대사회와 행정」, 28(3): 1-31.
- 치안정책연구소. (2022). &안전속도 5030& 종합 효과 분석 연구. 「치안논총」, 38.
- 한상진·엄기종·송원방. (2023). 도시부 도로의 안전속도 5030 시행 전후 통행시간 변화 분석. 「대한교통학회지」, 41(4): 470-482.
- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies. *Journal of the American Statistical Association*, 105(490), 493-505.
- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- Barber, R. F., Candes, E. J., Ramdas, A., & Tibshirani, R. J. (2021). Predictive inference with the jackknife+. *Ann. Statis*, 49(1), 486-507.
- Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200-230.
- Card, D., & Krueger, A. B. (1993). Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4): 772-793.
- Davies, H. T., & Nutley, S. M. (Eds.). (2000). *What works?: Evidence-based policy and practice in public services*. Policy Press.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2): 254-277.
- Head, B. W. (2010). Reconsidering evidence-based policy: Key issues and challenges. *Policy and society*, 29(2): 77-94.
- Kahn-Lang, A., & Lang, K. (2020). The promise and pitfalls of differences-in-differences:

- Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics*, 38(3): 613–620.
- Lopez Bernal, J., Cummins, S., & Gasparrini, A. (2018). The use of controls in interrupted time series studies of public health interventions. *International Journal of Epidemiology*, 47(6): 2082–2093.
- Pawson, R. (2002). Evidence-based policy: The promise of realist synthesis. *Evaluation*, 8(3): 340–358.
- Pinto, R., Valentim, R., da Silva, L. F., de Souza, G. F., de Moura Santos, T. G. F., de Oliveira, C. A. P., Santos, M. M., Miranda, A. E., Cunha-Oliveira, A., & Atun, R. (2022). Use of interrupted time series analysis in understanding the course of the congenital syphilis epidemic in Brazil. *The Lancet Regional Health–Americas*, 7: 1–12.
- Roth, J., SantAnna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2): 2218–2244.

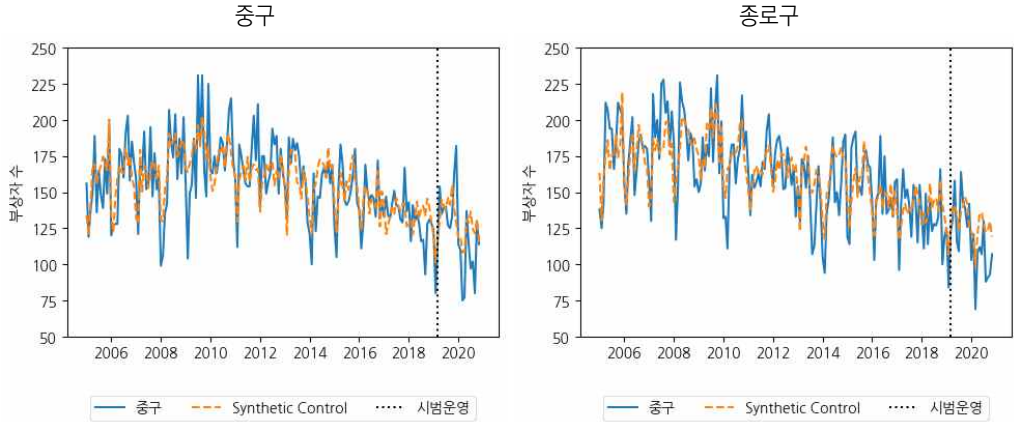
〈웹자료〉

- 경찰청·국토교통부. (2021.8.11.). & 안전속도 5030 & 시행 100일…보행 사망자 16.7% 감소. <https://www.korea.kr/news/policyNewsView.do?newsId=148891484> (검색일, 2025.7.4.).
- 경찰청·행정안전부·국토교통부. (2021.4.16.). 전국, 17일부터 & 안전속도 5030 & 본격 시행. <https://www.korea.kr/briefing/pressReleaseView.do?newsId=156447701> (검색일, 2025.7.4.).
- 대전일보 (2023.4.4.). '안전속도 5030' 전면 재수정…제한속도 상향 추진. <https://www.daejonilbo.com/news/articleView.html?idxno=2056603>(검색일, 2025.7.4.).
- 서울 열린데이터광장. <https://data.seoul.go.kr/>
- 서울특별시 교통운영과 (2018.12.6.). 서울 사대문안 안전속도 5030 전면시행…대도시 도심 첫 사례. https://news.seoul.go.kr/traffic/archives/500370?tr_code=sweb (검색일, 2025.7.4.).
- 서울특별시 도시교통실 (2022.3.28.). 안전속도5030 탄력 운영… 한강교량 등 20개 도로 제한속도 60km/h로 상향. https://www.seoul.go.kr/news/news_report.do#view/359668?tr_code=snews (검색일, 2025.7.4.).
- 한국도로교통공단 (2021.4.19.). &안전속도 5030& 4월 17일 전면 시행. https://www.koroad.or.kr/main/board/6/7554/board_view.do?sv=%EA%B5%90%EC%B0%A8%EB%A1%9C&cp=6&listType=list&bdOpenYn=Y&bdNoticeYn=N(검색일, 2025.7.4.). 한국일보 (2022.3.27.). 서울 속도 규제 일부 풀다… 한강다리 등 시속 50→60km. <https://www.hankookilbo.com/News/Read/A2022032711070001461>(검색일, 2025.7.4.).

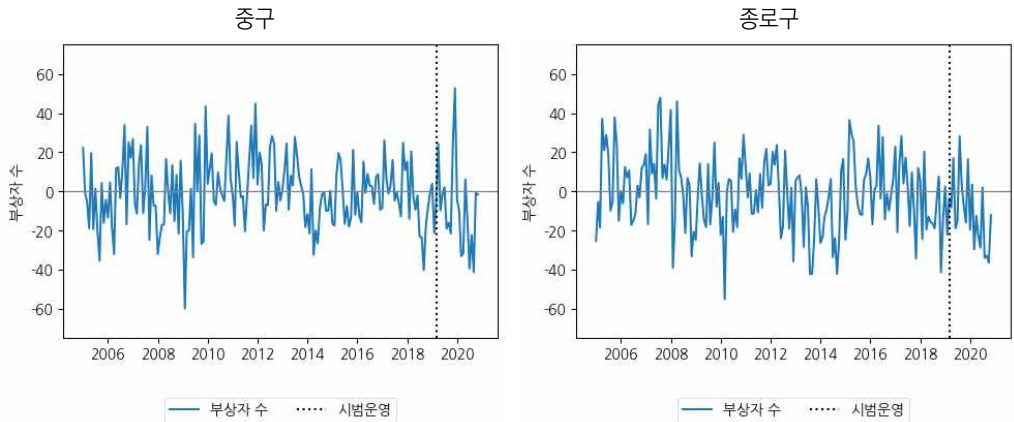
별첨 1

〈그림 1-1〉 통제집단합성법 분석결과 시각화

[통제집단합성법 분석결과]



[정책효과]



ABSTRACT

Rethinking Pre-Post Comparison and Traditional Difference-in-Differences Approaches: A Focus on Safe Speed 5030

Kilkon Ko, Youngjoo Byeon & Minjun Hong

This study examines how the estimation of policy effects, an essential component of evidence-based policy, can yield significantly different results depending on the methodological choices made, using the case of the "Safe Speed 5030" policy. This policy offers a suitable case for quantitative evaluation, as its implementation timing is clearly defined and changes in traffic accident injuries before and after the intervention are observable in quantitative terms. While early evaluations based on pre-post comparisons and difference-in-differences (DiD) analyses reported significant policy effects, this study empirically demonstrates that the estimated effects may appear minimal or statistically insignificant depending on the analytical choices made. These findings suggest that relying solely on intuitively persuasive statistics to justify policies may be misleading and that improving the quality of evidence requires methodological choices that reflect the characteristics of the data and the policy context. Moreover, the study highlights that both pre-post comparison and DiD should not be treated as inherently valid or fixed tools, but rather as analytical strategies that must be flexibly adapted and supplemented according to the specific research setting.

【Keywords: evidence-based policy, pre-post comparison, difference-in-differences, quasi-experimental design, the safe speed 5030】