

<국외훈련 결과 보고서>

데이터 분석기술을 이용한 성과감사 추진전략 연구

2020 년 7 월

감 사 원

김 현 표

국외훈련개요

1. 훈련국 : 미국

2. 훈련기관명 : 텍사스 대학교(University of Texas at Austin)

3. 훈련분야 : 감사 분야

4. 훈련기간 : 2018. 6. 25. ~ 2020. 6. 10.

훈련기관개요

훈련기관 개요

○ 훈련국가: 미국

○ 훈련기관명: 텍사스 대학교(University of Texas at Austin)

○ 주소 : 110 Inner Campus Drive STOP G0400, Austin, Tx 78712

훈련기관 소개

○ 훈련기관에 대한 개략적인 설명

- 텍사스 대학교 오스틴 캠퍼스는 연구 중심의 명문 공립대학교로 단일 캠퍼스의 학생 수 기준으로 미국 대학 가운데 다섯 번째로 규모가 큰 대학임
- 경제학과가 포함된 인문대학을 비롯하여 LBJ 행정대학원, McCombs 경영대학, Moody 커뮤니케이션대학, 공과대학 등 17 개 단과 대학·대학원, 평생학부로 구성되며, 다양한 학제간 프로그램 운영
- 주 캠퍼스는 텍사스 주도인 오스틴의 다운타운 북쪽에 위치

○ 소속학과 소개

- 경제석사 과정은 경제학적, 분석적, 통계학적 전문성이 요구되는 민간 및 공공 분야의 직업을 구하거나, 박사과정 지원을 위한 경제학, 수학 등 배경 보완을 위한 경제학 수업 제공
- 2018-2019년 기준 56명의 학생 중 47%가 중국, 인도 등 외국인이며, 경제학, 경영학, 공학, 수학 및 통계학, 정치학 등의 배경을 가짐

□ 입교교섭자료

○ 지원절차

- 지원 마감일이 UT Austin 의 다른 석사과정(보통 전년도 12 월 중순)에 비해 늦은 당해년도 3 월 1 일
- 지원 서류는 학부성적표, GRE 점수, TOEFL 또는 IELTS 점수, 추천서 3 장, SOP, RESUME 와 함께 수학 관련된 수강 내역을 요구
- 어학점수는 최소 기준 점수(TOEFL 은 PBT 550, CBT 213, IBT 79 점, IELTS 는 overall 6.5) 이상 요구하며 GRE 는 Quantitative 영역을 강조

○ 비자관련

- F-1 비자 발급을 위해 등록비용(\$1,000) 납부 후 I-20 를 발급받음
- 4 월 중 합격자 발표 후 수업은 7 월부터 시작하여 시간이 촉박함

요 약

- 4 차 산업혁명으로 촉발된 데이터 분석기술의 발전은 공공부문에
많은 기회를 제공하고 있고, 이와 같은 데이터 분석기술이 감사원의
성과감사 활성화에 어떻게 이용될 수 있을지 확인 필요

- 데이터 분석은 데이터 마이닝, 기계학습, AI, 통계추론 등 다양하게
언급되는데 다양한 통계적 기법과 패턴 인식 기술 등을 이용하여
데이터에서 유의미한 규칙 및 연관성을 확인하는 방법을 의미함
 - 컴퓨터 성능 개선 등 제반 기술을 발달은 기존의 회귀분석 이외에
인접분석, 의사결정나무, 인공신경망 분석, 군집분석 등 다양한
분석을 가능하게 함
 - 계량경제학에서는 연관관계를 보여주는 회귀분석에서 처치(정책,
사업 등)의 효과를 확인하기 위한 인과관계를 찾아내기 위해
다양한 기법을 발전시킴
 - 도구변수(IV) 기법은 처치에 영향을 주지만 결과에 영향을 주지
않는 변수를 삽입하여 처치에 의한 효과만 분리할 수 있음
 - 회귀단절모형(RDD) 기법은 처치 전후를 구분하여 회귀분석 시
확인되는 유의미한 차이를 이용하여 효과 측정
 - 패널분석은 같은 개체를 여러 시간동안 관측한 패널 데이터를
이용하여 미관측 요소를 제거하여 효과 측정
 - 이중차분분석(DiD)은 처치군과 통제군의 처치 전후 효과를 각각
비교하여 시간, 개체별 추세를 제거하는 방법으로 효과 측정

- 감사원은 회계검사와 직무감찰 위주의 합법성 감사에서 행정의 효율성을 추구하는 성과감사로 이행 추구
 - 감사원은 '94 년 성과감사를 도입, '05 년 성과감사 운영 매뉴얼을 발간, '07 년 성과감사 실무지침 작성 등 성과감사 활성화를 위한 제도적 기반 마련
 - 그간 성과감사 활성화 노력은 주요 시책의 성과 제고에 대한 관심을 고조하고 사업의 문제점을 조기 발견하는데 기여하였으나,
 - 지난 10 년간 성과감사의 빈도는 6% 내외로 실시 빈도가 낮음
 - 이는 합법성 감사에 익숙한 실무자의 선호도, 합법성 감사와 달리 확정된 지표를 이용하지 않음에 따른 정치적 논란, 주요 이해관계자의 성과감사 수요 발굴 어려움 등에 기인함
 - 또한, 성과감사 내용 측면을 살펴보면 실시된 성과감사의 경우도 사업의 효과를 확인하기 보다는 효과 달성을 위한 어떠한 노력이 수행되고 있는지, 목표와 지표 설정의 적정성 위주로 확인

- 미국 감사원(GAO)은 응용연구방법팀(ARM) 등의 전문가팀 지원 아래 계량 분석을 이용한 사업이나 정책의 효과 분석에도 적극적임
 - GAO 감사는 재무감사, 재무제표 검증, 성과감사로 구분되며, 성과감사는 효과성 · 효율성 · 경제성 평가, 합법성 감사 등으로 세분화
 - 효과의 평가에는 유사 실험 기법이 많이 이용되며, 회귀단절모형 분석이나 이중차분분석 등의 계량적 기법이 활용됨
 - (사례 1) 감기약 등에서 얻을 수 있는 필로폰 제조 성분 PSE 에 대한 정부의 판매 통제의 효과를 평가하기 위하여 통제를 이미 시행한 오레건 주와 나머지를 대상으로 한 이중차분방법을 이용

- (사례 2) 학교 현장에서 학생의 인종 등 특성과 관계없이 정당한 차별을 받고 있는지에 대하여 평가하기 위하여 구성 학생의 특징(흑인비율, 남녀비율 등)에 따른 차별 강도 관계를 분석
 - (사례 3) 증가하는 소득과 부의 편중이 고령자들의 생활에 어떠한 영향을 미치는지 분석하기 위하여 생존분석 기법 등의 예측적 분석 기법을 동원하여 인종별, 최종학력별 다양한 양상을 보임을 밝힘
- 정책의 효과 평가를 위하여 국내 초등학교 성적이 주변 아파트 가격에 미치는 효과를 이중차분분석을 이용하여 평가
- 학교 성적과 주택 가격의 관계는 상호간에 영향을 주는 요인으로 인과관계가 오인되지 않기 위한 설계가 중요함
 - 본 연구에서는 2011년 공개된 초등학교 학업성취도 분석 결과를 이용하여 우수 초등학교 통학구역 내 아파트 가격이 공개 전후 어떻게 변화하였는지를 분석
 - 우수 초등학교 주변의 아파트가 다른 아파트에 비해 2% 더 높은 가격이 상승한 것으로 확인
- 데이터 분석기술의 발전은 계량적 분석기법을 고도화할 뿐 아니라 계량분석에 이용될 데이터를 생성하는데도 도움을 줄 수 있어 정책, 사업의 효과를 평가하는데 많은 기여를 할 것으로 판단됨
- 이러한 계량적 분석기법의 발전은 정책이나 사업의 효과, 효율성 및 경제성을 살펴보는 성과감사를 활성화·고도화 할 수 있으므로,
 - 빅데이터, 통계 추론 등 데이터 분석 기술의 발전 동향을 지속적으로 파악하고 장기적인 관점에서 적용방안 마련 필요

목 차

I. 서론	13
II. 데이터 처리 기술의 발전 동향	16
1. 데이터 분석기술	16
가. 데이터 분석기술의 발전 동향	16
나. 데이터 분석기술의 종류 및 기본이론	22
(1) 지도학습	24
① 회귀 분석	
② 인접 분석(KNN)	
③ 분류 분석(Logistic Regression)	
④ 의사결정나무	
⑤ 인공지능망	
(2) 비지도학습	33
① 군집 분석(K-means)	
② 주성분 분석(PCA)	
③ 연관 분석	
2. 계량경제학 – 인과관계 분석을 중심으로	39
가. 회귀 분석과 인과관계	39

나. 인과 추론의 다양한 방법론	41
① 도구 변수(IV)	
② 회귀단절모형(RDD)	
③ 패널데이터 분석 – 고정효과 모형	
④ 이중차분분석(DiD)	
Ⅲ. 성과감사 추진 현황과 과제	49
1. 우리나라 감사원의 성과감사 추진 현황과 과제	49
가. 감사원의 감사 제도 운영 경과	49
나. 감사원의 성과감사 운영 경과	52
다. 감사원의 성과감사 운영 성과와 한계	53
2. 미 GAO 의 성과감사 추진실태	57
가. 미 GAO 조직 및 임무	57
나. GAO 의 감사 유형	60
다. GAO 성과감사의 특징 – 정책 및 사업의 ‘효과성’ 평가	63
3. 미 GAO 의 데이터 분석 기법을 이용한 효과성 평가 사례	76
가. 주요 메탐페타민(필로폰) 성분	
에 대한 접근 통제가 마약조제 실험실에	
미치는 영향(GAO-13-204)	76
나. 흑인, 남자 또는 장애 학생에 대한 차별의 적정성 검토(GAO-18-258)	77
다. 은퇴 후의 생활 보장 – 소득과 부의 불균형(GAO-19-587)	79

IV. 데이터 분석 기술을 이용한 정책의 효과 분석 예시	
- "학업성적이 우수한 초등학교가 주택 가격에 미치는 영향"	81
1. 분석 개요	81
2. 선행 연구 분석	81
3. 국내 교육제도 현황 및 학업성취도 평가	85
4. 자료 수집	87
가. 우수한 학교와 해당 학교의 통학범위 정보 수집	87
나. 주택 거래 정보 수집	91
5. 모델 구성 및 분석 결과	94
가. 이론	94
나. 모델 구성	97
다. 분석 결과	98
V. 결론	103
[참고문헌]	106
[첨부] "학업성적이 우수한 초등학교가 주택 가격에 미치는 영향 분석"시 이용한 R code	110

[표] 목차

[표 1] 종속변수와 독립변수	20
[표 2] 뉴저지(NJ)와 펜실베니아(PA)의 최저임금 차이에 따른 고용률 변화	44
[표 3] 감사원의 감사유형	48
[표 4] 감사 유형별 감사 실시현황(2004-2017)	50
[표 5] 성과감사 목록과 감사 목적(2017)	52
[표 6] 효과성을 목적으로 한 성과감사 현황과 감사초점(2017)	53
[표 7] 성과감사의 유형별 주요 감사 내용	58
[표 8] 설계 요소의 선정을 위한 질문 안내	63
[표 9] GAO 경제 분석 보고의 기본 요소	68
[표 10] 모형의 독립변수와 종속변수	75
[표 11] 분석 주제별 활용된 분석기법 현황	77
[표 12] 각 구별 초등학교, 우수 초등학교 현황	87
[표 13] 아파트 매매 정보에 대한 통계 요약	90
[표 14] 이중차분 방법에 대한 개념	93
[표 15] 단순 분석 모형 실행 결과	96
[표 16] 연도별 통제 변수가 반영된 모형의 실행 결과	99

[그림] 목차

[그림 1] 데이터 분석기술 혁명	14
[그림 2] 인공신경망 개념도	17
[그림 3] 선형회귀분석 예시	22
[그림 4] 텍사스 지역 일평균기온에 따른 전기 수요	24
[그림 5] KNN 을 이용한 분류 분석 예시	25
[그림 6] 로지스틱 회귀 분석 개념도	26
[그림 7] 의사결정나무를 이용한 회귀분석 개념도	27
[그림 8] 의사결정나무를 이용한 분류 분석 개념도	28
[그림 9] 인공신경망 개념도	29
[그림 10] 다층 퍼셉트론의 학습 개념도	30
[그림 11] 간단한 clustering 사례	31
[그림 12] K-means 분석 개념도	32
[그림 13] 주성분 분석 개념도	33
[그림 14] 연관 분석 개념도	34
[그림 15] SDO 와 ATE 의 차이점	38
[그림 16] 회귀단절모형 사례	41
[그림 17] 뉴저지(NJ)와 펜실베니아(PA)의 고용률 변화	44
[그림 18] 미국 GAO 조직도	56
[그림 19] 사업 또는 정책의 논리 모델에 대한 개념도	62

[그림 20] 이중차분분석의 개념도	71
[그림 21] 회귀단절설계 기법 개념도	72
[그림 22] 메탐페타민 사건 발생 현황(2002-2010)	74
[그림 23] 구성 학생의 특징에 따른 처벌 강도 변화	75
[그림 24] 학구도안내서비스 예시	85
[그림 25] 이중차분 방법에 대한 개념도	92
[그림 26] 우수 학교 주변 아파트와 나머지 아파트의 단위 가격 변화	97

I. 서론

현대 사회에서 정부의 기능이 고도화되고 행정의 국민 생활에 미치는 영향이 증가하면서 우리나라 감사원을 포함한 주요 선진국의 감사기구들은 합법성 위주의 감사에서 정책집행의 성과와 책임성을 제고하기 위한 성과감사를 강화하는 방향으로 감사 패러다임을 전환하고 있다.

미국 최고 감사기구인 GAO(United States Government Accountability Office)는 감사의 기능은 '부패척결', '책임성 확보', '경제성·효율성·윤리성·형평성·효과성 강화', '통찰력 제고', '예측력 촉진'의 순으로 성숙한다고 제시하고 있다.

우리나라의 경우도 감사원은 과거 부정부패 척결에 중점을 둔 합법성 감사 위주로 진행되다가 1994년 성과감사가 처음 언급된 이후 사업이나 정책(프로그램)의 효과성, 효율성, 경제성을 살피는 성과감사에 대한 관심이 높아졌고, 2000년대 시스템 감사가 도입되면서 성과감사가 본격적으로 진행되어 왔다.

특히 이번 정부는 100대 국정과제 중 하나로 성과감사를 중심으로 감사원의 역할을 강화하고자 하고 있어 감사원이 성과감사를 강화하기 위한 중요한 기회가 되고 있다.

한편 최근 인공지능, 기계학습, 빅데이터 처리, 사물인터넷 등 정보처리기술의 발달에 기인한 4차 산업혁명이 촉발되면서 관련 산업과 서비스가 증가하고 있다. 인터넷과 스마트폰의 보급으로 많은 정보가 온라인에서 생산되고 있고 사물인터넷 등 센서 기술의 발달로 정보의

디지털화가 진전되면서 대량의 데이터가 쌓이고 있고, 빅데이터 기술은 이런 대용량의 데이터를 관리할 수 있게 해 주었다. 또한 기계학습 및 인공지능에 기인한 데이터 처리 기술은 이렇게 관리된 대용량의 데이터를 처리 및 분석하여 유용한 정보를 제공하고 있다.

이러한 기회를 맞이하여 세계 각국의 정부 등 공공분야는 4 차 산업혁명으로 인해 제공되는 기회를 이용하여 행정 서비스의 품질을 제고하는 방안을 마련하는 한편 공공 부문에서 쌓인 데이터를 개방하여 활용할 수 있도록 제공하고 있으며, 4 차 산업과 관련된 기술의 확보와 산업의 발전에 많은 투자를 하고 있다.

우리 정부에서는 국정과제로 과학기술 발전이 선도하는 4 차 산업혁명을 마련하면서 지능정보 핵심기술에 투자하는 한편 데이터의 개방 및 유통의 활성화에 역량을 강화하고자 하고 있어 공공부문을 포함한 다양한 분야의 데이터를 활용할 수 있는 기반이 조성되고 있다.

해외 선진국은 이러한 기술 발전과 산업변화를 적극 활용하여 데이터 분석기술을 정부의 의사결정 지원 성과평가 등에 도입하고 있으며, 학계에서도 빅데이터 등 데이터 분석기술을 정부 정책의 평가 등에 적용하는 연구가 많이 진행되고 있다.

특히, 미국 GAO는 과학기술, 경제 분석, 데이터 분석 및 예측 등의 분석기술을 제공하는 전문 조직인 ARM(Applied Research and Method) Team 을 운영하면서, 데이터 마이닝, 계량 분석, GIS 분석 등 다양한 통계 및 데이터 분석 기법을 동원하여 부조리 사례를 적발하거나 정책사업의 성과와 효율성을 평가하는 등 계량적 분석에 적극 활용하고 있다.

성과감사는 규정의 준수 여부를 확인하는 합법성 감사와 다르게 사업이나 정책의 효과, 효율성 및 경제성(3E: Efficiency, Effectiveness, Economy)을 지표로 평가를 하고 있어 분석에 이용할 수 있는 데이터를 수집하고, 가공하여 분석할 수 있는 데이터 분석기술에 따라 분석 결과의 품질이 좌우된다고 할 수 있다.

본 연구에서는 빅데이터 등 첨단 데이터 분석 기술의 기본 이론과 발전 동향 및 활용에 대해 알아본 후, 감사원 성과감사 추진 현황과 발전을 위해 필요한 조건과 해외의 성과감사 추진 현황에 대해 살펴보겠다.

또한, 이를 바탕으로 현재 진행 중인 데이터 분석기술의 발전이 성과감사를 활성화하는데 어떻게 이용될 수 있을지 알아보도록 하겠다.

Ⅱ. 데이터 분석기술의 발전 동향과 기본 이론

1. 데이터 분석기술

가. 데이터 분석기술의 발전 동향

인공지능, 사물인터넷, 빅데이터 등 디지털 기술의 발전이 새로운 산업의 변화를 이끌고 있다는 4 차 산업혁명은 2016 년 1 월 스위스 다보스에서 열린 세계경제포럼(World Economic Forum: WEF)에서 Klaus Schwab 가 처음 언급하였다.

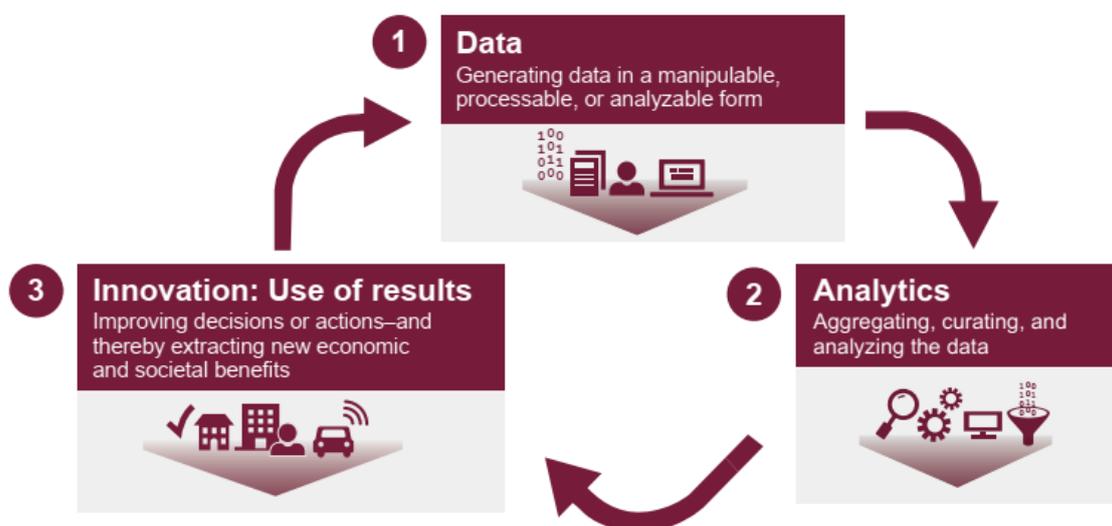
4 차 산업혁명에 대한 정의는 학자에 따라 조금씩 다르지만 공통적으로 언급되는 키워드나 관련 기술 동향을 위주로 살펴보면, 인터넷과 스마트폰 보급으로 구현된 초연결사회에서 사물인터넷 등의 센서 기술을 통해 현실 세계의 정보가 빠르게 데이터화 되고, 이렇게 수집된 대량의 데이터를 빅데이터 기술로 관리하는 한편, 기계학습으로 대표되는 인공지능 기술을 이용하여 분석, 가공하여 3D 프린팅이나 로봇 기술들로 다시 현실로 구현되는 정보통신기술의 융합으로 이루어진 산업 체계로 설명할 수 있다.

4 차 산업혁명이라는 용어는 2016 년 1 월 나타난 것이지만 제반 기술과 산업은 여러가지 키워드와 함께 지속적으로 언급되어 왔다. 특히 빅데이터 처리, 기계학습 등으로 대표되는 데이터 분석기술(Data Analytics)의 발전과 보급은 보다 많은 사람들이 실생활에서 데이터 분석 기술을 경험하거나 직접 활용하도록 하고 있어, 4 차 산업혁명에 대한 언급

이전부터 민간 다양한 영역 뿐 아니라 정부 및 공공부문에서도 데이터 과학, 데이터 마이닝, 데이터 분석 등 다양한 표현으로 언급되면서 관심을 가져왔다.

미국 GAO 는 이러한 추세 동참하여 2016 년 9 월 “Data and Analytics Innovation” 을 주제로 포럼을 개최하여 현재의 데이터 분석 기술의 현황을 확인하고 그 발전이 향후 어떤 기회를 제공하고 예상되는 문제점은 최소화할 수 있는지에 대하여 확인하였다. 여기에서는 데이터 분석 기술의 혁명은 [그림 1]과 같이 ① 대용량의 다양한 데이터가 많은 분야에서 활용 가능해지고, ② 데이터 분석 기술의 발전은 더욱 강력한 데이터 분석을 지원하며, 데이터 분석 가능성을 변경시켰으며, ③ 분석 결과는 인간의 의사결정을 지원하거나 자동으로 임무를 수행하고 그 결과는 다시 data 로 피드백되는 3 단계의 혁명 절차로 나타난다고 보았다.

[그림 1] 데이터 분석기술 혁명



자료: GAO

데이터 분석은 대용량의 데이터 속에서 알고리즘과 수학적 처리과정을 거쳐 패턴을 찾아내어 유의미한 결론을 도출하는 과정으로 데이터 마이닝(Data Mining), 기계학습(Machine Learning), 인공지능(Artificial Intelligence), 통계학습(Statistical Learning), 통계추론(Statistical Inference) 등 다양한 개념으로 언급된다.

데이터 분석기술이 우리 실생활에 밀접하게 다가와 있다는 것은 다양한 사례를 통해 확인할 수 있다. 많이 알려진 사례는 미국의 유통기업인 타겟(Target)에서 아직 10 대인 딸에게 출산용품에 대한 할인쿠폰을 보내서 부모가 항의하러 갔더니 실제로 딸이 임신 중 이었고, 타겟은 소비자의 구매 패턴에 대한 데이터 분석을 통해 맞춤형 홍보를 하였던 것이다. 또한 동영상 재생 서비스를 제공하는 YouTube 에서 사용자가 플레이한 영상 기록을 이용하여 추천 영상을 제공해주거나 스마트폰에서 촬영한 인물 사진이 자동으로 인물별로 구별되어 관리되는 등의 서비스가 모두 데이터 분석기술로 구현한 것이다.

현대적인 데이터 분석 기술은 1946 년 컴퓨터의 발명, 1950 년 Alan Turing 의 튜링 테스트(Turing Test) 설계 및 1956 년 Dartmouth 학회에서 인공지능(Artificial Intelligence: AI)라는 용어 제안으로 본격 시작했다고 볼 수 있다.

이후 1956 년부터 1974 년 까지를 황금기라고 불리며 많은 연구가 진행되고 투자가 이루어졌다. 이 시기에 최초의 인공 신경망 분석(Neural Networks)인 Perceptron 이 설계되고, 인접 분석(The nearest neighbor), 군집분석(Cluster Analysis), 의사결정나무(Decision Tree)

등 분석 방법이 제안되었으며, 1959 년 기계학습(Machine Learning) 개념이 소개되었다.

이렇게 활발하던 인공지능에 대한 연구는 컴퓨터 연산 능력의 부족, 현실 세계와의 차이점 등에서 기인한 한계가 드러났다. 특히 인공 신경망 분석인 Perceptron 은 AND/OR 같이 선형 관계는 해결할 수 있지만 XOR 문제의 해결은 불가능하다는 것이 1969 년 수학적으로 증명되었다. 이에 따라 미 국방부는 AI 연구자금을 지원을 중단하는 것을 계기로 인공지능 연구에 대한 투자가 전반적으로 감소하였고, 1974 년부터 1980 년까지 첫번째 인공지능의 겨울(The first AI Winter)이 찾아온다.

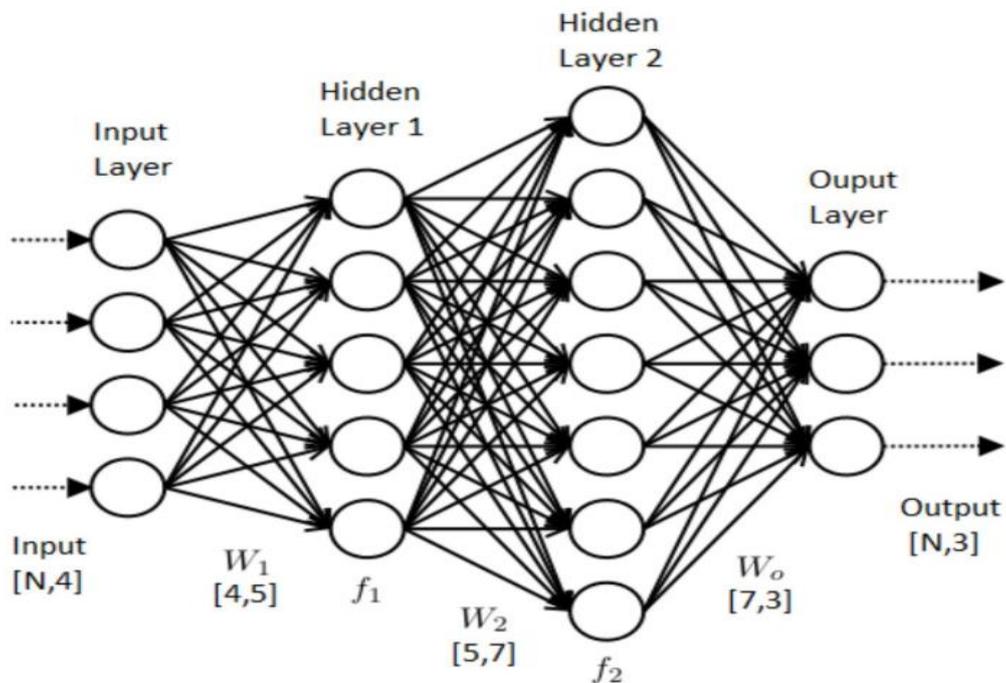
이후 인공지능에 대한 연구는 통계적 기법에 집중되어 실험 설계와 통계분석을 통해 데이터마이닝이라는 개념으로 발전되었고, 이는 현재의 빅데이터 기술로 이어지고 있다.

한편 인공지능에 대한 연구는 1980 년 보다 실용적인 분야로 세분화하여 전문가 시스템(Expert System)의 형태로 도입되면서 상업적인 분야에 많이 적용되었다. 지식과 경험을 데이터베이스화 하여 의사결정 추론을 통해 사용자에게 정보를 제공하는 형태로 구성되었고, 의사결정 추론은 베이즈(Bayes)의 확률적 방법과 0 과 1 사이 다양한 실수값을 가질 수 있는 퍼지(Fuzzy)이론이 주로 활용되었다.

이 추세는 전문가 시스템이 복잡해지면서 지식 추출의 병목 현상(knowledge acquisition bottle-neck)이라는 문제가 발생하여 한계가 드러났으며 짧게 조성된 AI 의 봄은 1987 년 이후 1993 년까지 두번째 인공지능의 겨울이 찾아왔다.

한편 1960년대 소개된 인공신경망 분석은 1985년 신경망의 처리를 거치는 숨겨진 레이어(Hidden Layer)를 도입하면서 [그림 2]와 같이 인간의 뇌를 모델로 하는 알고리즘이 완성되었으나 당시의 컴퓨터 연산 능력의 한계로 실용화될 수 없었다.

[그림 2] 인공신경망 개념도



1995년 기계학습에 대한 통계적 접근 방법인 Support Vector Machine(SVM) 개념이 소개되면서 이론적 배경이 부족했던 기계학습과 인공지능에 엄격한 수학적 분석이 적용될 통로를 제공하였다.

또한 같은 1995년 의사결정나무 기법을 여러 개 모아 각 단계에서 최적의 답을 찾아가는 기법인 Random Forest 기법이 제안되었고, 1997년 IBM에서 개발한 인공지능 Deep Blue가 세계 체스 챔피언을 이기는 사건이 있었다.

2000년 이후 컴퓨터 연산 성능의 개선과 빅데이터 처리 기술의 발전 등으로 기계학습 모델에 대한 관심은 더욱 증가하였고, 많은 IT 업체들이 인공지능과 기계학습 또는 데이터 처리기술에 대한 투자를 통해 관련 서비스를 제공하기 시작하였다. 또한 2016년은 우리 나라에서도 큰 화제를 모았던 구글의 인공지능 AlphaGo가 이세돌과의 바둑 대결에서 승리하면서 우리나라에도 인공지능에 대한 관심이 매우 높아졌다.

오늘날에는 빅데이터와 인공지능은 단순 마케팅 용어가 아니라 인공지능 기반의 자율주행, AI 스피커, AI 를 이용한 상품 추천 및 고객관리 서비스 등 이를 기반으로 한 서비스가 많이 제공되고 있다. 뿐만 아니라 정부의 사업 설명자료와 장기계획에는 빅데이터와 인공지능에 대한 언급이 빠지지 않고, 연구개발 과제 대부분이 빅데이터와 인공지능에 한발짝 걸치고 있을 정도로 인공지능의 전성기가 도래하였다.

시간을 조금 더 과거로 돌려 초기 개념의 인공지능 또는 데이터 분석의 연혁을 찾아보면 그 시작은 통계적 추론의 기본 이론인 베이즈 이론(Bayes' Theorem)이 소개된 1763년 까지 거슬러 올라간다. 베이즈 이론을 통해 불확실성 아래에서 관측된 값을 이용하여 의사결정을 수행할 있었다. 이후 1800년대에는 데이터를 모형에 fitting 하는데 이용되는 최소제곱법이 소개되고 이를 이용한 회귀분석법(regression analysis)이 도입되었다. 회귀분석은 통계적 가정 사항에 기반하고 있어 사회과학 분야에서는 통계적 가정을 약화시키면서 상관관계 또는 인과관계를 확인하기 위한 연구가 지속되었다.

나. 데이터 분석기술의 종류 및 기본이론

데이터 분석(Data Analysis)은 다양한 통계적 기법과 인공지능을 이용한 패턴 인식 기술 등을 이용하여 데이터에서 유의미한 패턴, 규칙 또는 관련성을 발견하는 방법이다.

데이터 분석은 많은 분야에서 활용되는데 고객은 누구이고, 어떤 물건을 구매하고, 해당 물건에 대한 평가는 어떠한 지 등의 정보 같은 고객 정보를 이용하여 고객별, 제품별 판매 패턴 등을 발견(Mining client information)할 수 있고, 온라인 서비스를 제공하는 사람은 방문자가 누구이고, 얼마나 자주 또는 오랫동안 머무르며, 어떤 물건을 얼마나 구매하는 지 등의 온라인 행동 정보를 이용하여 온라인 구매 고객의 소비를 예측하거나 소비 촉진을 위한 nudge 를 발견(Online behavior)할 수 있다. 또한 당신이 하는 행위를 누가 하는지에 대한 정보로 선호도를 예측하여 추천해주거나(Collaborative filtering), 블로그, 트위터, 메일 및 뉴스에 나온 글의 주요 용어, 글의 감정상태 등을 포착(Text Mining)할 수 있다.

데이터 분석을 위해서는 데이터의 수집과 데이터의 전처리 및 변환(cleaning), 분석(analysis), 문서화 및 시각화 등의 과정을 거치는데, 분석(analysis)도 중요하지만 실제 데이터 분석에서는 데이터와 수집과 전처리 및 변환 과정에 많은 시간과 노력이 소모되고, 도출된 결과를 간결하고 분명한 문장과 아름다운 도표로 생산해내는 것도 중요하다.

데이터 분석은 다량의 데이터, 즉 관측 값을 이용하여 모델을 만들고 테스트하는 것을 반복하여 최적의 모델을 도출하는데 이 과정을 학습

(Learning)이라고 한다. 데이터 분석 이론은 이 학습 과정의 특징에 따라 지도 학습(Supervised Learning)과 비지도(Unsupervised) 학습으로 구분된다.

지도학습은 우리가 예측하고자 하는 목표 변수 y 값을 다양한 관측 변수 x 값들로 예측할 수 있는 모델을 만드는 것으로, x_i 의 입력 값 대해서 y_i 가 도출된다는 결과를 반복하면서 학습을 시키는 것이다. 대표적으로 회귀분석(regression analysis)과 분류분석(Classification analysis)이 있고, 기본적으로 다음의 식으로 표현된 관계식에서 y 값이 연속된 수치(연속형)이면 회귀 분석, 구분된(categorized) 불연속 값(범주형)이면 분류 분석으로 구분된다.

$$y_i = f(x_i) + e_i$$

여기서 우리가 관측한 데이터인 x 와 y 값을 이용하여 미지의 함수인 f 를 구하는 것이며, 현실의 데이터를 이용하는 만큼 x 값에 대응하는 y 값을 설명해주는 함수가 모든 값에 대하여 정확히 일치할 수 없기 때문에 발생하는 오차 e 가 존재하며, 학습은 이 오차를 최소화시키는 방향으로 진행한다.

여기에서 x 를 독립변수(Independent Variable), y 를 종속변수(Dependent Variable)라고 하는데, 통계, 경제, 컴퓨터 공학 등 다양한 분야에서 [표 1]과 같이 다양한 이름으로 언급된다.

[표 1] 종속변수와 독립변수

Y	X
Dependent variable	Independent variable

Explained variable	Explanatory variable
Response variable	Control variable
Predicted variable	Predictor variable
Regressand	Regressor

자료: Introductory Econometrics (Jeffrey Wooldridge)

비지도학습은 지도학습과 다르게 x 값에 대응하는 y 값의 관계를 주지 아니하고 y 값의 분포만 가지고 비슷한 유형의 데이터를 모으는 등 패턴을 찾아내는 분석 방법이다. 대표적으로 군집 분석(Clustering) 방법이 있다.

중요한 지도 학습과 비지도 학습의 분석 방법에 대한 기본 이론과 내용을 살펴보면 다음과 같다.

(1) 지도학습

① 회귀분석(Regression Analysis)

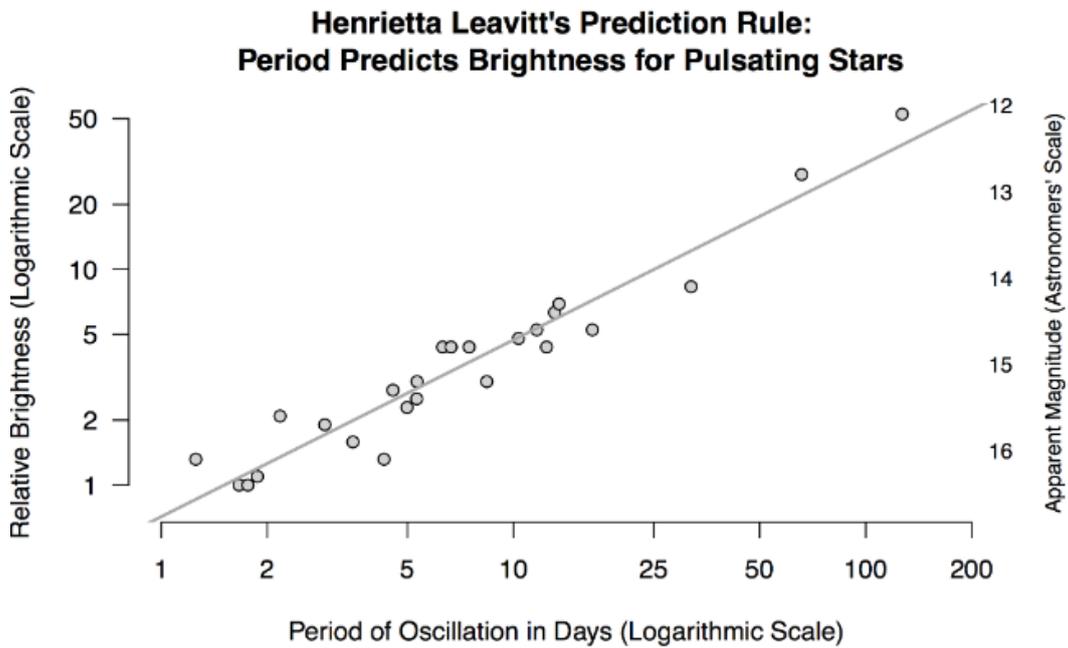
회귀분석은 가장 오래된 분석 방법 중 하나로 단순하고 직관적이기 때문에 데이터 분석이라는 틀을 벗어나서 과학과 산업, 경제 및 사회과학 분야에서 가장 많이 이용되는 분석방법이다.

가장 간단한 선형회귀분석(Linear Regression)을 기준으로 기본 이론을 설명하면, 회귀분석은 다음의 식에서 각 (x, y) 값에 따른 오차 e 를 가장 작게 만드는 식을 구성하는 계수(coefficient) β 를 찾는 과정이다.

$$\begin{aligned}
 y &= \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p + e & x &= (1, x_1, x_2, \dots, x_p) \\
 &= x \cdot \beta + e & \beta &= (\beta_0, \beta_1, \beta_2, \dots, \beta_p)
 \end{aligned}$$

이때 계수 β 를 찾는 과정에서 이용되는 방법이 최소제곱법(Least Squared Error)으로 n 개의 관측 값 $(1, x_1, x_2, \dots, x_p; y)$ 을 넣었을 때 발생하는 각 오차 e 의 제곱의 합이 최소가 되는 계수 β 의 집합이 최적의 모델이다. 다음 [그림 3]에서 점은 각 관측 값이고 직선은 관측 값을 이용하여 추정된 모형이다. 이때 각 관측 값은 모형으로 추정한 값과 동일하지가 않는데 각 관측 값에서 직선까지 떨어진 거리가 오차이며, 이 오차의 제곱의 합을 최소로 하는 최적의 직선 기울기 β 를 찾아낸다.

[그림 3] 선형회귀분석 예시



선형 회귀 분석은 모델이 단순하여 모델의 분산이 가장 낮다 - 즉, 새로운 값 또는 특이점(outlier)이 발견되어도 모델이 크게 변하지 않는다.

또한 단순한 구조로 인해 모델을 해석하기 용이하다. 예를 들면 여기에서 계수 β 는 x 가 1 만큼 변할 때 y 가 β 만큼 변한다는 탄력성을 나타내고, 양쪽에 로그를 취한 변수로 모델을 구성하였다면 계수 β 는 x 의 1% 변화에 따른 y 의 변화율을 의미하는 등 직관적으로 해석이 가능하다. 아울러 회귀분석 모형은 강력한 통계적 이론에 근거하고 있어서 결과값에 대한 신뢰도가 높다.

그러나 회귀분석은 위 [그림 3]에서 확인하였듯이 대부분 예측값(직선)이 관측값(점)과 일치하지 않는다. 즉 모델의 분산은 작지만 예측값의 분산은 크다. 또한 통계 이론에 근거하고 있지만 해당 모델이 최적의 모델이기 위해 필요한 가정 사항이 대부분 만족시킬 수 없기 때문에 예측값에 편차가 있을 수 있다.

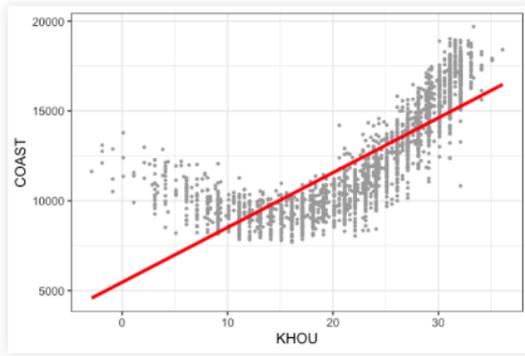
② 인접 분석(K-Nearest Neighbors Analysis)

인접 분석(KNN: K-Nearest Neighbors Analysis)은 회귀분석과 다르게 계수를 이용하지 않는 (Nonparametric) 방법으로 x 값에 대응하는 y 값을 예측하기 위하여 가장 가까이 있는 K 개의 관측 값(x, y)을 이용하는 분석 방법이다.

회귀분석과 다르게 $y = x * \beta + e$ 같은 식을 별도로 구성하지는 아니하고 x 값에 대응하는 예측 값 y 의 집합으로만 구성된다.

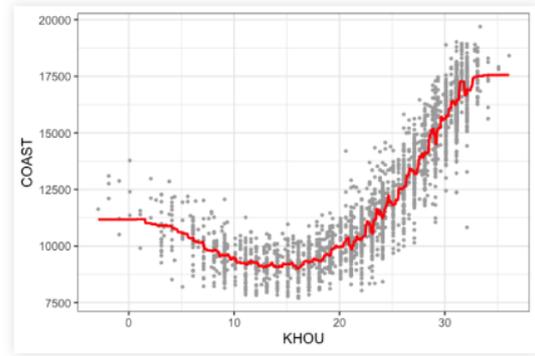
회귀분석과 KNN 의 차이를 살펴보면, 다음 [그림 4]에 나타난 도표는 텍사스주 전기 사용량에 대하여 일 평균 온도(KHOU)에 따른 전기 수요(Coast)를 회귀분석과 KNN 분석을 통해 나타내고 있다.

[그림 4] 텍사스 지역 일평균기온에 따른 전기 수요



Parametric:

$$f(x) = \beta_0 + \beta_1 x$$



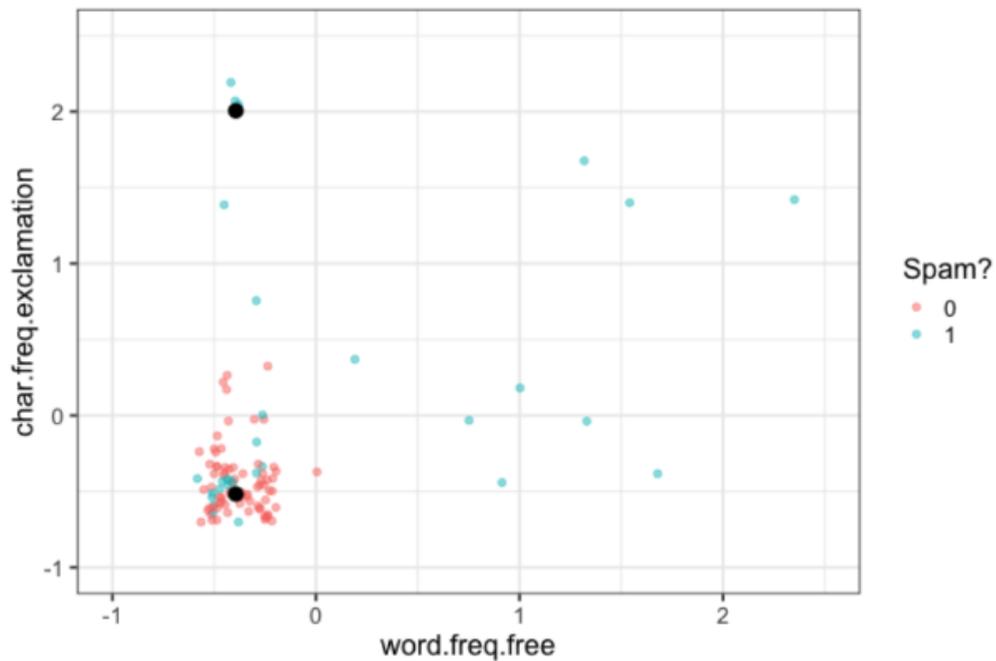
Nonparametric (k-nearest neighbors):

$f(x)$ = average y value of the 50 points closest to x

KNN 분석에서는 전기 사용량을 예측하고자 하는 날의 일 평균기온 값 주변에서 가까운 순으로 K 개의 관측 값을 추출한 후 평균을 구하면 우리가 예측하려는 값이 도출된다. 이때 K 값이 작으면 예측 값은 정확하지만 모델이 복잡하고 변동성이 커서 새로운 데이터에 의해 값이 급격히 변할 수 있고, K 값이 크면 예측 값은 완만하게 변하지만 정확성은 낮아진다. 이와 같이 참조할 주변 관측값 K 개의 선정에 따라 결과 값이 달라지므로 최적의 예측 결과를 K 값을 미리 선정하는 것이 필요하다.

한편 KNN 분석에서 y 값이 연속된 값이 아닌 범주형 데이터이면 분류 분석(Classification)을 실시할 수 있다. [그림 5]는 메일에 작성된 문자 특성을 이용하여 스팸 메일을 분류하고 있는데, 새로운 메일인 검정색 점 주변의 K 개의 메일이 평균적으로 스팸이면 스팸메일로 구분하는 방법이다.

[그림 5] KNN 을 이용한 분류 분석 예시

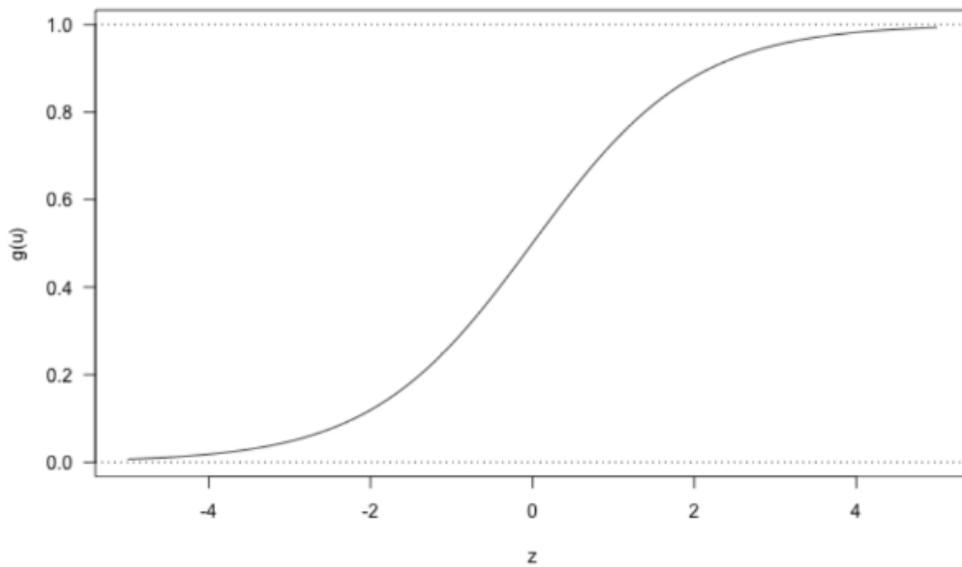


③ 분류 분석 (Classification): Logistic Regression 모델

분류는 회귀 분석과 다르게 종속변수 y 값이 연속 값이 아닌 범주형 데이터이다. 즉 0 또는 1로 표현된다. 로지스틱 회귀 분석은 기본적으로 회귀 분석을 이용하여 x 값에 대하여 y 값을 예측한다. 이때 y 값은 0 또는 1이므로 모형에서 도출된 y 의 예측 값도 0 또는 1로 배정하여야 한다.

예를 들면 회귀 분석을 통해 나온 모델로 예측한 y 값이 0.7이 나오면 1과 가까운 값이므로 1이고, 0.2라고 나왔으면 0과 가까우므로 0으로 예측할 수 있다. 하지만 이렇게 연속적으로 나온 값을 범주형 값으로 변환하는 과정에서 0 또는 1로 변하는 과정이 임계 값에서 분명히 드러나게 하기 위하여 [그림 6]과 같은 변환 함수를 이용하는 것이 로지스틱 회귀 분석(Logistic Regression: Logit)이다.

[그림 6] 로지스틱 회귀 분석 개념도



A standard choice is $g(z) = e^z / (1 + e^z)$.

- At $z = 0$, $g(z) = 0.5$.
- When $z \rightarrow \infty$, $g(z) \rightarrow 1$, and when $z \rightarrow -\infty$, $g(z) \rightarrow 0$.

④ 의사결정나무 (Decision Tree)

의사결정나무 분석 방법은 각 단계(node)에서 해당 항목의 관측 값에 대한 목표 값을 연결해 나가면서 예측 값을 찾는 분석 방법으로 논리의 흐름이 나무 모양을 하고 있어 의사결정나무라고 한다.

의사결정나무는 각 단계에서 질문에 따라 오른쪽 또는 왼쪽을 선택하여 아래로 내려가는 방식으로 결과에 도달하도록 모델이 구성되고, 예측된 결과가 최선의 값이 아닐 수는 있지만 논리의 흐름이 직관적이어서 이해하기 용이하다.

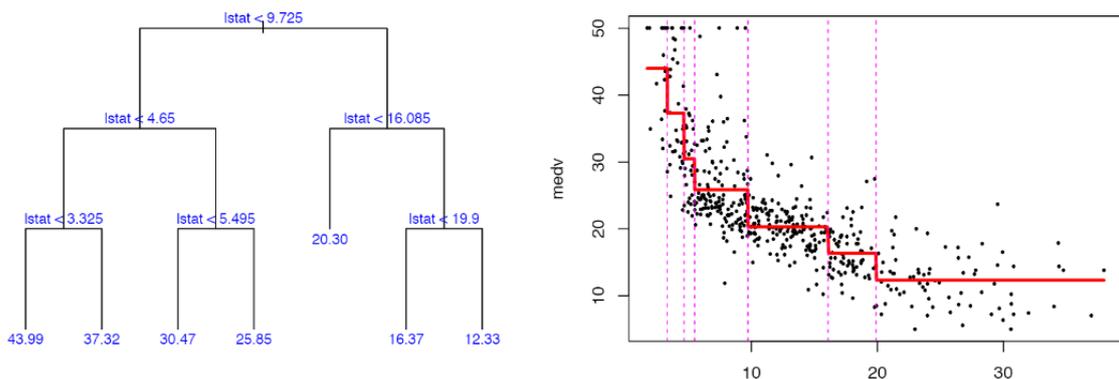
이와 같은 의사결정나무는 각 단계에서 분류되는 기준을 어떻게 구성하느냐에 따라 예측의 정확성과 속도가 증가하며, 모델의 구성과정은 바로 이러한 분류 기준을 배치하는 절차이다.

각 node 에서 양쪽으로 비슷하게 나누어질 분류 기준을 우선 적용하여야 최종 결과에 빠르게 도달할 수 있고, 재귀적 분기와 가지치기 (prune) 기법을 통해 이러한 분류기준의 배치를 효율적으로 진행한다.

의사결정나무의 구조상 분류 분석에만 적용될 것으로 생각할 수 있지만 분류 분석 뿐 아니라 회귀 분석에도 이용된다.

회귀 분석을 살펴보면 [그림 7]과 같이 각 노드에서 기준에서 예측하고자 하는 값이 어디에 해당하는지 선택되어 최종 예측 값을 찾는다. 이렇게 나온 함수는 선형 회귀분석과 다르게 직선이 아닌 계단형 모양으로 구성된다.

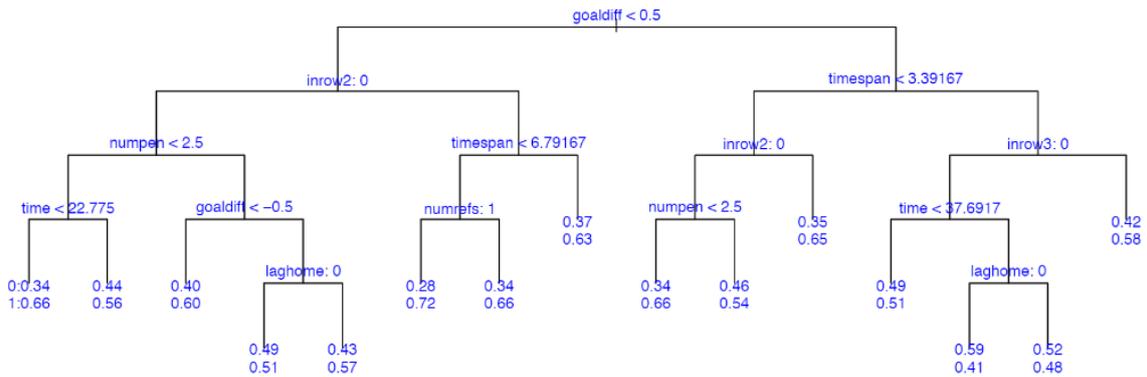
[그림 7] 의사결정나무를 이용한 회귀분석 개념도



분류 분석의 경우 더욱 이해가 용이하다. [그림 8]과 같이 각 단계(node)에서 학습을 통해 분류 기준에 해당하는 항목에 대하여

의사결정나무가 작성되며, 각 단계에서 질문에 대답을 하면서 예측 값을 찾아내면 된다.

[그림 8] 의사결정나무를 이용한 분류 분석 개념도

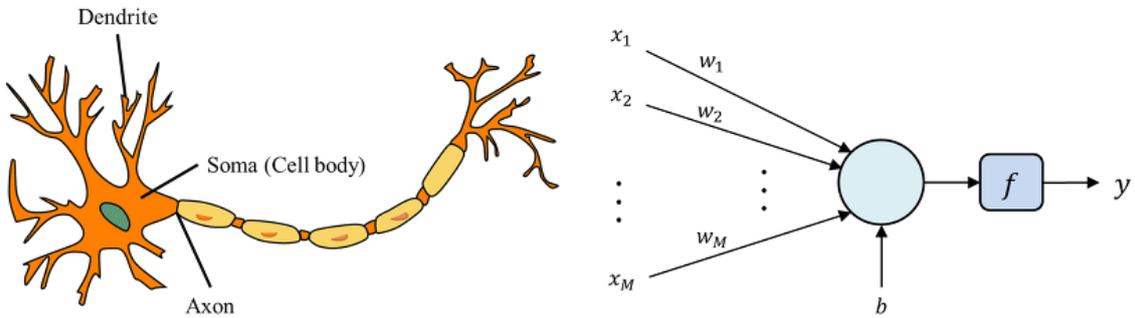


⑤ 인공신경망(ANN, Multi-Layer Perceptron)

인공신경망 분석(Neural Network Analysis)은 생물체의 신경 세포망을 따라 전기적 신호가 전달되어 하나의 정보로 형상화되는 과정을 추상화한 분석 기법이다.

[그림 9]의 왼쪽 그림은 생물체의 신경 세포망을 형상화하고 있는데, Dendrite 를 통해 전달되어 온 개별 전기 신호가 Soma 에서 합산되어 Axon 에서 하나의 정보로 형상화되어 뇌로 전달된다. 이와 마찬가지로 [그림 9]의 오른쪽 그림에서 주어진 데이터 x 값들은 각각의 가중치 w 가 반영된 연산을 수행하는 함수를 거쳐 결과 값 y 가 도출되는데, 이러한 개념으로 모델을 구성한 것을 인공신경망 분석 기법이라고 한다.

[그림 9] 인공신경망 개념도



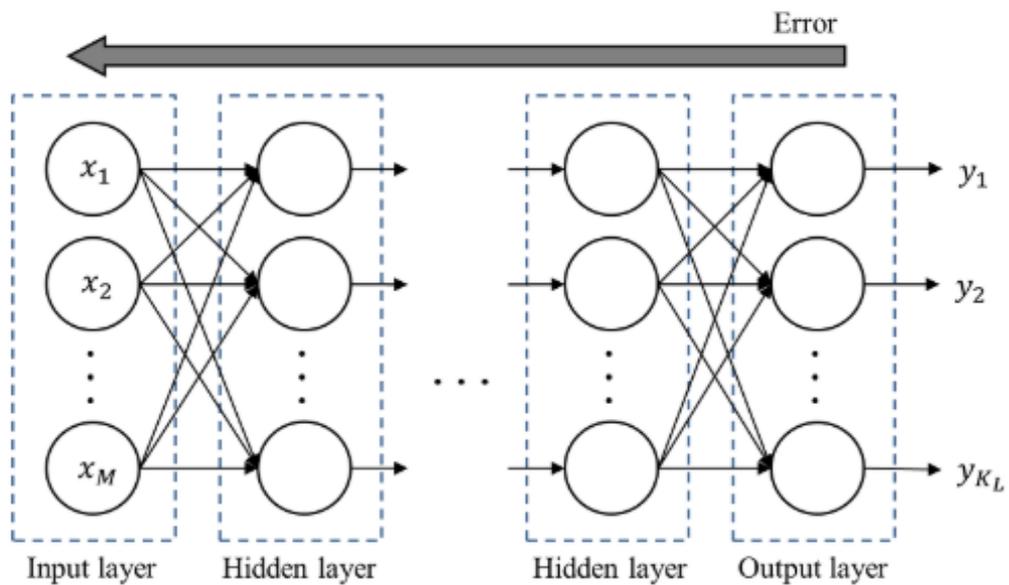
인공신경망 분석에서 학습은 입력 노드에서 개별 관측 값이 입력되면서 가중치 w 값이 조정되는 것으로 학습 데이터(Training Data)에 초기 가중치(w)를 반영하여 계산된 결과값이 임계 값을 초과하면 출력 노드에는 1 또는 -1 로 예측 값이 도출되고, 이때 나온 예측 값이 실제 값과 다를 경우 가중치를 업데이트하여 다음 학습 데이터를 통해 같은 과정을 반복한다.

이러한 과정을 반복하는 동안 가중치는 계속 갱신되면서 가중치의 변화가 일정한 값으로 수렴하게 되면 학습이 완성된다.

한편 이러한 단편적인 신경망은 앞서 인공지능의 역사에서 살펴본 바와 같이 선형적인 논리 분석이 아닌 XOR 관계를 처리할 수 없어 도입 초기에는 큰 성과를 보지 못하였다. 이러한 문제는 입력 값과 출력 값의 중간에 숨겨진 노드(Hidden Layer)을 추가하여 반복적인 관계식 구성을 통해 해결할 수 있고, 이 방법이 현재 가장 널리 사용되는 다층 퍼셉트론(Multi-Layer Perceptron) 분석 방법이다.

다층 퍼셉트론에서는 [그림 10]과 같이 각 노드마다 가중치가 반영되어 계산된 값이 다음 층에서 입력 값으로 이용되어 최종 예측 값이 나오면 실제 값과의 차이를 구한 후 다시 오차를 반대로 전파하는 방식으로 가중치에 대한 학습이 된다.

[그림 10] 다층 퍼셉트론의 학습 개념도



(2) 비지도 학습

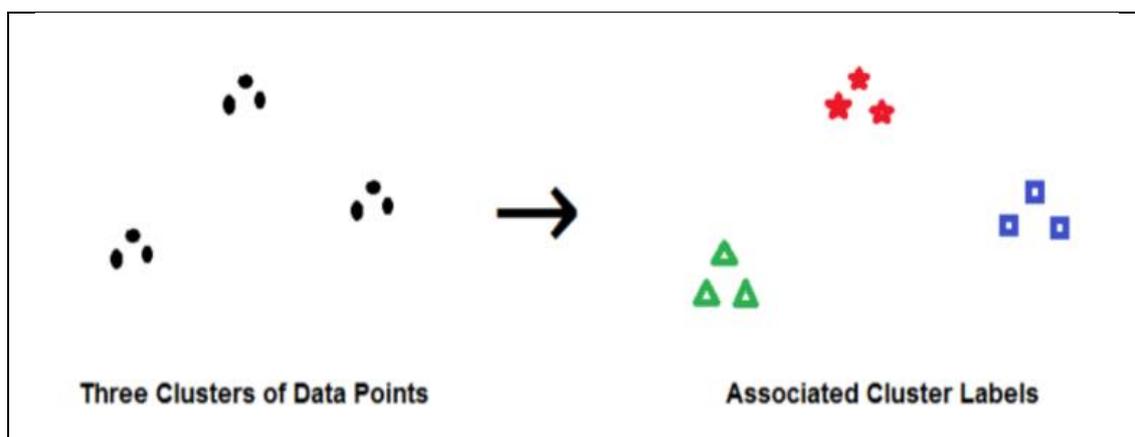
① 군집분석 (Clustering) - K-means

군집 분석 (Clustering)은 대표적인 비지도 학습법으로 관측 값들을 사전에 정의되지 않은 분류 기준으로 구분하는 방법으로, 분류 분석 (Classification)과 다른 점은 분류는 사전에 분류 기준이 명확하게

구분되어 있고, 군집 분석은 분석이 끝난 이후에 분류 기준을 알 수 있다는 점이다.

[그림 11]의 왼편에 위치한 관측 값들을 그룹화한다고 하면 다양한 기준으로 나누어 볼 수 있겠지만 가장 간단히 생각할 수 있는 것이 오른편과 같이 가까이 위치한 관측 데이터들끼리 그룹으로 묶는 방법이다. 따라서 관측 값들 사이의 거리를 측정할 수 있다면 우리는 관측 값들을 분류할 수 있을 것이다.

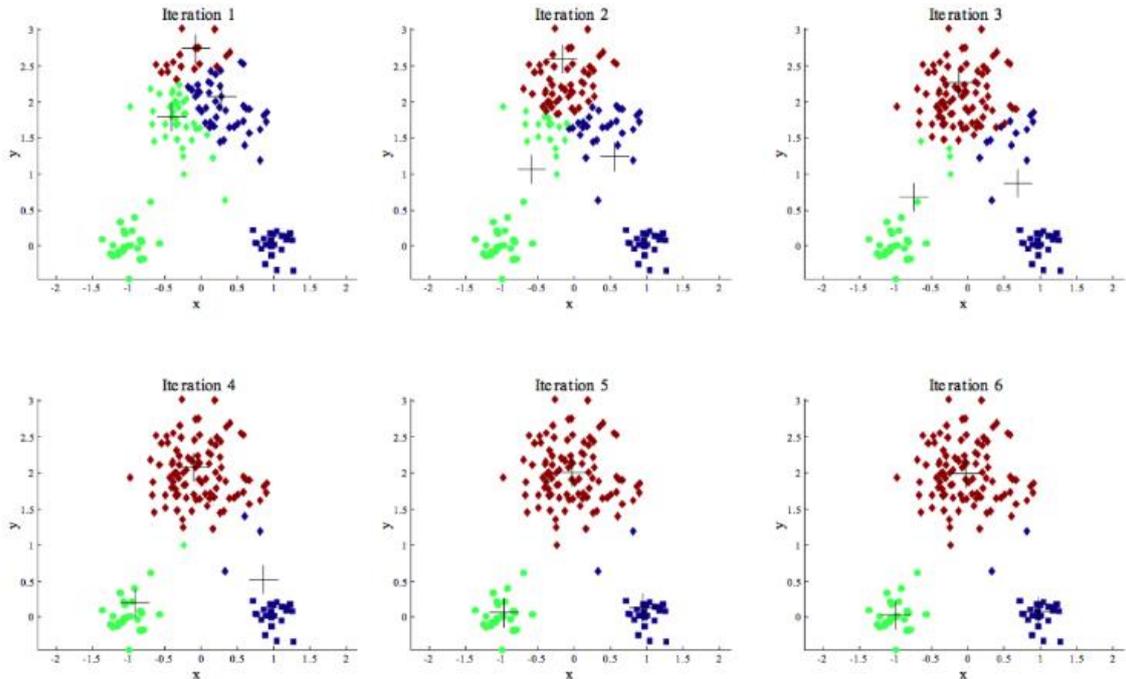
[그림 11] 간단한 clustering 사례



이러한 군집 분석 기법 중 간단하면서 많이 사용되는 분석 방법은 K-means clustering 이 있다.

K-means 는 [그림 12]와 같이 K 개의 중심점을 임의로 선택한 후 거리 함수를 이용하여 모든 관측 값을 K 개의 중심점에 할당한다. 이때 중심점은 할당된 관측 값들의 평균에 해당하는 값이 된다. 이후 새로운 K 개의 중심점을 임의로 선택한 후 앞의 과정들을 반복하며, 중심점을 바꾸면서 군집화를 반복하다 보면 더 이상 군집이 변하지 않게 되고 학습이 종료된다.

[그림 12] K-means 분석 개념도



K-Means 군집분석은 개념은 단순하지만 초기 중심점에 따라 군집이 수렴되는 방향이 달라질 수 있고, 중심점의 개수 K 값에 따라서 달라질 수 있다.

따라서 초기 중심점을 어떻게 선택할 것이고, 중심점의 개수인 K 값을 무엇을 할 것인지를 결정하여야 한다. 초기 중심점의 경우는 K 개의 점들이 최대한 멀리 떨어져 선정되면 특이점을 향해 군집이 수렴하는 경우를 예방할 수 있어 합리적인 군집을 형성하도록 할 수 있다. 또한 초기 중심점의 개수는 K 값을 변경해가면서 분석 결과를 검토하여 최적의 결과를 보여주는 K 를 선정한다.

이때 군집분석은 비지도 학습으로 학습시킬 데이터를 이용하여 오차를 비교하는 지도학습 기법으로 최적 하이퍼파라미터 값을 찾는 방법을 이용할 수 없다. 이 경우는 분석 값의 배치를 이용한 개념으로

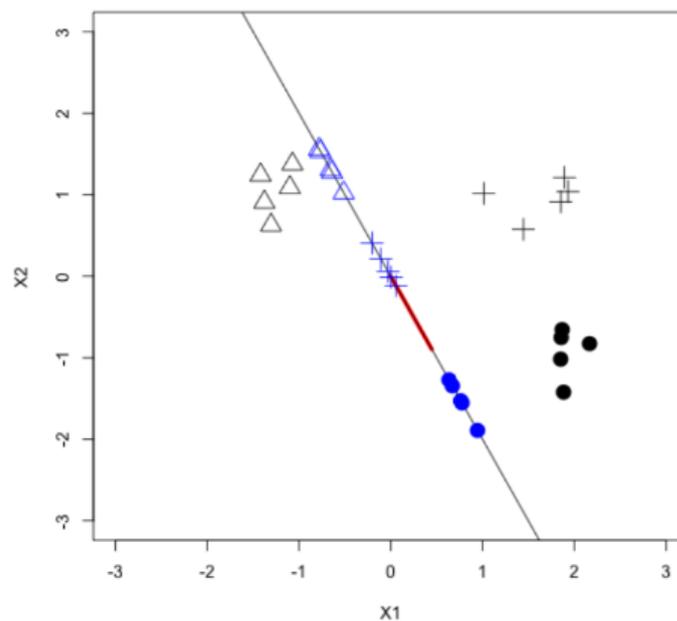
elbow plot 이나 CH index 또는 Gap Statistic 을 이용하여 구분이 명확해지는 K 값을 찾는다.

② 주성분 분석(PCA)

주성분 분석(Principal Components Analysis)은 높은 차원을 가진 데이터를 낮은 차원으로 요약하는 분석 기법으로 데이터들을 가장 잘 설명해줄 수 있는 주요 성분을 찾아 해당 주 성분을 기준으로 각 데이터를 설명하는 방법이다.

주성분 분석은 각 데이터들을 주성분에 투영한다는 개념으로 생각할 수 있다. [그림 13]에서 삼각형, 십자, 점으로 표시된 관측 값이 X-Y 평면에 표시되어 있지만 직선으로 표시된 주성분에 투영되면 각 관측 값들은 주성분 내에서 단순한 거리로 표시된다.

[그림 13] 주성분 분석 개념도



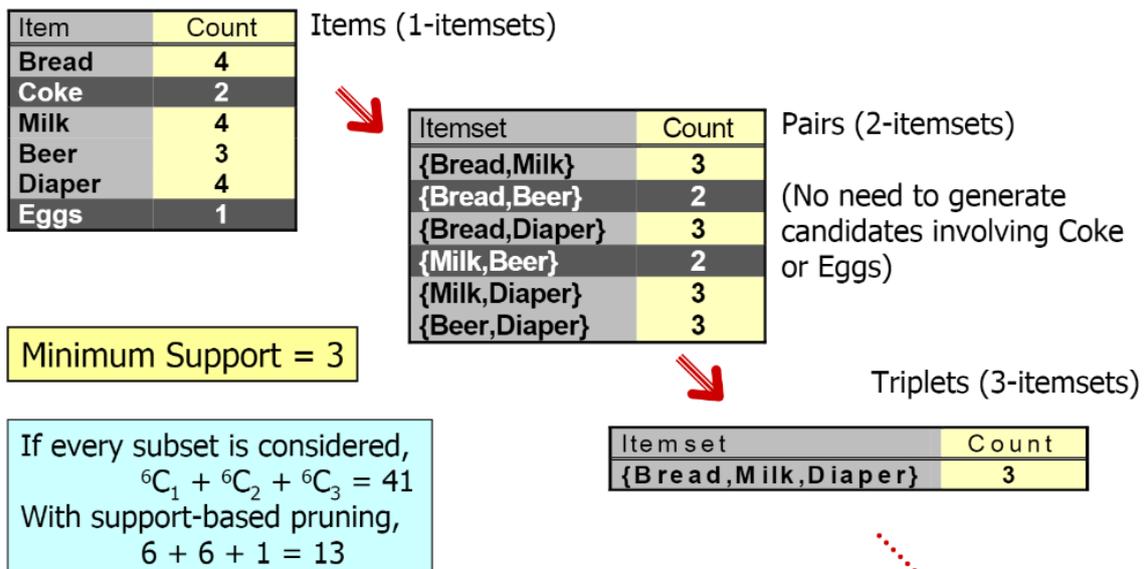
여기에서 주성분은 각 관측 값들이 최대한 명확하기 떨어질 수 있는 성분을 정하며, 처음 주성분으로 설명되지 않는 나머지 오차를 그 다음 주성분을 결정하여 점들을 구분하는 방식으로 주성분을 결정한다.

이때 소수의 주성분으로 데이터 간 차이를 최대한 많이 설명할 수 있다면 그 주성분 분석 모델의 성능이 좋은 것으로 판단할 수 있다.

③ 연관분석 (Association Rule)

연관분석 (Association Rule)은 대규모의 데이터들 중에서 서로의 연관성을 확인하여 상관 관계를 찾는 기법이다. 일반적으로 음악감상 사이트나 영상 사이트에서 음악이나 영상을 추천하거나 고객들의 구매패턴을 활용한 마케팅에 많이 이용된다.

[그림 14] 연관 분석 개념도



연관 분석의 이론은 [그림 14]와 같이 장바구니를 이용하여 설명한다. 예를 들면, 고객들의 구매 내역을 확인해보니 일반적으로 빵과 우유를 구매한 고객은 달걀과 콜라를 구매할 확률이 높았다는 분석 결과를 확인한 후에, 어떤 고객이 빵과 우유를 구매하면 그 고객도 달걀과 콜라를 구매할 것으로 추정하여 할인 쿠폰의 제공 등 타겟 마케팅을 하는 것이다.

연관분석에서는 다양한 개별 데이터의 묶음 조합 가운데서 거래가 빈번하게 일어나는 조합을 발견하는 것이 중요하다. 이러한 조합의 선택을 계량화 된 지표를 이용하여 수치화하며, 주로 이용되는 지표로 지지도 (support), 신뢰도 (Confidence), 향상도 (Lift)를 이용한다.

‘X 이면 Y 다’ 라는 연관 규칙과 관계하여 지지도, 신뢰도, 향상도 개념을 설명하면, 지지도는 전체 데이터 묶음 중 X, Y 를 모두 포함하는 데이터 묶음의 비율을 말한다. 즉 지지도가 높다는 것은 해당 규칙이 빈번하고 구성비가 높다는 것으로 지지도가 낮은 규칙을 제거하는 기준으로 활용된다. 신뢰도는 X가 포함된 데이터 묶음 중 Y도 포함된 데이터 묶음의 비율(일종의 조건부 확률)을 의미한다. 따라서 신뢰도가 높다는 것은 X 이면 Y 라고 할 수 있어 해당 규칙이 유용하다는 것을 보여준다. 향상도는 X가 없을 때 Y의 확률 대비 X가 있을 때의 Y의 확률의 증가 비율로 향상도가 1 이라면 X와 Y는 서로 독립되어 아무 관련이 없다는 것을 나타내며(X 유무와 관계없이 Y 확률이 동일), 향상도가 클수록 X가 존재할 때 Y가 존재하는 비율이 높아 둘이 동시에 존재하는 게 우연이 아닌 X 이면 Y인 연관관계가 유의하다고 해석할 수 있다.

2. 계량경제학 - 인과관계 분석 (causal inference) 을 중심으로

계량경제학(Econometrics)은 경제 현상에 대한 실증적인 내용을 제공하기 위하여 통계 기법을 이용한 추론을 통해 관측된 경제 데이터에서 경제 이론과 관련된 경험적인 결과를 도출하거나 관측에 기반하여 경제 현상을 분석하는 정량적인 분석 방법이다.

데이터 분석에서는 다양한 기계 학습 방법을 이용하여 모델을 구성하는데 계량경제학에서는 통계학적 이론 기반이 확실하게 있는 회귀분석(Regression Analysis)을 주로 이용한다. 또한, 데이터 분석이 예측을 목적으로 하였다면 계량경제학은 주로 추론에 집중하여 도출된 계수 β 를 통해 효과를 추정하는 것을 목적으로 한다.

가. 회귀분석과 인과관계

계량 경제학은 사건으로 인해 발행하는 영향을 측정하는 기법으로 설명할 수 있고, 회귀분석(Regression analysis)을 통해 계산된 계수 (Coefficient) β_1 를 통해 그 영향이 어느 정도 인지를 확인한다.

$$y = \beta_0 + \beta_1 \times x + u$$

하지만 기본적으로 회귀식에서 도출된 계수는 x 값과 y 값의 연관성 (Correlation)을 나타낼 것이지만 반드시 인과관계 (Causality)를 의미하지는 않는다.

예를 들어서 도시별로 인구 대비 경찰 인력의 수와 범죄율의 상관관계를 확인하였을 때, 인구대비 경찰 인력의 수가 많을수록 범죄율이 높았다는 결과가 나올 수 있다. 이 결과는 경찰이 많으면 오히려 범죄가 더 많이 발생하니 경찰을 줄여야 한다는 결론을 의미하지는 않는다. 경찰 인력과 범죄율의 상관관계가 높다는 것이 확인되었을 뿐이지만 범죄율이 높은 도시이기 때문에 경찰 인력의 증원이 많았던 것인지 그 반대인 것인지를 확인하는 것이 바로 인과관계를 확인하는 것이다.

데이터 분석 기법은 대부분 x 와 y 값 사이의 상관 관계에 대해서만 관심이 있고 특히 현대의 복잡한 기계학습 기법들은 왜 그러한 관계가 있는지에 대해서는 관심이 없고 단지 그러한 패턴이 보인다는 것만 확인한다. 하지만 계량 경제학은 데이터 분석과 다르게 예측보다는 추론을 목적으로 하고 그렇다면 계수 β 가 의미하는 것이 인과 관계가 있다는 것을 확인하여야 한다.

자연 과학에서는 처치(Treatment)가 있을 때 어떤 결과가 나오는지 확인하기 위하여 처치가 있는 그룹과 별도로 통제 그룹인 처치가 없는 그룹을 구분하여 실험을 실시한다. 이와 같이 통제된 그룹과 처치된 그룹의 결과를 비교하여 처치와 결과 사이의 인과 관계나 처치의 효과를 확인할 수 있다.

하지만 사회 과학에서 제기되는 질문은 자연 과학과 같이 직접적인 실험을 실시하기 어려운 경우가 많다. 사람들의 생각과 행동을 확인하기 위해 너무 많은 시간과 비용이 요구되거나 처치 여부의 선택에 있어 윤리적인 문제가 제기될 수 있어 직접적으로 실험이 실시되는 경우는 드물고, 통제군을 구분하여 실험을 실시하더라도 통제가 확실하지 않은

경우가 많다. 따라서 이러한 한계에서 처치와 결과의 인과관계 또는 처치의 효과를 밝히기 위하여 여러 기법들이 활용되어 왔다.

나. 인과 추론(causal inference)의 다양한 방법론

사회 과학에서 인과추론이 어려운 이유는 처치에 따른 효과를 파악하기 위해서는 처치가 있었던 그룹의 결과가 만약 그 그룹에 처치가 없었다면 발생할 결과와 어떤 차이가 있는지 확인하여야 한다. 하지만 이때 처치가 없었다면 발생하였을 결과는 실제로는 발생하지 아니한 사실(Counterfactual)이기 때문에 결과를 관측할 수 없다.

만약 실험에 참여하는 개체가 모두 동일하다면 통제 그룹을 설정하여 처치를 통제함으로써 효과를 바로 확인할 수 있지만 실제로는 실험 개체별 조건이 다르고 처치 그룹과 통제 그룹을 구분할 때에도 무작위로 되었다고 할 수 없기 때문이다.

[그림 15] SDO 와 ATE 의 차이점

$$\underbrace{E_N[y_i|d_i = 1] - E_N[y_i|d_i = 0]}_{\text{SDO}} = \underbrace{E[Y^1] - E[Y^0]}_{\text{Average Treatment Effect}} + \underbrace{E[Y^0|D = 1] - E[Y^0|D = 0]}_{\text{Selection bias}} + \underbrace{(1 - \pi)(ATT - ATU)}_{\text{Heterogenous treatment effect bias}}$$

[그림 15]에서 우리가 확인하여야 하는 처치의 효과는 평균 처치 효과(Average Treatment Effect: ATE)이지만 우리가 관측할 수 있는

결과는 결과의 단순 차이 값(Simple Difference in mean outcomes: SDO)이다. 만약 처치 그룹과 통제 그룹을 선택하는데 편차가 없고 처치의 효과가 균일하게 나타난다면 SDO 값이 ATE 와 동일하다고 볼 수 있다.

즉, 인과관계 분석에서는 표본의 추출 과정이 무작위이며 표본들 간에 처치의 효과는 균일하게 나타난다는 가정을 통해 SDO 값을 평균 처치효과로 이용하며, 이러한 가정을 충족시키거나 편향을 제거하기 위한 다음과 같은 방법론이 존재한다.

① 도구 변수(Instrumental Variable)

도구 변수는 처치(X)의 효과(D)가 결과(Y)로 나타날 때, 만약 우리가 관측할 수 없는 요인(u)이 있어서 처치와 결과에 모두 영향을 미친다면 처치로 인한 결과의 변화가 처치에 의한 것인지 아니면 관측되지 않은 그 요인에 의한 것인지 알 수 없다.

이 경우 처치에 영향을 미치지만 결과에는 영향을 주지 않는 도구 변수(Instrumental Variable: IV)를 삽입하여 Z가 변할 때 Y의 변화를 살펴봄으로써 관측되지 않은 요인 u에 의한 영향을 제외하고 처치에 의한 효과만 추적할 수 있다.

대표적인 도구 변수 활용 사례로는 경찰 인력 증가로 인한 범죄 발생 감소 효과를 측정한 사례가 있다. 경찰 인력이 증가하면 범죄율이 감소하는 것이 합리적인 결과지만 범죄율이 높은 도시에는 경찰 인력의 증원이 발생하기 때문에 경찰 인력의 증가와 범죄 발생 비율의 변화 사이의 상관 관계를 분석하면 경찰 인력의 증가와 범죄 발생 비율의 증가 사이에 양의

상관관계가 있다고 나올 수 있다. 즉 경찰이 많을수록 범죄율도 증가한다는 모순된 결과가 나올 수 있다.

이러한 문제를 해결하기 위하여 범죄율과 관련이 없지만 경찰 인력의 변화에 영향을 주는 도구변수(IV)를 삽입하여 범죄율 증가에 따른 경찰 인력 변화 효과를 삭제할 수 있다.

Steven Levitt(1997)은 시장 선거의 주기를 도구변수로 이용하여 이를 해결하였는데, 시장 선거가 있는 해는 지지율 상승을 위해 경찰 인력의 증원이 활발한 경향이 있다는 점에 착안하여 해당 시기의 경찰력 증가는 범죄율과 관련 없고, 이에 따른 범죄 감소 여부를 확인하여 경찰력과 범죄 발생 사이의 인과관계를 확인하고자 하였다.

또한 워싱턴 D.C. 지역은 주요 인사가 참석하는 행사가 빈번하게 개최되는데 이 경우 주요 인사의 경호와 질서 유지를 위해 경찰 인력 증원 발생한다. 이러한 경찰 인력 증가는 범죄율과 무관한 인력 증가이므로 이를 이용하여 경찰 인력의 증가와 범죄 발생 사이의 인과관계를 확인할 수 있다.

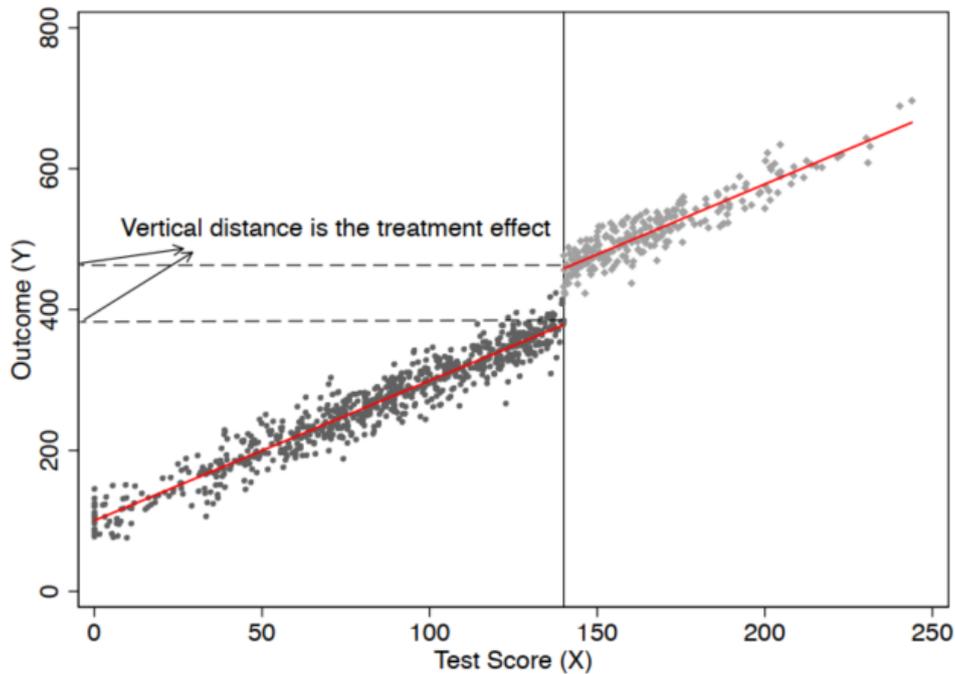
이와 같은 도구 변수 기법은 얼마나 합리적으로 적당한 도구 변수의 선정하느냐가 중요하다. 도구 변수의 결정에는 결국 경제 이론과 논리적인 추론을 통해 찾아내므로 연구자의 직관력이 중요하다.

② 회귀단절모형 (Regression Discontinuity Design)

회귀단절모형은 사건의 발생 또는 처치로 인하여 결과가 변경되는 것을 처치 전후의 데이터만 이용하여 각각 회귀모형을 도출하여 회귀모형이 서로 단절되었다는 것을 통해 효과를 확인하는 기법이다.

회귀단절모형은 [그림 16]과 같이 처치가 실행된 시간 전후의 데이터를 구분하여 각각 회귀 분석을 실시하며 이때 두 그래프 사이에 유의미한 간격(Gap)이 발생하면 처치의 효과가 있다고 볼 수 있다. 이때 유의미한 간격은 두 그래프가 신뢰범위 안에서 서로 만나지 않는 것을 의미하고 그때의 차이 값이 처치 효과가 된다.

[그림 16] 회귀단절모형 사례



또한 이때 나타난 단절이 처치에 의한 것이 분명한지 확인하기 위하여 처치가 없는 통제 그룹에 대한 회귀단절모형을 실행하여 통계적으로 유의미한 단절이 발생하지 않는다는 것을 확인하는 것이 중요하다.

③ 패널데이터 분석(Panel data analysis)-고정효과(Fixed Effect) 모형

패널 데이터는 같은 개체(개인, 회사, 학교, 미국의 주, 나라 등)에 대하여 여러 시간동안 관측된 데이터를 의미하며, 이 경우 결과 값에 포함되어 있지만 우리가 확인할 수 없었던 관측하지 못한 요소(unobserved factor)는 각 개체별로 시간의 경과에 따라 동일하다고 가정하고 패널 처리를 통해 미관측 요소를 제거하는 방법이다.

이러한 미관측 요소가 시간의 경과에 관계없이 동일하다고 가정하는 고정효과(fixed effect) 모형과 시간 경과에 따라 무작위로 변경된다고 보는 임의 효과(random effect) 모형이 있다. 일반적으로 경제학에서는 미관측 요소를 고정된 효과로 가정하는 고정효과 모형을 이용한다.

고정효과 모형에 따라 패널데이터를 이용한 미관측 효과를 제거하는 효과를 확인하면, 다음 식에서 우리가 관측하지 못한 요소를 c 라고 할 때 다음과 같이 시간의 경과에 따라 평균을 내면 다음과 같이 미관측 요소 c 가 양쪽에서 남아있다. 이때 이 두 식의 차이를 구하면 미관측 요소 c 는 삭제되고, 우리가 확인하고자 하는 계수 β 를 구할 수 있다.

$$\begin{aligned} y_{it} &= x_{it}\beta + c_i + \varepsilon_{it} \\ \bar{y}_i &= \bar{x}_i\beta + c_i + \bar{\varepsilon}_i \end{aligned}$$

$$\begin{aligned} (y_{it} - \bar{y}_i) &= (x_{it} - \bar{x}_i)\beta + (c_i - c_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \\ \dot{y}_{it} &= \dot{x}_{it}\beta + \dot{\varepsilon}_{it} \end{aligned}$$

④ 이중차분분석 (Difference in Difference: DiD)

처치(Treatment)의 효과(Effect)가 무엇인지 확인하기 위해서는 처치 전후의 차이를 비교하거나 처치 이후에 처치군과 통제군이 차이를 비교하는 방법을 생각해 볼 수 있다. 하지만 처치 전후의 차이만 비교할 경우 처치 그룹에 대한 결과의 관측 값이 시간 경과 또는 주변 여건 변화에 따라 변경되었을 수 있지만 이를 반영할 수 없다.

또한 처치 이후의 처치 그룹과 통제 그룹을 비교할 경우 처치 그룹과 통제 그룹별 개개의 특성에 따라 발생하는 차이가 반영되지 않을 수 있다.

$$\text{Outcome}_{it} = \beta_1 + \beta_2 \text{Treat}_i + \beta_3 \text{Post}_t + \beta_4 (\text{Treat} \times \text{Post})_{it} + \varepsilon_{it}$$

이러한 경우 처치 그룹과 통제 그룹 각각 처치 전후 시간에서 결과값의 차이를 구한 후 처치 그룹의 차이 값과 통제 그룹의 차이 값의 차이 값을 구하면 관측되지 않았던 시간 요인과 개별 특성 요인 모두 소거된 처치 효과를 도출할 수 있다.

이중차분 분석을 이용하여 뉴저지와 펜실베니아의 최저임금 인상의 차이에 따른 고용률의 변화를 살펴본 연구가 유명하다. [표 2] 및 아래 식과 같이 각각 뉴저지(NJ)와 펜실베니아(PA)에서 뉴저지의 최저임금 인상 전후의 고용률 변화는 각각 λ , $\lambda + \delta$ 이다. 양쪽 주에서 고용률 변화 값의 차이를 계산하면 δ 는 2.76 으로 나오는데, 이 값은 뉴저지의 최저임금 인상으로 인한 고용률 변화 효과를 의미하고 최저임금 인상으로 오히려 고용률이 2.76 만큼 상승한 것을 보여주고 있다.

$$Y_{its} = \alpha + \gamma NJ_s + \lambda d_t + \delta(NJ \times d)_{st} + \varepsilon_{its}$$

PA Pre: α

PA Post: $\alpha + \lambda$

NJ Pre: $\alpha + \gamma$

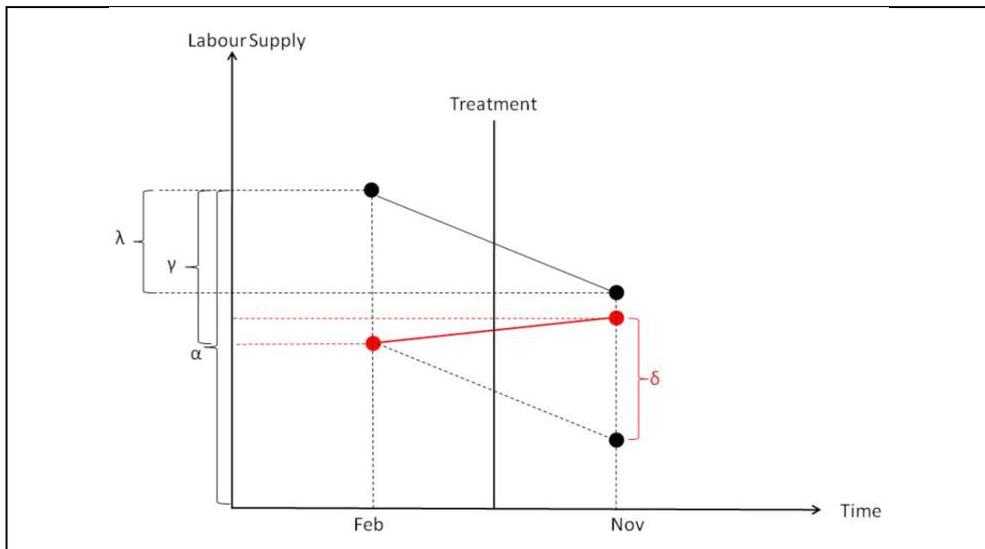
NJ Post: $\alpha + \gamma + \lambda + \delta$

$$(NJ \text{ Post} - NJ \text{ Pre}) - (PA \text{ Post} - PA \text{ Pre}) = \delta$$

[표 2] 뉴저지(NJ)와 펜실베니아(PA)의 최저임금 차이에 따른 고용률 변화

Variable	Stores by state		
	PA (i)	NJ (ii)	Difference, NJ - PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	-2.89 (1.44)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	-0.14 (1.07)
3. Change in mean FTE employment	-2.16 (1.25)	0.59 (0.54)	2.76 (1.36)

[그림 17] 뉴저지(NJ)와 펜실베니아(PA)의 고용률 변화



이를 그래프로 나타내면 [그림 17]과 같이 뉴저지(NJ)와 펜실베니아(PA)의 고용률은 각각 $\lambda + \delta$, λ 만큼 변화하고, 실제 뉴저지의 최저임금 인상의 효과 δ 를 확인하기 위해서는 뉴저지의 고용률 변화에서 펜실베니아의 고용률 변화를 차분하여 계산한다.

Ⅲ. 성과 감사 추진현황과 과제

1. 우리나라 감사원의 성과감사 추진 현황과 과제

가. 감사원의 감사 제도 운영 경과

우리나라 감사원은 해방이후 회계검사 및 결산검사를 수행하는 심계원과 직무감찰을 수행하는 감찰위원회 두 기관의 기능이 1962년 헌법 개정과 1963년 감사원법 제정으로 감사원으로 통합되면서 1964년 설립되었다.

이러한 감사원의 설립 배경 상 초기 감사원은 회계질서를 확립하고 공무원 비위를 척결하는 등 공무원의 기강 확립에 중점을 두었고, 규율의 준수 여부인 합법성 위주의 감사가 진행되었고, 1970년대에는 감사의 종류를 평면적이고 통상적인 정기감사에 해당하는 일반감사와 국민 생활에 밀접한 관련이 있는 기관에 대한 특별감사로 구분하여 감사를 실시하였다.

이후 1980년대를 거치면서 합법성 위주의 감사를 진행하면서도 역점감사, 총괄감사, 전문감사, 사전예고제 등 다양한 감사 접근 방법을 도입하여 감사업무를 효과적으로 수행할 수 있는 방법을 모색하였다. 이 시기에는 일반감사는 정기감사로, 특별감사는 특정감사와 계통감사로 구분되어 감사를 실시하였다.

1990 년대는 민주화의 영향으로 일정 자격을 갖춘 자의 공익적 목적을 위한 감사청구에 의하여 감사원이 감사를 실시하는 공익감사 청구제도가 도입되었고, 감사 업무의 독립성 확보를 감사의 기본 원칙으로 하였다. 이 시기에는 일반감사와 특정감사의 두 줄기를 유지하면서 1994 년 “예산편성 및 행정의 효과성, 능률성, 경제성을 제고하기 위한 성과감사” 를 지향하는 것을 감사원 운영의 기본 방향으로 수립하면서 성과감사라는 감사종류가 공식적으로 도입되었다.

2000 년대는 정보화가 급속도로 진행되었던 시기로 행정 분야도 전자정부를 구축하는 등 IT 기술을 적극 도입하였고, 감사원은 2004 부터 2006 년까지 e-감사시스템을 구축하여 감사업무를 전산 처리하는 체계를 갖추었다. 또한 정보화 물결과 함께 개방화를 겪으면서 정부혁신을 추구하였다. 감사원은 이러한 변화에서 국민의 신뢰성을 높이고자 2001 년 국민감사청구제도를 도입하고, 2003 년 국회감사청구제도를 도입할 뿐 아니라 2003 년부터 감사결과를 공개하였다. 또한 감사원은 운영기조는 부조리와 낭비의 요인이 되는 문제점과 제도의 개선을 촉구하고 행정의 투명성과 효율성을 높이는 생산적인 감사를 요구하는 등 기존의 합법성 감사를 지속하면서 행정의 효율성을 추구하는 감사가 점차 강조되고 있었다.

특히 전윤철 감사원장 시기에는 주요 프로그램 등에 대하여 상시 모니터링을 통하여 문제점을 체계적으로 수집하여 종합적이고 근본적인 개선책을 요구하는 시스템 감사를 강조하였다. 또한 2005 년 주요 정책 및

사업의 효율성 분석과 평가와 관련된 감사제도 및 방법 등을 조사 및 연구하는 평가연구원을 설립하였다.

아울러, 2004 년 대상기관의 업무 전반에 대하여 감사를 하던 일반감사를 폐지하고 헌법상 기본 책무인 결산검사와 회계검사에 집중하는 재무감사와 대상기관의 주요 업무를 집중 감사하는 기관운영감사를 도입하였고, 2007 년 특정과제감사의 명칭이 특정감사로 변경되면서 현재와 같이 재무감사, 기관운영감사, 특정감사, 성과감사 등 [표 3] 과 같이 4 개의 감사로 구성된 감사유형 체계가 완성되었다.

[표 3] 감사원의 감사유형

구분	정의 및 특징
재무감사	감사 대상기관의 재무적 자원 활용의 적정성을 분석 및 평가하기 위해 예산의 운용실태 및 회계처리의 적정성 여부 등에 대한 검토와 확인을 위주로 정기적으로 실시하는 감사
기관운영감사	기관 운영의 건전성과 효율성을 확보하기 위해 기관의 조직, 인사, 예산의 운영과 주기능 및 주임무를 진단·평가하는 등 대상기관의 주요 활동과 업무에 대한 적법·타당성을 점검하는 감사
특정감사	국민 및 국회 감사 청구사항 등 사회·경제적 현안과제와 같은 다양한 감사수요에 대응하여 원인 및 책임소재를 규명하고, 개선대책을 마련하기 위해 실시하는 감사
성과감사	주요 정책, 사업, 제도 및 업무시스템 등에 대한 문제점을 체계적으로 진단, 분석, 점검하여 경제성, 능률성, 효과성을 평가하고 종합적이고 근원적인 개선대책을 제시하기 위해 실시하는 감사

나. 감사원의 성과감사 운영 경과

앞에서 살펴본 것과 같이 감사원은 1994 년 성과감사를 도입 하였지만 성과감사에서 도모하고자 하는 정책 및 사업의 효율성, 효과성 개선을 위한 감사는 1980 년대부터 일부 반영되어 진행되고 있었다.

하지만 성과감사가 공식적으로 도입된 1994 년 이후에도 대부분 감사는 합법성 감사에 초점을 맞추고 진행되었다. 효율성, 효과성, 경제성을 종합적으로 분석하고 평가하는 감사는 활발하지 않았으나 2004 년 정부성과관리를 강조한 노무현 정부의 출범 이듬해인 2004 년, 정부업무 평가기본법의 제정 이듬해인 2007 년 성과감사의 집중 실시되었다. 또한, 2005 년 ‘성과감사 운영 매뉴얼’ 을 발간하여 성과감사를 수행하는데 필요한 기본 절차를 마련하고 2007 년 성과감사 실무지침을 작성하는 등 성과감사 활성화를 위한 노력을 지속하였다.

그러나 2004 년부터 2017 년까지 감사원에서 실시한 감사 중 성과감사의 비중을 살펴보면 [표 4]와 같이 2004 년과 2007 년을 제외하면 성과감사의 비중은 전체 감사의 6% 이내에 머무르고 있다. 하지만 성과감사의 실시 비중이 낮더라도 절반 정도의 비중을 차지하는 특정감사에서 합법성 준수여부와 함께 효율성 및 효과성을 제고하기 위한 방법을 점검하고 있어 성과감사의 기능을 수행하고 있다고 볼 수 있다.

[표 4] 감사 유형별 감사 실시현황(2004-2017)

년도	재무감사		기관운영감사		성과감사		특정감사		합계	
2004	55	(46%)			12	(10%)	52	(44%)	119	(100%)
2005	19	(26%)	10	(14%)	4	(6%)	39	(54%)	72	(100%)
2006	11	(16%)	15	(22%)	0	(0%)	43	(62%)	69	(100%)
2007	26	(19%)	34	(25%)	23	(17%)	56	(40%)	139	(100%)
2008	8	(5%)	110	(64%)	1	(1%)	52	(30%)	171	(100%)
2009	8	(6%)	62	(44%)	2	(1%)	69	(49%)	141	(100%)
2010	14	(9%)	31	(20%)	5	(3%)	106	(68%)	156	(100%)
2011	9	(7%)	31	(25%)	5	(4%)	77	(63%)	122	(100%)
2012	20	(14%)	31	(21%)	6	(4%)	88	(61%)	145	(100%)
2013	21	(17%)	5	(4%)	2	(2%)	97	(78%)	125	(100%)
2014	15	(11%)	34	(25%)	2	(2%)	85	(63%)	136	(100%)
2015	17	(12%)	25	(18%)	3	(2%)	96	(68%)	141	(100%)
2016	16	(12%)	40	(29%)	8	(6%)	72	(53%)	136	(100%)
2017	16	(10%)	31	(20%)	9	(6%)	101	(64%)	157	(100%)
합계	255	(14%)	459	(25%)	82	(5%)	1,033	(57%)	1,829	(100%)

자료: 감사원 감사연보(2004-2017년)

다. 감사원의 성과감사 운영 성과와 한계

최재해(2017)는 감사원의 역할이 ① 회계질서 및 공직기강 확립, ② 행정 및 재정운영의 개선 및 향상, ③ 주요 시책 및 사업의 성과 제고, ④ 정책실패에 대한 행정책임성 확보 등 4 가지로 크게 구분하였다. 이 기준에 따라 해방 이후 감사 유형별 감사 수행의 추세를 확인한 결과 초기에는 합법성 감사에 해당하는 ① 유형의 감사가 거의 대부분을 차지하였으나 1980년 이후 성과감사의 특징이라고 할 수 있는 ③ 유형의 감사가 증가하고 있다는 것을 확인하였다.

이러한 관점에서 윤기웅과 유승현(2018)은 성과감사는 감사원의 주요 시책 및 사업의 성과 제고에 대한 관심을 불러일으키고 실제 주요 시책과 사업의 성과 향상을 유도했고, 특히 성과 제고에 대한 감사원의

관심으로 주요 시책 또는 종합대책 시행에 따른 문제점을 조기에 발견하여 개선하는 데 기여하였다고 평가하고 있다.

또한 이렇게 감사원의 감사 초점이 합법성 감사에서 사업 및 정책의 효과, 효율성, 경제성 등 성과에 대한 평가까지 확장되면서 감사를 수감 받는 대상기관에서 행정의 책임 범위가 단순한 법률, 규정의 준수에서 보다 효율적이고 효과적인 사업 수행과 정책 수립 및 집행까지 확장된다고 볼 수 있겠다.

하지만 위와 같은 성과감사의 성과에도 불구하고 아직 성과감사가 성공적으로 정착되었다고 보기는 어렵다. 이러한 이유로 감사원에서도 주기적으로 성과감사 활성화를 위하여 각종 매뉴얼을 개정하거나 절차를 개발하는 등의 노력을 수행하고 있다.

공병천과 윤기웅(2017), 윤기웅과 유승현(2018) 등은 성과감사가 아직 다음과 같은 한계를 가지고 있다고 분석하였다.

첫째, 감사 실무자가 성과감사를 피하는 경향이 있어 감사 실시 빈도가 낮다. 둘째, 법규의 준수 여부를 확인하는 합법성 감사의 관행에 익숙하여 성과감사 기법의 활용을 어려워한다. 셋째, 주요 이해관계자들에게 필요한 정보를 효과적으로 제공하지 못하는 측면이 있다. 넷째, 정치적 논란 회피를 위해 명백한 잘못 또는 법규 위반사항의 적발에 초점을 맞추는 경향이 있어 거시적인 문제를 간과할 우려가 있다. 다섯째, 성과감사 결과의 기여도에 대한 점검 지표가 미흡하다.

정길영(2017), 공병천과 윤기웅(2017), 공병천, 유승현 및 윤기웅(2018)은 이와 같은 어려움에 대하여 성과감사를 기피하는 것은 실무자가 성과감사에 익숙하지 않은 데 기인하므로 감사 방법론의 개발과 교육을 통한 인식 제고가 필요하다고 제시하고 있다. 또한 익숙하지 않은 감사기법을 실시하는데 따른 위험에 대한 적절한 보상이 요구되는 한편, 주요 이해관계자들의 필요로 하는 성과감사 수요를 적극 발굴하려는 노력이 필요하며, 정치적 논란을 극복하기 위해서는 감사원의 독립성 확보가 필요한 등의 조치를 제시하고 있다.

성과감사 내용적 측면을 확인해보면 김형아(2019)는 감사원에서 2017년 수행된 9개의 성과감사를 분석하면서 감사 목적이 성과감사의 필수 요소라고 할 수 있는 효과성, 효율성 및 경제성을 반영하고 있는지 확인하였다. [표 5]와 같이 9개 감사 중 명시적으로 효과성, 효율성 및 경제성을 파악한 감사는 8건이었고, 특히 성과감사의 핵심이라 할 수 있는 효과성을 감사 목적으로 제시하고 있는 감사는 [표 6]과 같이 4건에 그쳤다고 확인하였다.

[표 5] 성과감사 목록과 감사 목적(2017)

2017년 성과감사 목록	감사 목적
교원양성 및 임용제도 운영실태	교원양성, 교원정원관리, 교원임용제도 운영과 사립학교의 교사 채용실태에 대한 종합적인 점검과 분석을 통해 개선방향 제시
노인의료 지원사업 추진실태	노인의료 지원사업의 성과를 점검하고 효율성 제고
농축산물 안전관리실태	농축산물 안전관리 실태에 대한 체계적 점검을 통해 효율적인 관리방안을 마련하여 농축산물 안전에 대한 소비자 신뢰를 제고
다문화가족정책 추진실태	다문화가족정책의 성과와 한계를 심층 분석, 진단하여 다문화가족정책의 실효성(효율성/효과성)을 제고하기 위한 방향제시
대형 연구개발사업 추진실태	대형 국가연구 개발사업의 효율성과 효과성 확보를 지원

일·가정 양립 지원제도 운영실태	일·가정 양립 지원사업의 성과를 점검하고 효과성 제고
주요 원자재 비축관리 실태	성과와 한계를 진단하고, 사업 운영의 효율성과 원자재 수급차질 등 위기 대응능력을 제고하기 위한 개선방향 제시
주택금융 위험요인 관리실태	주택금융 실태를 점검하여 향후 금리 인상 시 위험요인을 완화
취약계층 직접일자리사업 추진실태	직접일자리사업 전반에 대한 종합적/객관적 분석을 실시하고 확인된 문제점 개선을 유도하여 사업운영의 효율성과 성과 제고

[표 6] 효과성을 목적으로 한 성과감사 현황과 감사초점(2017)

성과감사 목록	감사 초점
노인의료 지원사업 추진실태	치매관리체계가 효과적으로 구축되어 있는가? 치매환자 관리를 위한 인프라는 적정하게 구축되어 있는가? 수요자 선호를 반영한 서비스가 제공되고 있는가? 장기요양기관에 대한 관리는 적정한가? 장기요양기관 등에 대한 시설안전관리는 적정한가?
다문화가족정책 추진실태	정부 지원의 사각지대를 분석·진단하여 제도 설계시 반영하고 있는가? 사업성과(효과성)를 제고하도록 관련제도가 운영·관리되고 있는가? 사업수요·목적에 맞게 지원사업이 효율적으로 운영되고 있는가? 지원사업이 사회·경제적 여건을 고려하여 형평성 있게 추진되는가?
대형 연구개발사업 추진실태	세부 연구과제가 사업단 목표와 부합하고 단계별 일관성을 유지하고 있는가? 사업단 차원의 목표 및 지표가 명확하게 설정·관리되고 있는가? 사업단 방식의 추진체계가 효율적으로 설계·운영되고 있는가? 연구과제 평가가 공정하게 수행되고 평가 후속조치가 제대로 이행되고 있는가? 21세기 프론티어 사업의 연구성과 활용실태가 정확히 파악되고 있는가? 21세기 프론티어 사업의 성과르 활용하기 위한 사후관리가 적정한가?
일·가정 양립 지원제도 운영실태	지원제도 설계 시 취약분야 등 지원대상에 대한 검토가 적정한가? 지원사업이 적정하게 집행되고 있는가? 성과·평가·환류 등 성과관리는 적정하게 이루어지고 있는가? 지도·감독은 적정한 대상을 선정하여 이루어지고 있는가?

또한 김형아(2019)는 효과성 확인을 목적으로 하는 4 건 감사 사항의 경우도 해당 프로그램이 어떠한 효과를 보였는지 평가하기 보다는 ‘효과를 달성하기 위하여 프로그램 운영이 적절히 수행되고 있는지’, ‘효과적인 사업 수행을 위한 목표와 지표의 설정과 관리가 제대로 되고 있는지’ 등 주로 문제점을 확인하는 데 그치고 있고, 감사를 수행하는 방법도 계량적인 분석 방법을 적용하기 보다는 통계 자료를 작성하거나 자료를 비교하는 정도로 진행되었다고 평가하였다.

정부의 프로그램은 효율적이고 경제적으로 추진되어야 하지만 공공부문의 존재 이유가 공공성에 있음을 고려하면 해당 시책이 효과적으로 추진되는 것이 가장 중요하다. 이러한 측면에서 성과감사의 목적인 효과성, 효율성, 경제성 중에서 효과성을 측정하여 평가하는 것이 중요하다고 볼 수 있다.

이에 성과감사를 적극적으로 수행하고 있는 미국 최고 감사기구 GAO의 성과감사 추진 현황을 살펴보고, 현재의 데이터 분석기술의 발전으로 효과성 평가를 중심으로 한 성과감사의 발전 방향을 살펴보겠다.

2. 미 GAO의 성과감사 추진실태

가. 미 GAO 조직 및 임무

미국의 최고 감사 기구인 GAO(Government Accountability Office)는 연방정부 입법부에 소속된 독립적이고 초당적인 전문 기관이다. 1921년 예산회계법(the Budget and Accounting Act of 1921)에 따라 설립되어 연방정부에 대한 회계감사, 평가 및 조사 서비스를 의회에 제공하며, 연방정부의 성과개선과 책임성 확보를 통해 의회가 헌법상 책무를 수행하는 것을 지원한다.

GAO의 인력규모는 1970년대 후반에는 5천여명 수준까지 증가하였으나 조직의 슬림화를 거치면서 인력이 감소되었으며, 2019년 기준 워싱턴 D.C.에 위치한 본부와 미국 전역에 위치한 11개 현장 사무소에 총 3천여 명의 인력이 근무하고 있다.

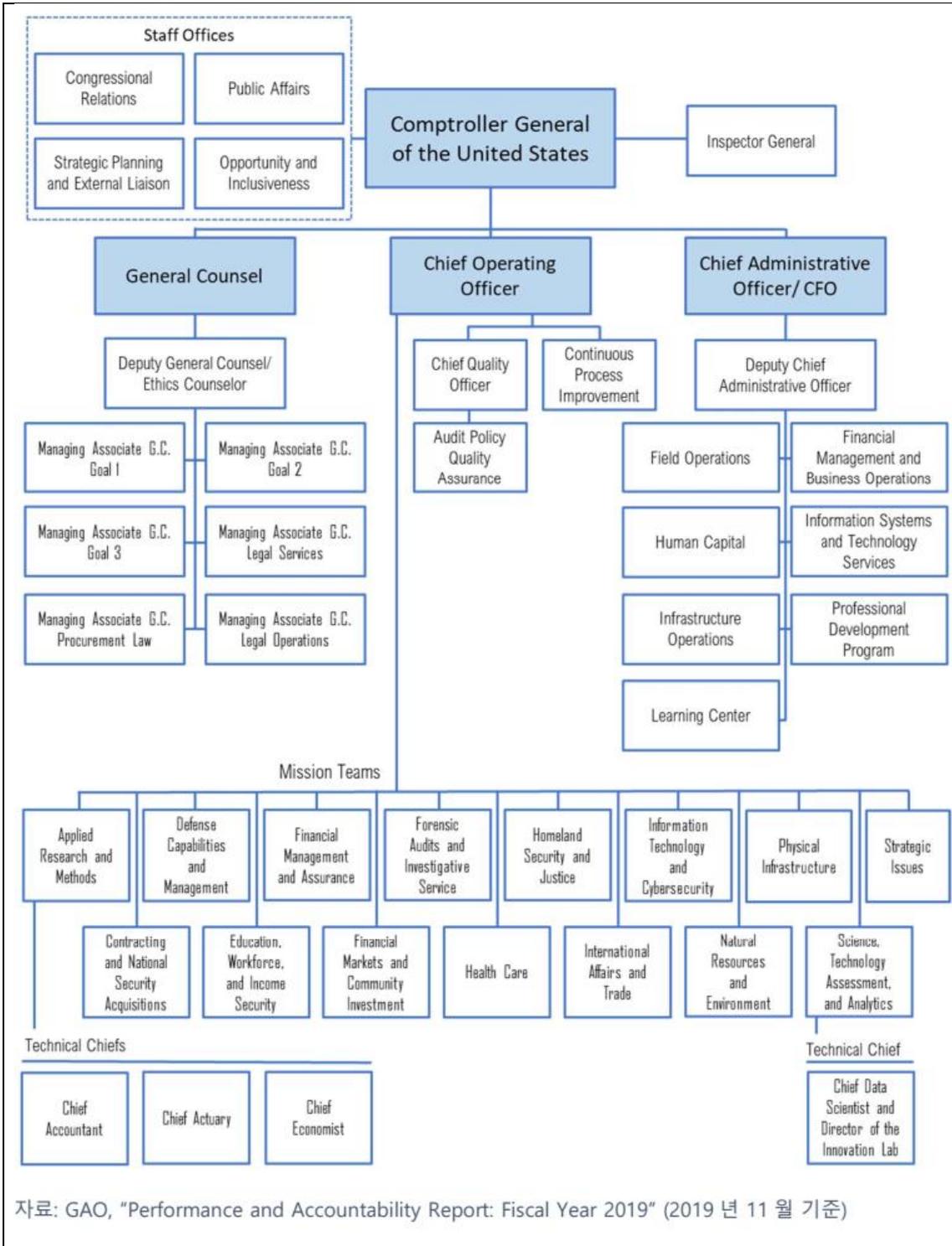
구체적인 구성은 [그림 18]과 같이 감사원장(Comptroller General) 아래 사무총장(Chief Operating Officer), 법무실장(General Counsel), 행정재무실장(Chief Administrative Officer/CFO)을 두고, 감사원장 직속으로 홍보, 의회 대응, 전략계획 및 대외관계, 공정한 근무환경 담당 참모 부서와 감찰실을 운영하고 있다.

사무총장 아래에는 국방, 재무, 정보보안 등 12 개 분야별 감사팀, 포렌식팀, 전문가 지원 부서인 응용 연구·방법팀(Applied Research and Method: ARM)과 과학기술 평가·분석팀(Science Technology Assessment and Analytics) 등 15 개 임무팀(Mission Teams)과 감사품질실, 절차개선실을 운영하고 있다.

또한 법무실은 법률전문가들에 의한 입찰 이의신청, 의회 및 연방정부에 대한 세출관련 법률지원, 계약 검토 등 기타 법률지원 부서로 구성되어 있으며, 행정재무실은 현장운영, 인적자본, 인프라 운영, 정보시스템, 재무 관리, 전문성 개발팀 및 교육센터로 구성되어 있다.

GAO 는 의회의 위원회 또는 소위원회의 요청, 법률 또는 위원회 보고서에서 따라 강제되거나 감사원장의 지시에 의하여 감사가 진행하며, 그 대상은 정부 기관과 그 사업, 활동 및 기능 뿐 아니라 정부와 계약한 개인, 기관, 비영리기관 및 기타 비정부기관 까지 포함한다.

[그림 18] 미국 GAO 조직도



자료: GAO, "Performance and Accountability Report: Fiscal Year 2019" (2019년 11월 기준)

GAO 의 임무를 구체적으로 살펴보면, 연방 자금이 효율적이고 효과적으로 사용되는지 확인하기 위한 대상 기관의 운영 전반에 대한 감사와 불법 또는 부적절한 행위의 혐의에 대한 조사를 수행한다. 또한 정부의 사업 또는 정책이 그 목적에 얼마나 부합하는지 보고하고, 정책 분석을 수행하여 의회에서 고려할 대안에 대한 아웃라인을 제시한다. 아울러 의회나 행정부 기관장에게 정부를 보다 효율적이고 효과적으로 만들기 위한 방안을 제시하며, 정부 입찰이나 계약 관계에 대한 이의신청 등 법률적인 결정이나 의견을 제시한다.

나. GAO 의 감사 유형

The Yellow Book 이라고 알려진 GAO 의 정부 감사 기준(2018)에 따르면 GAO 가 수행하는 감사 유형은 재무감사(Financial Audit), 증명업무 및 재무제표 검증(Attestation Engagements and Reviews of Financial Statements), 성과감사(Performance Audits)로 구분된다.

재무감사는 기관의 재무 상태가 재무보고 형식에 따라 공정하게 작성되었는지 확인하는 것으로 재무보고에 대한 내부 통제, 재무제표에 영향을 미치는 법률, 규정, 계약 및 보조금 조항의 준수 여부를 검토한다.

증명업무 및 재무제표 검증은 사용자 요구에 따른 대상의 문제 또는 주장(assertion)에 대한 광범위한 재무 또는 비재무 목표에 대하여 해당 주제 또는 주장을 적절한 기준에 따라 측정하거나 평가한다.

성과감사(Performance Audits)는 객관적인 분석과, 조사결과 및 결론을 제공하여 감사 대상기관의 기관장 또는 관리 및 감독권자에게

프로그램의 성과를 제고하고, 운영을 개선하며 비용을 절감하도록 시정 조치를 하는 의사결정을 촉진하고 공공 책임성에 기여하도록 하는 감사로 정의하고 있다. 성과감사의 목적은 다양하지만 일반적으로 사업 또는 정책의 효과성, 효율성 및 경제성 (Effectiveness, Efficiency, economy: 3E) 평가, 내부통제 (internal control), 규정의 준수 여부 (compliance), 전망적 분석 (prospective analyses) 을 포함한다.

GAO(2018)는 사업 및 정책의 효과성, 효율성, 경제성 평가는 사업의 목표 달성 정도 (effectiveness)를 측정하며, 경제성 및 효율성 평가를 통해 사업 목표 달성을 위한 비용을 분석하는 것으로 규정한다. 또한 내부통제는 기관을 효율적·효과적인 운영하고, 보고의 신뢰성을 확보하며, 규정 준수를 보장하도록 설계된 기관의 내부 통제 시스템에 대한 평가를 의미한다. 규정의 준수여부는 대상기관이 법률, 규정, 계약 및 보조금 협정 및 기관의 자산획득/보호/사용/처분과 기관의 서비스 품질/적시성/원가에 영향을 미치는 기타 요구사항에 의해 확립된 기준을 준수하는지 여부에 대한 평가로 합법성 감사에 해당한다. 마지막으로 전망적 분석은 미래에 발생 가능한 사건에 기반하여 기관의 대응 가능한 행동에 대한 분석과 결과를 제공하는 감사를 의미한다. 위 4개 유형의 성과감사 목적에 대한 구체적인 내용은 [표 7]과 같이 구성된다.

[표 7] 성과감사의 유형별 주요 감사 내용

구 분	주요 내용
효과성, 효율성, 경제성 평가	<ul style="list-style-type: none"> • 입법, 규정 또는 조직의 목표 달성 정도 평가 • 사업의 성과 개선 또는 효과를 저해하는 요소의 제거를 위한 대안적 접근 방법의 상대적 능력 평가

	<ul style="list-style-type: none"> • 성과 또는 혜택에 기반한 사업의 상대적인 비용/효과 분석 • 사업의 성과가 당초 목표와 일치하는지 여부 판단 • 사업이 다른 사업과 중복되거나 충돌하는 정도 평가 • 해당 기관이 건전한 조달 관행을 따르고 있는지 여부 평가 • 사업의 효과, 결과 또는 경제성 및 효율성과 관련된 성과 측정과 재무 정보에 대한 신뢰성, 유효성 및 관련성 평가 • 정부의 서비스와 혜택이 적정 수혜자에게 접근 가능한지 여부 • 사업의 단위 비용 감소 및 생산성 제고 가능성 또는 방법 평가
내부통제	<ul style="list-style-type: none"> • 조직의 임무, 목표 및 목적의 효과적, 효율적인 달성 여부 • 자원 사용의 법률, 규정 및 기타 요구사항에 대한 준수 여부 • 민감 정보를 포함한 자원의 무단 취득/사용/처분에 대한 보호 • 전산화된 시스템에서 확보한 정보의 신뢰성 • 성과 측정 같은 관리 정보의 완전성, 정확성 및 일관성 • 기능 중단을 방지하기 위한 정보시스템의 비상 계획
규정의 준수여부	<ul style="list-style-type: none"> • 사업의 목적, 수행 방식, 제공되는 서비스, 결과 또는 수혜 대상이 법, 규정, 계약, 보조금 협정 및 기타 요구사항에 부합하는지 여부 • 정부의 서비스와 혜택을 받는 대상의 자격이 적정한지 여부 • 발생 또는 제안된 비용, 또는 확보된 수입은 관련 법률, 규정, 계약 또는 보조금 협정을 준수하는지 여부
전망적 분석	<ul style="list-style-type: none"> • 정부의 사업 및 서비스에 대한 현재/향후 동향과 미래의 잠재적인 영향 및 정책 대안에 대한 영향 • 다양한 가정에 따른 사업 결과 예측을 포함한 사업 및 정책 대안 • 이해관계자의 견해 분석, 장단점을 포함한 정책 또는 입법 제안 • 예측된 미래 사건과 이에 대한 이해관계자, 관리자의 예측된 행동 등을 가정한 예산 및 전망 등

이와 같이 GAO의 성과감사에 대한 정의, 목적 및 주요 내용을 살펴보면 우리나라 감사원의 감사 유형에서 기관운영감사, 특정감사,

성과감사로 정의될 수 있는 감사를 모두 성과감사라는 틀에서 정의하고 있다.

또한 우리나라 감사원에서 성과감사로 분류하고 있는 사업의 효과성, 효율성, 경제성 평가 목적의 경우 감사 초점의 항목을 확인하면 “사업이 다른 사업과 중복 또는 충돌되는지 여부”, “건전한 조달 관행의 준수 여부” 등 특정감사의 감사 초점으로 활용되고 있는 항목도 있다. 이와 같이 국내 감사원의 성과감사는 특정감사와 경계가 분명하지 아니한 측면이 있다.

다. GAO 성과감사의 특징 - 정책 및 사업의 ‘효과성’ 평가

GAO(2012)에 의하면 프로그램에 대한 평가(Evaluation)는 연구 방법론을 사용하여 데이터를 수집하여 분석하고 해당 프로그램이 얼마나 잘 작동하고 있고 그 이유는 무엇인지에 대하여 평가하는 체계적인 연구이다. 평가는 프로그램의 성과(Performance)에 대한 특정 질문에 답변하거나 사업의 운영 또는 결과의 분석에만 초점을 맞출 수 있다. 또한 평가 결과는 프로그램의 효과를 평가하고 어떻게 성과를 개선할 수 있을지 분석하며, 행정 자원의 할당에 기준점으로 사용될 수 있다.

평가를 실시할 때에는 평가의 품질, 신뢰도 및 활용성을 높이고 평가에 필요한 시간과 자원을 효과적으로 사용하기 위하여 사전 계획에 시간을 들여야 한다. GAO(2012)는 자료의 수집 이전부터 다음 다섯 단계의 절차를 거치면서 평가에 대한 설계를 하도록 제안하였다.

- ① 프로그램의 목표와 전략을 명확하게 이해한다.
- ② 관련성 있고 유용한 평가 질문들을 개발한다.
- ③ 각 평가 질문들에 대한 적절한 평가 접근법 또는 설계를 선택한다
- ④ 신뢰할 수 있는 관련 정보를 수집하기 위하여 데이터의 출처와 수집 절차를 식별한다.
- ⑤ 유용한 결론을 내릴 수 있도록 데이터를 분석하는 계획을 마련한다.

평가에 대한 설계는 프로그램의 목표와 전략을 명확하게 하는 것으로 시작한다. 프로그램이 실행될 때에는 보통 그 목적, 범위 그리고 목표가 무엇인지 제시되므로, 당시의 관련 문서를 확인하거나 담당 부서 공무원이나 이해관계자에게 목적과 목표에 대한 문의하는 방법으로 확인할 수 있다.

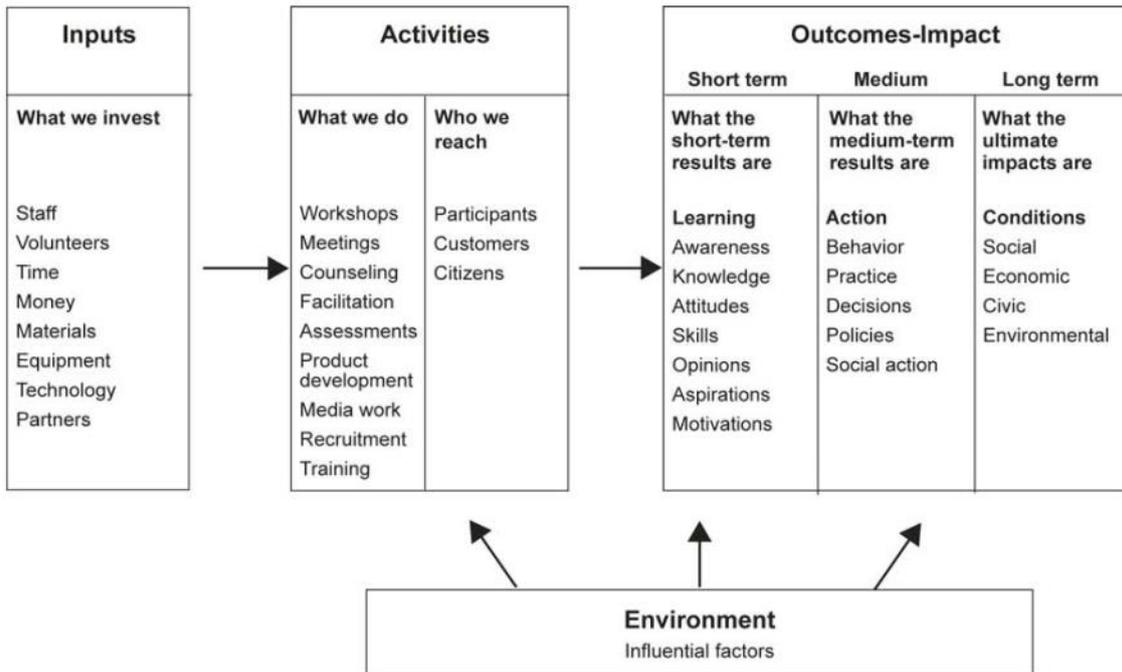
또한 [그림 19]와 같은 논리 모델을 이용하여 프로그램에 투입되는 자원, 활동과 수혜 대상, 결과와 영향 및 외부요인 등으로 프로그램의 흐름을 구조화하는 것이 도움이 된다.

이렇게 프로그램의 목표와 전략을 확인할 때에는 입법 과정의 역사, 담당 직원이나 관련 문서, 선행 연구나 홍보 자료 등 다양한 자원을 활용하는 것이 중요하다. 또한 당초 시행되었을 때의 정책적인 의도를 이해하고, 이후 중요한 외부 여건의 변화가 발생하였는지, 현대의 정책적인 고려가 무엇인지 확인하는 것이 중요하다.

평가 질문들을 개발할 때에는 프로그램의 이해관계자들이 프로그램 성과에 대하여 가지는 관심사항이나 우려하는 상황이 명확하게 설명되고 평가의 결과물이 유용하다는 것을 보장할 수 있도록 구성되어야 한다.

질문들은 평가 범위에 대한 틀을 구성하고 수집할 데이터와 비교할 대상이 설정되는 등의 평가의 설계를 유도하므로 적절하고 유용한 질문의 설정은 매우 중요하다.

[그림 19] 사업 또는 정책의 논리 모델에 대한 개념도



자료: GAO

프로그램 평가는 프로그램의 개선 방안을 마련하는 것을 지원하거나 프로그램에 대한 책임성 (Accountability) 을 확보하는 등 다양한 목적으로 요구된다. 이러한 프로그램의 평가 요청 목적에 따라서 평가의 초점은 달라질 수 밖에 없다. 예를 들면, 평가를 요청한 목적이 프로그램의 자원 할당에 대한 기준점을 마련하고자 하는 것이라면 질문은 어떤 프로그램 참여자가 가장 서비스를 많이 필요로 하는지, 어떠한 프로그램 활동이 원하는 결과를 달성하는데 가장 효과적인지 등에 대하여 확인하여야 한다.

의회에서 요청은 보통 매우 광범위한 범위의 관심 사항에 대하여 요청하므로 사전에 연구 가능한 질문들을 명확히 하고 집중시키는 것이 필요하다.

또한 프로그램의 진행 상태 또는 성숙도가 어떠한 지에 대하여 고려하는 것이 필요하다. 프로그램의 시행 초기 또는 막 시작하는 단계에서는 절차에 대한 모니터링이나 절차를 평가하고, 프로그램이 이미 안정화되고 성숙된 단계가 되었을 때는 결과에 대한 모니터링이나 평가, 영향에 대한 평가 등을 고려하여야 한다.

앞에서 식별된 질문에 대하여 주어진 시간과 자원을 가지고 신뢰할 만한 증거를 제공할 수 있도록 평가를 설계하여야 한다. 이러한 평가 설계의 기본 구성요소는 평가 질문, 목적 및 범위, 데이터 수집 방법 및 샘플링 절차, 분석에 대한 계획과 한계점 분석 등을 포함한다.

GAO는 평가 설계의 구성요소를 작성하기 위하여 설계 매트릭스 (design matrix)라고 하는 표준화된 도구를 이용한다. 디자인 매트릭스는 [표 8]과 같이 문제, 설계의 선택지, 평가자가 고려할 균형에 대한 입증할 수 있도록 구성되며, 이를 통해 이해관계자가 평가의 논리를 이해할 수 있도록 돕는다.

[표 8] 설계 요소의 선정을 위한 질문 안내

구 분	질 문	고려사항
조사 가능한 질문	어느 질문에 답변하는가?	질문은 구체적, 객관적, 중립적이고, 측정 가능하며, 답변 가능하여야 함

필요한 정보 및 출처	각 평가 문제 해결을 위해서 어떤 정보가 필요하고, 어떻게 확보할 수 있는가?	반드시 확보하여야 할 정보와 문서, 해당 정보 수집 계획, 출처 등 식별
범위와 방법론	어떻게 각각의 평가 질문에 답을 할 것인가?	필요한 정보와 데이터 수집(샘플링, 사례연구, 벤치마크 등) 전략, 전략별 계획된 범위(시간, 장소 등), 분석 기술(비용-편익분석, 민감도 분석, 회귀분석 등) 등을 명시
한계	설계의 한계는 무엇이고 평가 결과에 어떤 영향을 주는가?	데이터의 품질이나 신뢰성, 확보 불가능한 데이터, 비밀 분류 등 정보의 한계를 명시
예측되는 결과	이러한 분석으로 GAO 는 무엇을 말할 수 있을까?	평가 목적을 고려하여 예상하는 답변을 미리 작성

설계 요소를 선정하였으면 선행 연구나 관련 문헌에 대한 조사가 이루어 진다. 이 경우 평가자는 새로운 데이터를 수집하기 전에 기존 연구 결과의 조합을 통해 전체적인 프로그램의 효과나 프로그램 특별히 잘 또는 잘못 작동하는 특징을 살펴본다. 선행 연구 및 문헌 조사에는 다양한 통계 접근이 가능한데, 많이 사용되는 방법은 여러 연구의 효과의 크기를 분석하는 메타 분석(meta-analysis)이다.

프로그램이나 연구 질문에 따라서 평가 질문에 대한 증거를 확보할 수 있는 자료의 원천은 프로그램에 대한 행정 기록, 보조금 보고서, 성과 모니터링 데이터, 사업 참여자에 대한 설문 조사 및 기존 조사된 설문조사 결과 등 다양하다.

각 자료의 원천을 선택하는데 있어서 평가자는 해당 자료 원천이 평가 질문에 대한 답변과 결론을 도출하는 증거로 충분(sufficient)하고

적정 (appropriate)한 것인지에 대하여 고민하여야 한다. 충분하다는 것은 증거의 물량을 의미하고 적정하다는 것은 관련성, 입증 가능성 및 신뢰성을 의미한다.

평가 기준은 실제 성능 측정을 비교하고 평가하는 데 필요한 기준, 측정 또는 기대치이다. 따라서 평가 기준은 평가의 목적 달성을 위해 관련성 있고, 적절하며, 충분하여야 한다. 평가 기준으로 법이나 규정에 언급된 목적이나 목표, 담당 공무원에 의해 설정된 정책이나 절차, 전문적인 기준이나 규범, 전문가 의견, 선행 평가, 다른 기관에서 벤치마크 용으로 사용한 성과 기준 등을 활용하여 설정할 수 있다.

성과를 측정하는 몇몇 경우에는 모든 대상을 포함하여 연구할 수 있으나 보통은 자원이나 시간의 한계가 있어 샘플링을 하여 조사를 수행한다. 샘플링 방법은 통계 기법을 활용한 랜덤 샘플링(확률적 샘플)과 비확률 샘플링으로 구분되는데, 랜덤 샘플링이 대표성을 확보하기 쉽지만 실제 조사의 시간적 및 공간적 한계상 비확률 샘플링이 될 경우가 많다.

주요 데이터 수집이 진행되는 동안 전체 데이터를 수집하기 이전에 사전 테스트나 파일럿 연구를 하는 것은 중요하다. 이런 사전 테스트나 파일럿 연구를 통해 평가 설계를 정교하게 할 수 있고 수집할 데이터의 적절함, 수집 가능성, 현실성 등을 평가할 수 있다.

프로그램 실행 단계에 대한 평가는 프로그램이 의도한 대로 이행되는지, 실행 단계의 관리 문제 등은 없는지, 프로그램이 예상한 성과를 보여주지 못하는 이유는 무엇인지 등에 대한 답변을 찾는 과정이다.

반면 프로그램 결과 단계에서 평가는 프로그램이 그 목표를 어느 정도 달성하였는지에 대한 질문이다. 보다 구체적으로 프로그램은 의도한 목표와 목적을 달성하였는데, 이해관계자가 관심 가지는 중요한 부작용은 없는지, 결과가 프로그램의 접근, 구성요소, 프로그램 제공자 및 프로그램 수혜자에 따라서 달라지는 않는 지 등에 대한 답을 찾는 것이다.

정부 프로그램의 결과는 다른 연방, 주 및 지방 정부의 사업이나 정책, 경제 상황과 제반 여건 등 외부 요인에 영향을 받는다. 그래서 관찰된 결과는 다양한 영향들이 결합된 결과이다. 따라서 프로그램에 의한 고유 효과만 측정하기 위해서는 효과 연구가 신중하게 설계되어야 한다.

효과 평가(Impact evaluation)는 해당 프로그램이 없었다면 발생하였을 결과와 해당 프로그램이 있을 때의 결과를 비교하는 것으로 다양한 방법론이 있다. 직접 실험법, 유사 실험법(quasi-experimental), 비실험 설계로 크게 구분된다.

무작위 실험(Randomized Experiments)은 처치(Treatment or intervention) 그룹과 통제(Control or comparison)그룹을 무작위로 선정하여 실험을 하는 것으로 두 그룹 간의 차이점이 결과의 차이가 되며, 프로그램의 영향으로 볼 수 있다. 그러나 평가자가 프로그램 참여자에 의한 절차를 통제해야 가능한 것이므로 GAO에 의한 진정한 실험은 거의 드물다.

무작위 실험은 실제로 수행이 어렵기 때문에 유사 실험(Quasi-experiment) 기법이 많이 이용된다. 유사 실험은 처치 그룹과 하나 이상의 통제 그룹으로 나누어지는 것은 동일하지만 이 둘이 무작위로 선택되지는

않는다. 따라서 이러한 선택에 편차가 발생하여 결과에 영향을 미칠 수 밖에 없기 때문에 통계적 조정을 거쳐 이를 해소한다.

이러한 통계적 기법으로 회귀단절모형 (Regression Discontinuity Design) 분석이나 이중차분분석 (Difference-in-Difference Analysis) 등이 이용된다.

GAO 는 프로그램의 효과성, 효율성, 경제성을 평가하면 비용 및 효과 분석, 대안의 비교, 효과 평가 등 경제 분석을 수반한다. 경제 분석 (Economic Analysis)이란 결국 의사결정자와 이해관계자에게 법률이나 규정 제정, 프로그램의 시행 등 조치에 대한 경제적 효과를 설명하기 위한 분석이며, GAO(2018)는 경제 분석 결과를 의사결정자와 이해관계자에게 제시하기 위한 기본 요소를 발굴하여 경제 분석 보고를 위한 기본 구조로 활용하고 있다.

경제 분석의 기본 요소는 경제분석의 구조와 경제적 개념과 관련하여 제안되었고 [표 9]와 같이 ① 분석의 목적과 범위 (Objective and scope), ② 경제 효과를 조사하는데 사용된 방법론 (Methodology), ③ 경제효과 분석 (Analysis of effects), ④ 경제효과 분석의 투명성 (Transparency), ⑤ 분석의 문서화 (Documentation)로 구성된다. 이와 같이 경제 분석에서는 경제 효과를 분석하는 방법론과 분석을 실시하는 것도 중요하지만 분석의 목적을 분명히 하면서 그 목적을 달성하기 위한 범위를 설정하고, 경제 분석 결과의 신뢰성을 확보하기 위하여 데이터 소스에 대한 설명, 분석에 이용된 가정사항, 주요 가정사항이나 조건이 변경될 경우 결과의 변화 등을 반영하여 투명성을 높이는 전략이 필요하다.

[표 9] GAO 경제 분석 보고의 기본 요소

기본요소	경제적 개념
목적과 범위	경제 분석은 분석 대상이 되는 프로그램에 당위성과 정당성을 포함하여 설명하고, 분석의 목적을 언급함. 분석의 범위는 이러한 목적을 해결하도록 설계되어야 함. 특별한 요구가 없는 한, 분석은 미국 시민과 거주자에 영향을 미치는 효과에 집중하여야 하고, 분석의 시간 범위는 해당 조치의 주요 영향이 충분히 작용할 만큼 길어야 함
방법론	경제 분석은 대안 중 하나를 기준으로 하여 대안들과 비교를 통해 프로그램의 영향을 조사하며, 대안들은 아무런 조치가 없는 것을 포함한 모든 관련 대안을 고려함. 분석은 적절한 기준을 정의하는 것으로 시작하고, 각 분석 대안 별 평가는 대안에 따른 중요한 경제적 효과, 발생 시기, 해당 영향이 직접적인지 또는 부수적인지를 식별함
효과 분석	경제 분석은 가능하면 중요한 경제 효과를 정량화하고 기회 비용 개념을 사용하여 이익을 산출하여야 함. 또한 효과 또는 비용은 순현재가치 기준(net present value) 또는 관련한 결과 측정 기준을 적용하여 대안들 전반에 걸쳐 적용한 후 효과를 비교하며, 인플레이션을 통제하고 경제적으로 정당화된 할인율을 적용. 한편, 중요한 경제적 효과를 정량화 할 수 없는 경우 대안 비교에 미치는 영향을 설명함.

투명성	경제 분석은 사용된 분석 방법, 가정, 분석에 이용된 데이터를 설명하여 정당화함. 분석은 각각의 중요한 분석 방법과 가정 사항이 적절한 범위에서 변경되는 경우 경제 효과의 추정치와 대안의 비교 결과에 어떤 영향을 미치는지 평가함. 또한 분석에 이용된 데이터에 대한 주요 한계점과 이에 따른 영향을 설명함. 가능한 경우 경제 분석의 추정치에 기초한 주요 데이터 요소의 통계적 변동성이 추정치와 대안 비교의 결과에 어떤 영향을 미치는지 적절한 계량화가 필요.
문서화	경제 분석 결과는 명확하게 작성되어야 함. 사용된 데이터를 설명하는 명확한 레이블이 지정된 표 및 이러한 결과와 일치하는 결론으로 명확하게 작성되어야 함. 분석에 사용된 모든 출처와 이용 가능한 최상의 경제 정보를 기반으로 한다는 것을 인용하며, 외부 컨설턴트와 용역업체의 이용과 그들의 기여를 공개함

미국은 감사 결과를 예산의 배정과 밀접하게 연관시키기 때문에 GAO 이외에 다른 기구에서도 정부 프로그램의 효과성을 확인하는 활동이 활발하다. Haskins 와 Margolis(2015)에 따르면 오바마 정부는 근거에 기반한 정책(evidence-based policy)을 강조하여 근거에 따라 예산을 제공하였고, 이러한 정책기조로 GAO는 효과성을 평가하는 성과감사를 적극적으로 시행한 것으로 보인다.

김형아(2019)에 따르면 GAO는 정책의 효과성에 기반한 예산 배분을 담보하기 위하여 정책 및 사업에 대한 효과성을 다양한 관점에서

평가하고 객관적이고 과학적인 근거를 확보하고자 노력하였다. 또한 다양한 효과성의 평가 방법론을 적용하여 감사 결과 보고서를 작성하였다.

공통적인 방법으로는 사회과학에서 특정 행동에 따른 결과를 추정하는데 이용되는 무작위 통제 방법(Randomized Controlled Trial : RCT) 또는 무작위 실험(randomized experiment)방법을 이용하거나 유사 실험 방법(Quasi-experimental study)을 적용하고 있다.

이러한 방법론은 사회과학 분야에서 이론을 수립하거나 경험론적 증거를 확보하기 위하여 널리 이용되는 분석 방법론이나 우리 나라 감사원의 성과 감사에서는 거의 이용되지 않고 있는 실정이다.

미 GAO 은 무작위 실험 방법은 정책의 효과성을 확인할 수 있는 가장 강력한 도구로 인정하고 있다. 하지만 자연과학인 아닌 사회과학 분야에서 무작위 실험을 수행한다는 것은 분석 집단의 표본 크기가 제한적일 수밖에 없어 표본 추출 과정의 편차가 발생할 우려가 높고, 실험군과 통제군으로 구분하여 실험을 실시하여 장기간 추적하기에는 많은 시간과 자원이 소요되며, 실험군과 통제군으로 구분하는데 대하여 윤리적 또는 도덕적인 문제나 해당 실험에 참여하는 국민들에 대한 행정서비스의 형평성 문제로 실험 실시하는 것 자체에 많은 어려움이 있다.

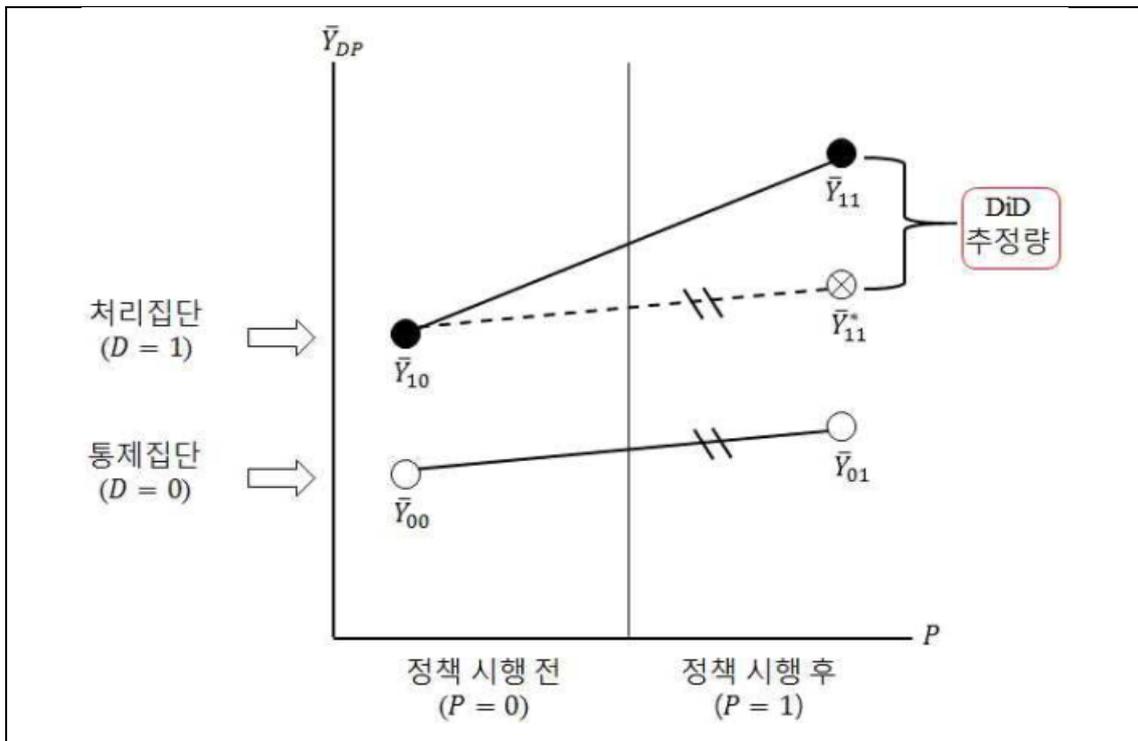
특히 성과감사는 정책 및 사업의 수행 주체가 아닌 외부 기관에서 비교적 짧은 시간내 수행하기 때문에 위와 같은 무작위 실험 방법을 적용하기는 불가능하다. 따라서 무작위 배정이 아닌 상황에서 정책의 효과성을 파악할 수 있는 수단이 필요하다.

이러한 수단으로 GAO(2012)는 유사 실험기법을 대안으로 제시하고 있고, 이중차분분석(Difference-in-Difference) 기법과 회귀

단절설계(Regression Discontinuity Design) 기법 같은 방법을 제안하였고, 손호성과 이재훈(2018) 및 한승훈과 안혜선(2019)은 각각 이중차분분석 기법과 회귀단절설계 기법을 이용한 정책효과의 추정이나 평가 방법을 제안하였다.

앞에서 살펴보았지만, 이중차분분석은 [그림 20]과 같이 정책 시행의 효과는 정책의 처치를 받은 그룹의 결과와 처치를 받지 않았다면 나타났을 결과의 차이가 되며, 이를 위해서 정책의 처치(Treatment)를 받는 그룹의 정책 시행 전·후의 결과 변화와 통제 그룹(Control)의 정책 시행 전·후의 변화의 차이를 계산하면 DiD 추정량이 나온다.

[그림 20] 이중차분분석의 개념도



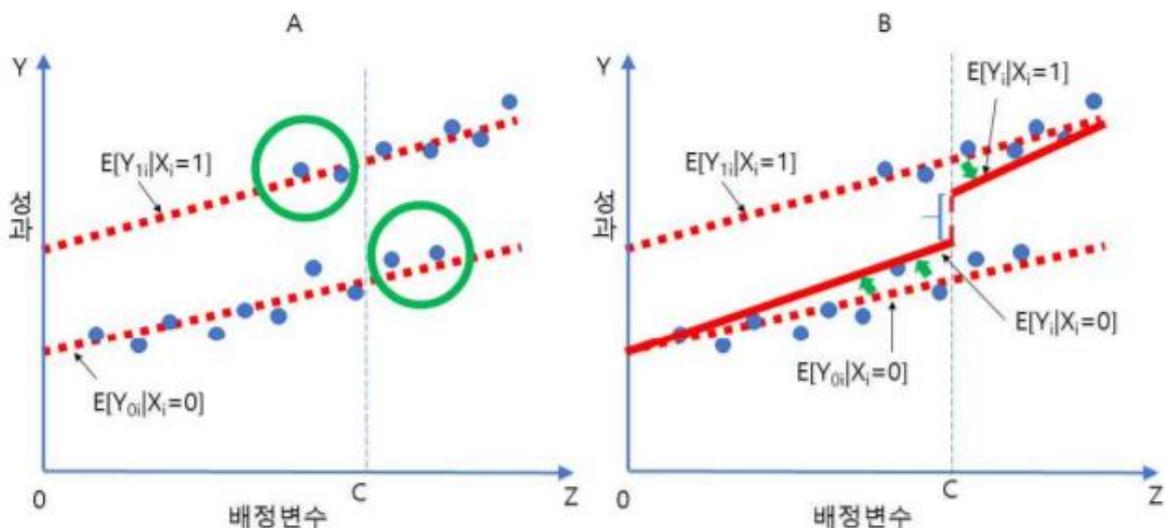
자료: 손호성과 이재훈(2018)

성과감사에서 이중차분분석을 적용하기 위해서는 해당 프로그램의 수혜 대상과 수혜를 받지 못하는 그룹을 구분하여 확인할 수 있는데, 프로그램의 대상이 특정 지역, 소득 구간, 연령 등을 대상으로 할 때 나머지 통제 그룹과의 차이를 정책 시행 전후의 비교를 통해 분석할 수 있다.

하지만 위와 같은 계산에는 [그림 20]에서 나타난 것과 같이 처치 그룹과 통제 그룹 모두 동일한 트렌드를 가지고 움직인다는 조건이 선행되어야 한다.

한편 회귀단절설계 기법은 정책의 시행 시기를 기준으로 시행 이전과 이후의 추세를 회귀분석으로 계산한 후 [그림 21]과 같이 두 회귀식 상이에 단절이 발생하는지 확인하는 방법으로 효과를 판단하는 것이다. 이 경우 주의해야 할 사항은 실제로 비선형 연결된 회귀식을 단절된 것으로 잘못 확인하거나 해당 단절이 다른 요인으로 발생한 경우를 검증해야 하는데, 보통 배정된 변수를 바꾸었을 때 단절되지 않았던 것이 정책의 시행 시점에 해당하는 값에서는 단절된다는 것을 확인하는 방법을 이용한다.

[그림 21] 회귀단절설계 기법 개념도



자료: 한승훈과 안혜선 (2019)

이상과 같이 국내 성과 감사의 현황과 한계를 확인하면서 프로그램의 효과를 평가하는 감사가 없었다는 점에서 미국 GAO의 성과감사를 확인하였고, 그 결과 GAO에서는 프로그램의 효과를 평가하기 위하여 준실험적 기법인 이중차분분석이나 회귀단절설계 기법을 적극 적용하는 것으로 나타났다. 다음에는 실제로 GAO에서 수행한 성과감사 사례를 알아보겠다.

3. 미 GAO의 데이터 분석 기법을 이용한 효과성 평가 사례

가. 주요 메탐페타민(필로폰) 성분에 대한 접근 통제가 마약조제 실험실에 미치는 영향(GAO-13-204)

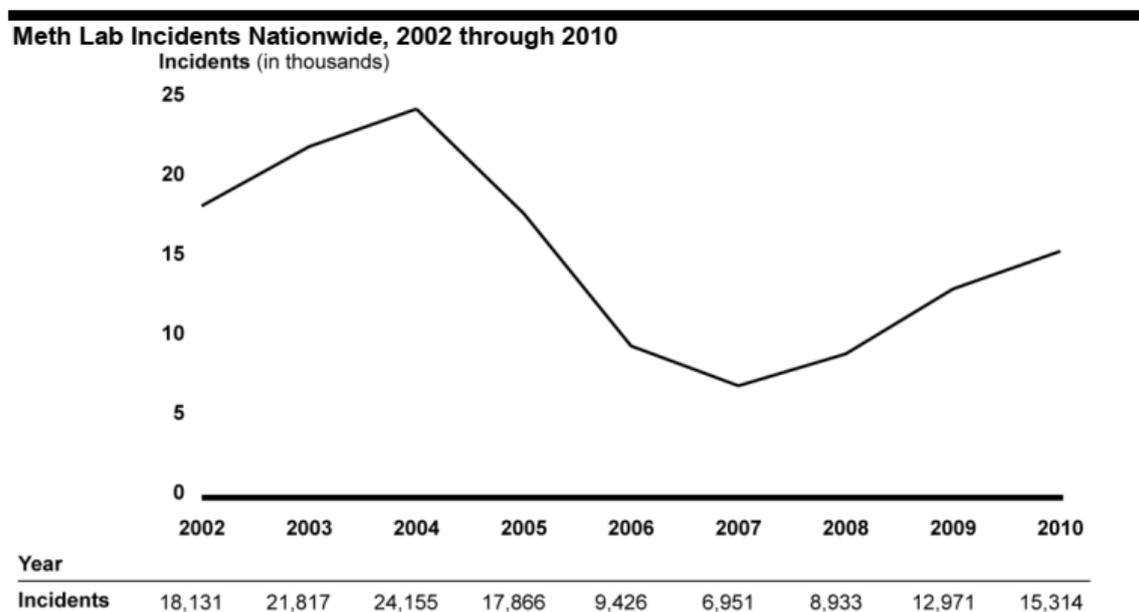
GAO는 감기약이나 알레르기약에서 일반적으로 얻을 수 있는 메탐페타민의 제조 성분 PSE(pseudoephedrine)에 대한 연방정부 및 주정부의 판매 통제에 따른 효과를 평가하기 위하여 마약단속국(Drug Enforcement Administration: DEA)의 자료를 분석하였다.

분석을 통해 [그림 22]과 같이 2007년까지는 PSE(pseudoephedrine) 성분에 대한 판매 금지로 메탐페타민(필로폰)의 실험실 사건이 감소하였으나, 이후 새로운 PSE 확보 방법의 개발과 효율적인 메탐페타민 생산 방법이(smurfing) 도입된 이후 다시 증가하고 있는 것으로 확인하였다.

GAO는 이와 같은 분석결과를 도출하기 위하여 처방전을 동반한 PSE 성분 판매 조치를 2006년 시행한 오레건 주를 대상으로 일반화 추정 방정식(Generalized Estimating Equations: GEE)을 이용한 다항회귀

분석 모형을 이용하여 실험실 사건의 추세를 확인하였다. 또한 이중차분방법(difference-in-difference)을 이용하여 위 주(처치 그룹: Treatment)에서의 결과를 통제된 다른 주(통제 그룹: Control)와 비교하여 그 효과를 도출하였다.

[그림 22] 메탐페타민 사건 발생 현황(2002-2010)



나. 흑인, 남자 또는 장애 학생에 대한 차별의 적정성 검토(GAO-18-258)

GAO 는 학교 현장에서 학생의 인종 등 특성과 관계없이 정당한 차별을 받고 있는지에 대하여 평가하기 위하여 2013 부터 2014 년까지의 학교 현황, 징계자료 및 학생 범죄 자료 등을 수집하여 학생의 인종, 성별, 장애여부 및 소득수준에 따른 징계 현황을 분석하였다.

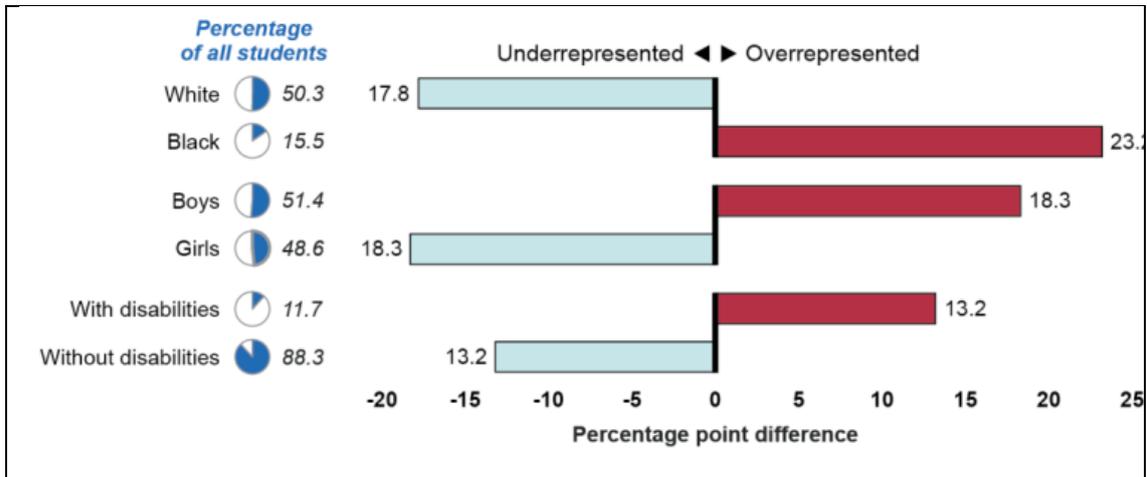
분석에는 일반화 선형회귀분석(Generalized linear regression)을 이용하였고, 나머지 학교 특성이 동일하다고 하였을 때 학교별 특징과 해당 징계를 받은 학생 비율 사이의 관계를 확인하였다.

이때 이용된 학교별 특성은 [표 10]과 같이 학생의 인종, 성별, 장애여부 등 특성에 따른 학교별 학생 비율, 학교의 유형별 특징과 교직원 구성 및 지리적 특성 등을 이용하였고, 학교별 특성이 변경될 때 학교 내 해당 징계를 받은 학생은 어떻게 변화하는 지를 분석하였다.

[표 10] 모형의 독립변수와 종속변수

독립변수	종속변수
다음 각각의 특징을 지니는 학생의 인구비율 : 남학생, 여학생, 백인, 히스패닉, 흑인, 동양인, 원주민, 혼혈, 장애학생, 점심 식사 지원받는 대상	다음 각각의 처벌을 받은 학생들의 비율 - 정학(교외)
학교 유형: 대안학교 여부, 특수 교육학교 여부, 자율형 공립학교 여부, 과학·컴퓨터·외국어·예술 분야 특별 공립 학교 여부	- 정학(교내) - 법
교직원 특성: 법집행관 존재여부, 학교상담교사 존재 여부, 2년 미만 경험을 가진 선생님 비율	집행기관에 이송
학년 운영구분: 6학년 이상, 5학년 이하	- 퇴학
인구 밀집도: 시골, 교외, 도시	- 체형
학생 인구학적 특성과의 상호작용변수 : 히스패닉 남자학생, 흑인 남자학생, 동양 남자학생, 원주민 남자학생, 혼혈 남자학생	- 학교 관련된 체포
고정효과: 주 단위 고정효과(주에 따른 법, 정책, 기타 처벌 결정에 영향을 끼칠 요인들 차이를 설명하기 위함)	

[그림 23] 구성 학생의 특징에 따른 처벌 강도 변화



분석을 실시한 결과 [그림 23]과 같이 학생의 장애 여부, 인종, 성별에 처벌을 받는 수위가 달라지는 것을 확인하였고, 이러한 불균형은 징계 유형, 소득 수준, 학교 유형과 관계없이 발생하는 것으로 확인되었다.

GAO는 이러한 분석 결과와 5개 주의 5개 학군의 학교들에 대하여 인터뷰를 실시하여 감사 결과를 구체화하였고 교육부(Department of Education)와 법무부(Department of Justice)에 결과를 통보하였다.

다. 은퇴 후의 생활 보장 - 소득과 부의 불균형(GAO-19-587)

GAO는 소득과 부의 배분이 고령자들의 은퇴 후 생활이 보장될 수 있는지에 대한 질문에 답을 하기 위하여 증가하는 소득과 부의 편중이 고령자들의 생활에 어떠한 영향을 미치는지 분석하였다.

이를 위해서 GAO는 시간의 경과에 따라 고령자들의 소득과 부의 분포가 어떻게 되어 있는지를 조사하고, 고령자들 사이의 소득 및 부와

장수 사이의 관련성, 각 특성 집단들이 나이가 들면서 소득과 부의 분포는 어떻게 변화하여 왔는가를 조사하였다.

미시간 대학교 사회문제 연구소의 연구 과제의 일환으로 수집된 건강과 은퇴 연구(the Health and Retirement Study: HRS) 자료 등 미국민의 건강 상태와 은퇴 생활과 관련된 자료 소스에서 분석에 필요한 자료를 추출하고 계량경제학 기법을 적용하여 [표 11]과 같이 데이터의 편향성 및 미관측 통제요인 등의 문제를 해소하면서 분석을 실시하였다.

[표 11] 분석 주제별 활용된 분석기법 현황

분석 주제	분석기법
시간의 경과에 따라 고령자들의 소득과 부의 분포	수집한 데이터의 복잡성, 누락, 인구분포상의 조정 등의 문제로 편중된 결과를 보여주므로 관측 값의 분포를 반복적으로 복사하는 통계적 기법으로 처리가 필요하며, Taylor Series 분산 추정 방법 사용
고령자들 사이의 소득 및 부와 장수 사이의 관련성	고령자를 자료가 시간이 지날수록 누락되어 불균형 자료의 문제가 있어 어떤 사건이 발생하기까지 걸린 시간을 분석 및 예측하는 생존 분석(survival analysis) 기법을 적용하고 Kaplan-Meier and Cox 의 비례위험 모형으로 생존 함수를 추정하여 분석
각 특성 집단들의 나이 증가에 따른 소득과 부의 분포 변화	가구 소득 및 재산의 중위 값 수준을 추정한 후, 각 그룹별 추정 값이 어떻게 변화하는 지 추정하여 분석

이러한 분석과정을 거쳐 나이가 들수록 가구별 소득의 불평등은 감소하지만 부의 불평등은 유지되며, 가구의 인종 별, 최종 학력별 차이에 따라 다른 양상을 보인다는 것을 확인하였다.

IV. 데이터 분석 기술을 이용한 정책의 효과 분석 예시

- “학업성적이 우수한 초등학교가 주택 가격에 미치는 영향”

1. 분석 개요

앞에서 알아본 GAO 감사 사례와 같이 GAO는 다양한 분석 기법을 이용하여 프로그램의 효과를 평가하고 있었다.

여기에서는 분석 기술들을 실제 데이터에 직접 적용하여 정책 등 조치가 어떤 결과를 가져왔는지에 대하여 확인하여 보았다.

국외 훈련 과정에서 정부의 프로그램의 효과를 평가하기에는 필요한 정보에 대하여 충분한 접근이 쉽지 않기 때문에 부득이하게 정부 정책과 관련이 높지만 접근이 용이한 자료를 가지고 계량 분석을 실시하여 보았다. 이러한 관점에서 국민들의 관심이 높고 교육 정책과 부동산 정책과 관련이 있는 “학교의 성취도가 주택 가격에 미치는 영향”을 주제로 하여 분석을 실시하였다.

2. 선행 연구 분석

우리 나라에서는 교육과 부동산은 국민들의 관심이 높은 주제이며, 이에 따라 부동산 가격과 교육 환경 및 성취도 사이의 관계에 대해서 많은 연구가 진행되었다. 또한 중국 등 다른 아시아권 국가 뿐 아니라 미국 등 서구권 국가에서도 많은 관심을 가지고 있는 주제라서 관련된 연구가 많다.

진영남과 손재영(2005)은 2002년 서울시 아파트 가격을 분석하여 사설학원의 수, 관할 교육청의 교육 예산 편성금액, 명문대 진학률이 주택 가격에 유의미한 영향을 미친다는 것을 발견하였다.

또한, 김예지와 이영성(2014)은 2010년부터 2011년까지 국가수준 학업성취도 평가 결과에 따른 서울시내 초·중·고등학교의 학업성도가 아파트 가격에 미치는 효과를 연구하여 중학교 학업성과와 서울대 진학률, 사설학원 수가 아파트 가격에 유의미한 영향을 준다고 확인하였다.

다른 나라의 경우에도 Wu and Huang (2017)은 중국의 베이징 광산(Fangshan) 지역의 주택 거래 데이터를 이용하여 명문학교(Key School)에서 거리에 떨어진 거리에 따라서 주택 가격이 달라진다고 밝혔다.

Hansen(2014)은 영국에서 토지 등록 데이터(Land Registry data)와 밀레니엄 코호트 연구(Millennium Cohort Study: MCS)¹의 데이터를 이용하여 유치원(pre-school) 시절에 발생하는 주거지 이동을 조사한 결과 영국의 부모들은 더 좋은 초등학교 근처의 주택을 구입하기 위하여 더 많은 비용을 지불할 준비가 되어있음을 발견하였다.

또한 Figlio and Maurice(2000)는 미국 플로리다주의 중간 규모 학군을 대상으로 학생들의 성적과 주택 가격을 조사하여 학교의 성적은 주택 가격에 큰 영향을 미친다는 중요한 증거를 확인하였다.

¹ 다양한 목적의 연구 조사를 위한 데이터를 생성하기 위하여 영국 런던대학에서 2000년부터 2001년 사이(21세기) 태어난 18,818명의 아이들을 대상으로 실시한 종단 조사.

학교의 평판과 주택의 가격은 높은 상관 관계를 가지고 있다는 것은 대부분의 사람들이 경험적으로 느끼고 있을 뿐 아니라 위와 같이 다양한 연구를 통해서 밝혀졌다. 하지만 학교의 명성과 주택의 가격 사이에는 높은 상관관계가 있다는 것을 확인할 수 있을 뿐 좋은 학교가 주택 가격의 상승을 유발하는 것인지, 아니면 비싼 주택 주변의 학교가 좋아지는 것인지를 확인할 수 있는 인과 관계에 대해서는 더 고민할 필요가 있다.

그 이유는 좋은 주거 환경을 가지는 지역에 위치한 학교가 일반적으로 학업 성취도가 높거나, 비싼 주거 비용을 지불할 수 있는 부모들이 자녀들의 교육에 더 많이 투자하여 높은 성취를 보여주는 경향이 있기 때문이다. 서울시내 중학교의 학업성과와 서울대 진학률 등이 아파트 가격 상승에 유의미한 영향을 미친다고 밝혔던 김예지와 이영성(2014)은 아파트 가격이 초·중·고등학교 학업성과와 서울대 진학률에 유의미한 영향을 준다고 확인하였다.

계량경제학에서 확인하고자 하는 문제는 특정 현상에 의한 효과가 어떠한 결과로 나타나는 것인지에 대한 인과관계를 규명하는 것에 있다. 또한 정부의 프로그램을 평가하여야 하는 성과감사의 경우에도 조치에 대한 성과가 어떤 결과를 나타내는지에 대한 인과관계를 확인하는 것이 중요하다. 하지만 회귀분석을 통해서 발견한 설명변수와 종속변수의 관계는 상관관계를 확인할 수 있을 뿐 인과관계를 보장하지 않고 있고, 이러한 인과관계를 규명하기 위한 많은 기법들이 발전되어 왔다.

여기에서 다루고 있는 학교의 학업 성취도와 주택 가격 사이의 관계에 있어서 상관관계와 인과관계를 구분하여 규명하는 것이 필요한데,

Black(1999)은 이 문제를 해결하기 위하여 보스턴 외곽에 위치한 Middlesex, Essex, Norfolk 카운티 지역을 대상으로 동일한 학군내 인접하여 위치하지만 통학구역 경계를 따라 양쪽에 위치한 주택들을 비교함으로써 주거 환경, 세금, 교육 예산 등 다양한 변수들의 영향을 제외하고 오직 초등학교의 차이에 따른 영향만 비교하도록 설계하였고, 학교 점수가 5% 상승시 부모는 2.1% 더 많은 주거 비용을 부담할 용의가 있음을 확인하였다.

또한 리처드 등(2018)은 뉴욕시의 학군이 주택 가격에 미치는 평균처치효과(Average Treatment Effect: ATE)를 확인하기 위하여 지리적인 회귀단절모형(Geographical Regression Discontinuity Designs: GeoRDDs) 기법을 적용하여 분석하였다.

반면 기존에 국내에서 실시된 연구에서는 학군 내 평균 주택가격과 평균 학업 성취도를 다른 학군의 값과 비교하여(김경민 외[2010], 안문영과 추준석[2017]) 동일 학군 내에서 학교별 차이에 따른 주택가격 변화를 확인하기 어렵거나, 학교에서부터 떨어진 거리를 통학범위로 간주하여 분석(김예지[2014], 남형권과 서원석[2016])하여 학교별 차이가 정확하게 반영되기 어려운 한계가 있었다.

이러한 경우 학군 내에서 변화하는 주거환경의 차이를 반영하지 못하거나 동일한 주거 환경에서 학교만 달라지는데 따른 주택가격 변화를 정확히 반영하지 못할 수 있다.

이에 본 연구에서는 통학범위가 구체적인 초등학교²에 대해서만 학교의 학업 성취도를 비교하였고, GIS 기반의 학구도안내서비스를 이용하여 초등학교 통학 범위에 포함되는 아파트를 식별하였다. 또한 2011 년 한 일간지에 의해서 단발성으로 공개(조선일보, 2011. 12. 2.)되었던 서울시 학교별 학업 성취도 결과를 유사 실험(quasi-experiment)으로 이용하여 초등학교의 학업 성취도에 따른 해당 아파트 가격의 변화를 분석하였다.

3. 국내 교육제도 현황 및 학업성취도 평가

우리 나라는 높은 학업성취도를 가지는 것으로 유명하며, 오바마 전임 미국 대통령이 수시로 한국 공교육의 우수성을 언급한 바 있다. 또한 한국인들은 자녀 교육에 높은 관심을 가지고 있어 “교육열”이라는 단어가 있을 정도이다. 이러한 교육에 대한 높은 관심과 열정은 우리나라가 단시간에 빠른 경제성장을 이룰 수 있도록 한 원동력이다.

하지만 교육에 대한 높은 집착은 학생들을 성적에 따라 서열화 하여 치열한 경쟁을 유도하고, 높은 10 대 자살율로 나타난 학생들의 행복을 악화시킬 뿐 아니라 과도한 사교육비로 인한 가계 부담을 가중시키는 등 부정적 외부효과를 보여주고 있다.

이와 같은 부작용이 지속되면서 2000 년 이후 교육에서 지나친 경쟁을 완화하고자 하는 움직임이 보여주고 있는데, 초등학교에서 시험을

² 초등학교는 일부 공통 통학구역을 제외하고 대부분 독점적인 통학구역을 가지고 있는 반면, 중학교와 고등학교는 공통된 통학구역 내 위치한 학교들 중 추첨으로 학교가 배정

폐지하고 학생을 평가하는데 있어 외부활동, 생활기록부, 동아리 활동 등 다양한 기준을 반영하는 등의 조치가 있었다. 이러한 추세는 2008 년부터 시작되었던 전국단위 학업성취도 평가에도 영향을 미쳤다.

교육부는 ‘기초학력미달 제로플랜’의 일환으로 2008 년부터 기존 표집 방식으로 실시하던 초·중·고등학생에 대한 국가수준 학업성취도 평가를 전수 평가로 전환하여 시행하였다. 시행 초기에는 학업성취도 평가결과는 학생들의 기본 학업성취수준을 측정하고 교육정책에 반영하는 내부적 검토 자료로 활용되었다. 하지만 2010 년부터 각 과목별 우수, 평균, 미흡한 학생 비율(교과별 3 단계 성취수준 비율)을 ‘학교 알리미 사이트 (www.schoolinfo.go.kr)에 공시하여 외부에 공개하였다.

이때 각 학교별로 개별적인 성취수준 학생 비율을 확인할 수 있으나 전체 학교의 정보를 일괄적으로 볼 수 있거나 학교별 순위 등은 공개되지 아니하고, 교육지원청 단위로 통계정보만 공개되어 지역별 평균적인 학생의 성취 수준 비율 정도만 확인할 수 있었다.

하지만 학업성취도 평가 결과에 대하여 학교 및 교육청 간 과열 경쟁과 학생들의 평가 부담이 증가하는 등 부작용이 발생하고 있었고, 민간 입시 컨설팅 회사들을 통해 지역별 학교 순위가 작성되어 공공연하게 순위가 알려지는 한편 2011 년 서울지역 학교들의 평가 결과는 국내 일간지를 통해 공개되면서 학교별 순위가 알려졌다.

이러한 논란들을 지속되고 있어 교육부는 2013년 4월 “2013년 국가수준 학업성취도 평가 기본계획”을 마련하면서 학업성취도 평가에서 초등학교는 제외하고 중학교는 평가과목을 축소하는 것으로 변경하였다.

그 결과 초등학교에 대한 국가수준 학업성취도 평가 결과는 2010년, 2011년, 2012년에 대해서만 공개되었고, 서울시 전체 초등학교에 대한 학교별 평가결과와 순위는 2011년 결과만 공개되었다.

본 연구에서는 서울지역 초등학교 학업성취도 결과가 2011년 한차례만 공개된 사실을 이용하여 자연 실험(Natural Experiment)에 의한 처지(Treatment)가 어떤 효과를 나타내는지 분석하여 초등학교의 학업 성취도가 주택가격 변화에 미치는 영향을 확인하였다.

4. 자료 수집

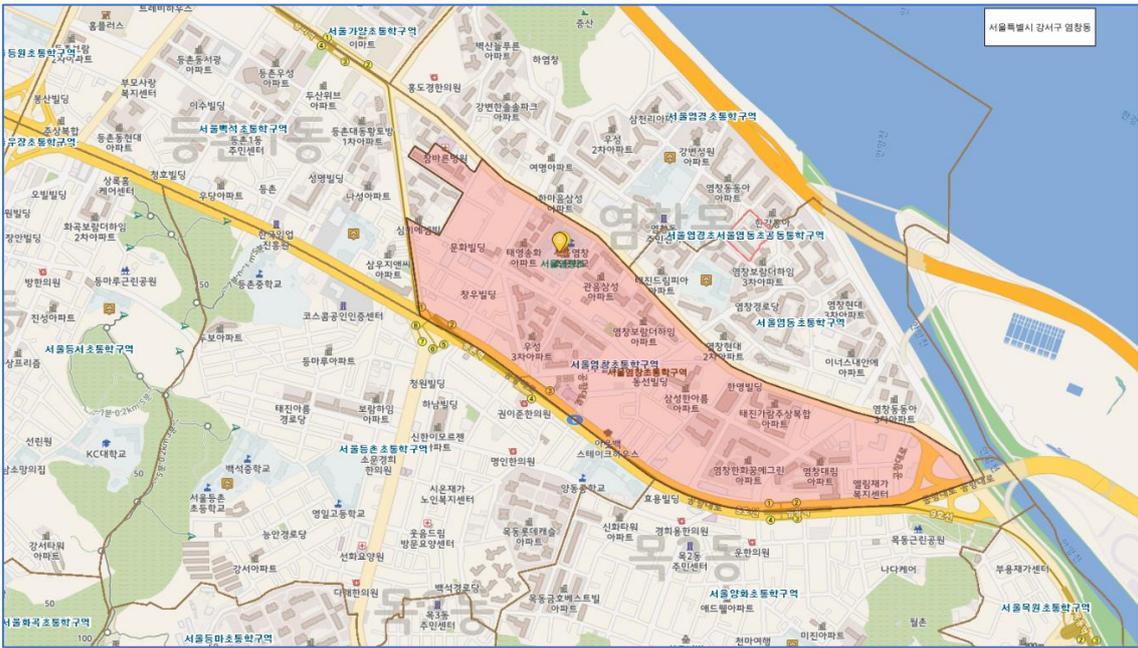
가. 우수한 학교와 해당 학교의 통학범위 정보 수집

학교의 학업성취도와 주택의 가격 변화 사이의 인과관계를 확인하기 위하여 학교의 통학범위 내 위치한 주택을 매칭하여야 하는데 교육부에서 운영하는 학구도안내서비스 (<https://schoolzone.emac.kr/gis/gis.do>)를 이용하여 학교별 통학범위를 확인할 수 있었다.

분석 대상은 초·중·고등학교 중 초등학교를 대상으로 하였다. 서울시의 경우 초등학교는 공통 통학구역을 가지는 일부 학교를 제외하고 독점적인 통학구역을 가지고 있어 해당 초등학교에 배정되는 주택의

범위를 결정할 수 있다. 반면 중학교나 고등학교는 각각 중학구 및 학군 범위 내에 위치한 학교들 중에서 지원 및 추첨 형식으로 학교가 배정되기 때문에 해당 중·고등학교에 배정되는 주택의 범위를 결정할 수 없다. 아울러 고등학교는 주거지의 위치와 관계없이 지원하여 들어갈 수 있는 과학고, 외국어고 등의 특수목적 고등학교나 자율형 사립학교에 지원할 수 있어서 고등학교의 학업 성취도와 주변 주택가격과의 인과관계가 높을 수 없다. 이와 같이 초등학교는 중·고등학교와 다르게 동일한 주거환경을 가지지만 학교에 따라서 주택 가격의 변화가 달라지는 정도를 확인하기에 적절하다.

[그림 24] 학구도안내서비스 예시



자료: 학구도안내서비스 (<https://schoolzone.emac.kr/gis/gis.do>)

또한 중학교와 고등학교에서는 여전히 전국단위 학업성취도 평가가 실시되고 그 결과가 공개되고 있고, 각각 과학고, 외국어고 등 특목고 입학

비율 또는 명문대 입학 비율, 수능점수 분포 등을 이용하여 학교별 학업 성취도를 확인할 수 있어 처지(Treatment)군과 통제(Control)군을 완벽하게 통제하기 어렵다. 하지만 초등학교의 경우에는 2011년 전국단위 학업성취도 결과 외에는 학교별 성적을 확인할 수 있는 대안이 없으므로 이는 단발성 쇼크 또는 처지에 해당하여 유사 실험(quasi-experiment)의 여건을 잘 갖추게 된다.

이와 같은 이유로, 김예지와 이영성(2014)의 경우 중학교의 학업성취도 결과가 주변 주택가격에 가장 큰 영향을 준다고 확인하였지만, 본 연구에서는 초등학교의 학업성취도 결과가 해당 학교의 통학범위 내 주택 가격의 변화에 미치는 영향을 확인하는 것으로 범위를 정하였다.

2011년 12월 공개된 서울시 학교별 2011년 전국단위 학업성취도 평가 자료에는 학교별 ‘보통학력 이상³’ 학생 비율과 서울시내 학교 중 순위, 각 기초자치단체(구) 내에서 순위 정보가 포함되어 있다.

우수 학교에 대한 정보는 학교별 순위나 보통학력 이상 학생비율 값이 아닌 우수학교 여부를 나타내는 더미변수로 적용하였고, 이때 우수학교 여부는 서울 시내 550 여개 공립⁴ 초등학교 중 상위 50 개 초등학교를 선정하였다. 또한, 우수 학교가 강남구, 서초구, 송파구 및 양천구에 주로 분포되어 [표 12]와 같이 상위 50 위 내에 해당하는 초등학교가 없는 구가 있으므로 각 구별 상위 10%에 해당하는 초등학교는 상위 50 위내 해당하지 아니하여도 포함시켰고, 총 60 개 초등학교를 우수 초등학교로 추출하였다.

³ 전국단위 학업성취도 평가 결과는 학교별 3 단계 성취수준 학생 비율을 제공하고, 3 단계 성취수준은 보통학력 이상, 기초학력, 기초학력 미달로 구분됨

⁴ 통학구역에 따라 배정되지 않는 사립 초등학교 제외

[표 12] 각 구별 초등학교, 우수 초등학교 현황

district	number of schools	number of 50th highest	proportion of outstanding
강남구	30	6	0.89
강동구	25	0	0.84
강북구	12	0	0.78
강서구	33	0	0.82
관악구	20	0	0.80
광진구	18	2	0.82
구로구	23	0	0.81
금천구	14	0	0.76
노원구	36	0	0.83
도봉구	19	0	0.82
동대문구	17	0	0.80
동작구	18	0	0.85
마포구	20	0	0.83
서대문구	14	0	0.83
서초구	18	4	0.92
성동구	17	0	0.81
성북구	23	0	0.81
송파구	37	8	0.87
양천구	25	3	0.87
영등포구	22	1	0.80
율산구	12	0	0.83
은평구	23	0	0.80
종로구	11	0	0.77
중구	9	0	0.77
종랑구	20	0	0.76

자료: 2011년 공개된 서울시 초등학교 전국단위 학업성취도 결과 분석

이렇게 추출된 60 개 초등학교의 통학 구역의 범위를 학구도 안내 서비스를 이용하여 확인하여 해당 범위내 포함되는 주택에는 우수학교를 나타내는 더미 변수의 형태로 우수 초등학교의 통학구역에 해당한다는 것을 표시하였다.

나. 주택 거래 정보 수집

우리나라에서 가장 일반적인 주거형태는 아파트이며, 특히 서울의 경우 대부분 아파트에 거주한다고 볼 수 있다. 또한 아파트는 단지별로 아파트 면적에 따라 유사한 가격 분포를 보여주므로 아파트 단지별 면적 유형에 따라 동일한 거래로 간주할 경우 매매 거래가 체결될 경우만 확인할 수 있는 거래 정보를 지속적으로 수집할 수 있다. 따라서 주택가격의 변화는 아파트 단지를 이루고 있는 아파트를 대상으로 하였다.

우리나라는 법률의 규정에 따라 2006 년 이후 체결되는 주택 거래는 신고하여야 하며, 이때 신고된 정보를 수집하여 국토교통부에서 운영하는 “실거래가 공개시스템” (<http://rtdown.molit.go.kr/>)을 통해 공개하여 제공하고 있다.

2006 년 이후 제공된 아파트 매매 정보는 아파트 단지의 이름, 주소(시, 구, 동 및 번지), 해당 아파트의 면적, 매매계약 체결일자, 가격, 층수, 아파트 연령 정보가 포함되어 있다. 본 연구에서는 우수 초등학교 정보가 공개된 2011 년을 기준으로 이전 6 년, 이후 6 년간 아파트 거래

내역을 분석 대상으로 하여 2006년부터 2017년까지 서울시내 아파트 거래 정보를 위 실거래가 공개시스템에서 추출하였다.

이렇게 추출한 데이터에서 아파트 단지의 이름과 아파트 면적을 결합하고 ID 를 부여하여 아파트 가격의 변화를 추적할 기본 개체로 활용하고, 각 개체별 연간 거래 가격의 중앙 값⁵을 대표 가격으로 선정하고 연간 거래 건수를 계산하였다.

또한 면적이 너무 작은 소형 평수는 초등학생을 자녀로 두지 않은 1인 가구나 신혼 부부가 거주할 확률이 높고, 면적이 너무 큰 대형 평수 아파트는 거래가 활발하지 않을 뿐 아니라 일반적인 매물과 다른 거동을 보일수가 있으므로 45m² 미만과 135m²를 초과하는 개체는 제거하였다.

아울러 거래가 자주 일어나지 않은 소규모 아파트 단지의 거래는 이상점이 발생할 확률이 높으므로 이를 제외할 필요가 있고, 분석 대상인 12년 기간 중 매매 기록이 누락된 시기가 있는 개체가 있을 경우 균형된 패널 데이터(balanced panel data)를 구성할 수 없으므로 연간 5회 미만 거래된 개체와 12년 기간 중 매매 기록이 누락된 개체는 제외하였다.

이러한 과정을 거쳐서 매매 거래가 많이 발생하여 이상점(outlier)에 대표 값이 왜곡되지 않고 안정적인 가격 변화를 보일 수 있는 세대수가 많은 아파트 단지의 개체만 추출하여 데이터를 구성하고 누락된 정보가 없는 균형된 패널 데이터를 구성할 수 있었다.

⁵ 거래 유형에 따라 특별히 가격이 높거나 낮은 거래가 이상점(outlier)을 형성하여 평균(mean)을 왜곡할 수 있으므로 중앙 값(medium)을 대표 값으로 이용

이렇게 구성한 데이터는 [표 13]과 같이 총 958 개의 개체(아파트 단지별 면적별 대표 아파트)의 12 년간 거래된 11,496 건 정보의 패널데이터로 각 개체별 위치, 연령, 면적 정보가 각 연도별 매매 가격의 중앙값, 단위 매매 가격 (단위 면적당 가격)의 중앙값, 거래 건수 등의 데이터로 구성되어 있다.

[표 13] 아파트 매매 정보에 대한 통계 요약

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
age	11,496	24.649	7.133	2	19	31	49
area_CT	11,496	71.628	16.890	50	60	80	130
mid_price	11,496	41,505.250	20,902.960	8,050	27,750	49,000	254,000
mid_Uprice	11,496	566.320	239.034	172.616	416.026	641.428	2,590.314
numb_contrat	11,496	28.442	23.765	5	13	35	323

year	dstrt	id	chk.y	sqM	priceC
2006 : 958	노원구 :1668	가락동가락(1차)쌍용아파트60 : 12	0:9108	L:1116	high:2914
2007 : 958	도봉구 : 768	가락동가락(1차)쌍용아파트80 : 12	1:2388	M:5028	mid :5799
2008 : 958	성북구 : 768	가락동가락금호아파트60 : 12	NA	S:5352	low :2783
2009 : 958	구로구 : 636	가락동가락금호아파트80 : 12	NA	NA	NA
2010 : 958	양천구 : 600	가락동우성60 : 12	NA	NA	NA
2011 : 958	영등포구 : 588	가산동두산60 : 12	NA	NA	NA
(Other):5748	(Other) :6468	(Other) :11424	NA	NA	NA

위 정보에 각 구별 특성을 설명할 수 있는 정보를 추가하기 위하여 각 구별 주택가격지수, 총 주택 거래량, 가구 수를 수집하여 포함하였다.

마지막으로 앞에서 추출한 우수 초등학교의 통학구역 내 해당하는 아파트 단지 정보를 이용하여 우수 초등학교를 나타내는 더미 변수를 포함시켰고, 전체 958 개의 개체 중 199 개 개체가 우수 학교의 통학구역 내 위치한 아파트로 처치 그룹(Treatment group)에 해당하고 나머지 759 개 개체가 통제 그룹(Control group)에 해당한다.

5. 모델 구성 및 분석 결과

앞에서 구성한 패널 데이터를 이용하여 우수한 초등학교가 주변 아파트 가격에 미치는 영향을 확인하였다. 우수한 초등학교 주변 아파트를 처치 그룹(Treated), 그 이외의 아파트를 통제 그룹(Control)으로 보고 우수 초등학교에 대한 정보가 노출된 2011년 12월, 즉 처치가 이루어진 전후에 처치 그룹과 통제 그룹의 움직임을 비교하여 처치에 따른 결과의 인과관계를 확인하는 이중차분방법(Difference-in-Difference: DiD)을 이용하여 분석을 실시하였다.

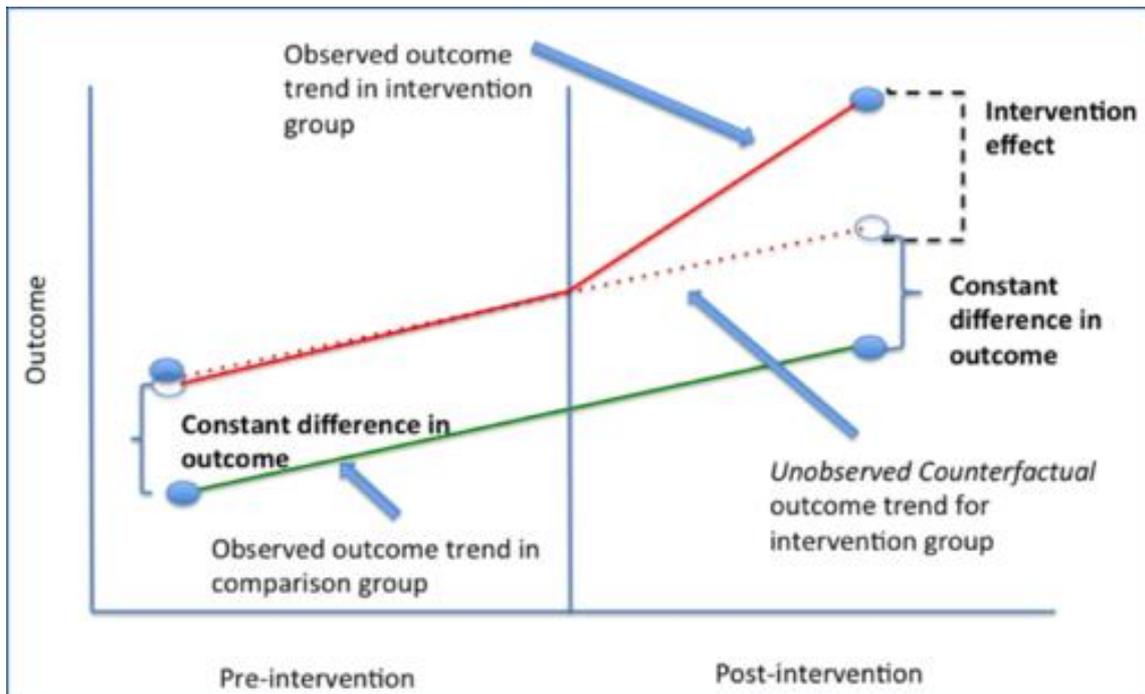
가. 이론

처치에 따른 결과의 인과관계를 확인하기 위해서는 처치가 이루어진 그룹이 처치로 인해 발생한 결과가 만약 처치가 없었다면 발생하였을 결과의 차이 값이 처치로 인한 효과이다. 하지만 처치가 이루어진 그룹에 만약 처치가 없었을 경우 발생하였을 효과는 실제로 발생한 사실과 반대되는(Counterfactual) 상황에서 값이므로 우리가 관측할 수 없고, 그 결과 처치로 인한 효과를 알 수 없다.

본 연구에서 적용한 이중차분분석(DiD) 방법은 처치가 발생하기 전과 후에 걸쳐서 처치가 이루어진 그룹과 처치가 이루어지지 않은 통제그룹의 움직임이 동일한(평행한) 모습을 보인다는 가정을 가지고 처치 그룹과 통제 그룹의 처치 전 후의 결과값 변화의 차이를 이용하여 처치의 효과를 구하는 방법이다.

처치 그룹(Treatment Group)을 A(빨간색), 통제 그룹(Control Group)을 B(초록색)이라고 하고, 처치 이전의 시간을 1(Pre-intervention), 이후를 2(Post-intervention)이라고 한다. [그림 25] 및 [표 14]와 같이 처치 이전에는 처치 그룹은 A, 통제 그룹은 B의 값을 가지고 있었으나 처치 전후 시간이 경과함에 따라 T 만큼 효과가 각각 반영되었고, 처치를 받은 처치 그룹은 처치로 인한 효과 E 만큼 반영이 된다.

[그림 25] 이중차분 방법에 대한 개념도



즉 우리가 알고 싶은 처치에 따른 효과 E는 처치 그룹의 처치 이후 결과값 $Y_{A,2}$ 에서 처치가 없었다면 발생하였을 결과인 $Y_{A,2}' (= A+T)$ 을 뺀 값인데, 이 값을 우리가 관측할 수 없으므로 통제 그룹의 처치 전후의 결과값 변화($Y_{B,2} - Y_{B,1}$)와 처치 그룹의 처치 전후의 결과값 변화($Y_{A,2} - Y_{A,1}$)의 차이 값을 구하면 처치 효과인 E를 구할 수 있다.

[표 14] 이중차분 방법에 대한 개념

구분	시간	결과	1 단계 차분	2 단계 차분
A	1	$Y_{A,1} = A$	-	-
	2	$Y_{A,2} = A + T + E$	T+E	E
B	1	$Y_{B,1} = B$	-	-
	2	$Y_{B,2} = B + T$	T	-

$$E = (\overline{Y_{B,2}} - \overline{Y_{B,1}}) - (\overline{Y_{A,2}} - \overline{Y_{A,1}})$$

이는 다음과 같이 처치에 대한 더미 변수인 Treat, 처치 전후에 대한 시간 경과를 나타내는 더미 변수인 Post 로 하여 회귀식을 구성하면 δ_1 의 값이 우리가 알고 싶은 처치에 따른 효과이다.

$$y_{it} = \beta_0 + \beta_1 Treat_i + \delta_0 Post_t + \delta_1 (Treat \times Post)_{it} + \epsilon_{it}$$

또한 시간에 따른 동일한 변화 추세를 가정하면 DiD 모델은 Two-way fixed effect 패널 모형으로 나타낼 수 있다. 처치 전후의 시간 경과에 따른 효과의 변화는 시간에 따른 특성 γ_t , 처치를 받은 그룹과 그렇지 않은 그룹의 차이는 개체에 따른 특성 γ_s 에 해당하고 처치 여부를 나타내는 더미변수를 D_{st} 라고 하였을 때 다음과 같이 회귀 식이 구성되며, 이때 산출된 δ 값이 우리가 구하고자 하는 좋은 학교가 주변 아파트 가격에 미치는 영향이다.

$$Y_{ist} = \gamma_s + \tau_t + \delta D_{st} + \epsilon_{ist}$$

나. 모델 구성

Cheng and Hoekstra(2013)는 정당방위 범위를 확대하는 법 (Castle Law) 실시에 따른 범죄 예방과 사망 사건 증가 효과를 확인하기 위하여 주별 Castle Law 적용 시기의 차이를 이용하여 이중차분(DiD) 분석을 실시하였고, Cornwell and Cunningham(2013)는 낙태 허용으로 인한 국민의 건강 상태 개선 효과를 확인하기 위하여 미국 주별로 낙태 허용 시점이 다른 점을 이용하여 이중차분(DiD) 분석 및 이를 확대한 삼중차분(Triple Difference) 분석을 실시하였다.

본 연구에서는 위 선행 연구의 분석 모델을 참고하여 이중차분 모델을 다음과 같이 Two-way Fixed Effect 모델로 구성하였다.

$$\log(\text{Unit Price}_{i,d,t}) = \beta_0 + \beta_1 \text{Treat}_{i,d,t} + \beta_2 \text{Yr}D_t + \beta_3 \text{Id}D_{i,d} + \beta_4 \text{Dist}_d + \beta_5 X_{s,t} + e_{i,d,t}$$

여기에서 i, d, t 는 각각 개별 아파트 개체(아파트 단지 및 면적별 대표 아파트), 구, 년도를 나타내는 첨자이며, 종속변수는 단위 면적당 거래가격의 로그 값, 독립변수로 우수 학교 주변에 위치한 아파트 여부를 나타내는 처치 더미 변수 $Treat$ 이다. 본 연구에서는 $Treat$ 변수는 주변에 우수 초등학교가 있는지 알게 된 이후와 나머지를 구분하는 더미 변수와 우수 학교 여부가 공개된 1~2 년 기간, 2~4 년, 4~6 년 구간을 각각 나타내는 더미 변수를 적용하여, 주변에 우수 학교가 있다는 정보가 공개된 처치에 대한 효과가 시간이 지날수록 어떻게 변하는지를 확인하였다.

이 외에 연도 더미 변수 YrD , 아파트 개체 아이디 IdD , 각 구별 특성을 나타내는 구별 통제 변수(구별 주택가격지수, 총 주택 거래량, 가구수) $Dist$, 그리고 기타 지역별 특징이나 가격별 특성의 연도별 추세를 보여주는 연도별 통제 변수 Xs 로 회귀 모델을 구성하였다.

다. 분석 결과

우선 연도별 통제 변수를 반영하지 아니한 간단한 형태의 모델을 분석하였다. 이때 회귀 모형의 독립 변수로 [표 15]와 같이 좋은 학교의 정보 공개 여부(GS)만 포함된 two-way fixed effect 모형(1), 여기에 각 구별 아파트 가격 지수를 포함시킨 모형(2), 좋은 학교에 대한 정보 공개 이후 시간 경과가 반영된 더미 변수($GS1, GS2, GS3$ 로 각각 정보의 공개 후 1~2 년, 2~4 년, 4~6 년 경과된 시점을 나타내는 더미변수)만 포함된 모형(3), 여기에 구별 아파트 가격 지수를 포함시킨 모형(4)을 분석하였다.

그 결과, [표 15]와 같이 4 개 모형 모두 우리가 예측과 다르게 주변에 좋은 학교가 있다는 정보가 공개된 이후 해당 아파트 가격이 다른 아파트에 비해 오히려 가격이 덜 상승한 것으로 나타났다. 정보 공개 이후 전체적으로 1.8% 내지 1.9% 만큼 가격이 덜 상승하였고, 시간 경과 시점에 따라서는 공개된 직후는(2012, 2013 년) 가격이 2.7~2.9% 만큼 덜 상승하고 시간이 지날수록 그 효과는 점차 완화되는 것으로 나타났고, 모든 결과 값은 통계적으로 유의미한 수준이었다.

[표 15] 단순 분석 모형 실행 결과

<i>Dependent variable:</i>				
	1_m_Uprice			
	(1)	(2)	(3)	(4)
GS	-0.018** (0.008)	-0.019** (0.008)		
GS1			-0.029*** (0.008)	-0.027*** (0.007)
GS2			-0.015** (0.008)	-0.016** (0.007)
GS3			-0.010 (0.010)	-0.014 (0.010)
District Price index	No	Yes	No	Yes
Years later Treatment	No	No	Yes	Yes
Observations	13,092	13,092	13,092	13,092
R ²	0.943	0.957	0.943	0.957
Adjusted R ²	0.938	0.953	0.938	0.953
Residual Std. Error	0.092 (df = 11989) 0.081 (df = 11988) 0.092 (df = 11987) 0.081 (df = 11986)			

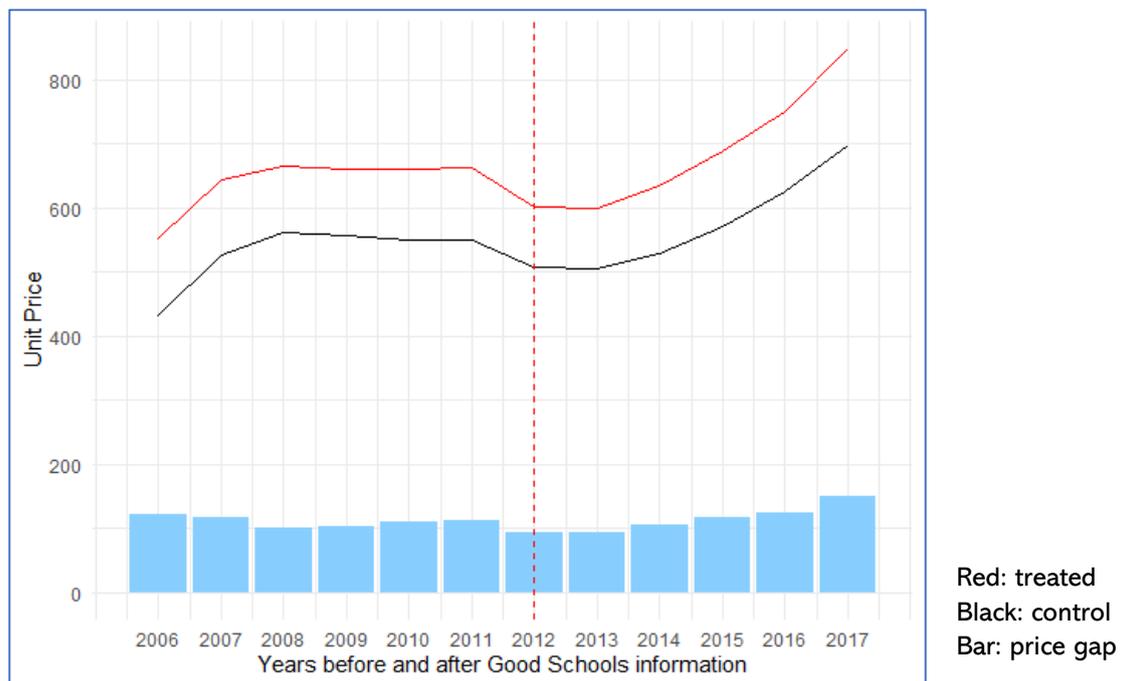
Note:

*p<0.1; **p<0.05; ***p<0.01

이와 같이 분석 결과를 확인하기 위하여 좋은 학교 주변의 아파트(처치 그룹)와 나머지 아파트(통제그룹)의 평균 가격의 변화를 [그림 26] 과 같이 확인하면, 좋은 학교 주변 아파트가 다른 아파트에 비해

가격이 비싸고 학업 성취도 평가 결과가 공개된 2011~2012 년 사이 아파트 가격은 전반적으로 하락하였다. 이러한 추세를 고려하면 당시 경제 여건에 의하여 가격이 비싼 아파트가 저렴한 아파트 보다 상대적으로 더 하락하는 등 통제되지 못한 요인이 영향을 미쳤다고 볼 수도 있었다.

[그림 26] 우수 학교 주변 아파트와 나머지 아파트의 단위 가격 변화



이와 같이 아파트 면적, 가격, 재건축 요인(연령) 등 아파트 단지에 따른 특성이 각 연도별 경제 상황이나 사회 여건에 따라 다르게 움직일 수 있으므로⁶ 이러한 요인을 통제하기 위하여 연도별 통제 변수 X_s 를 도입하여 분석을 실시하였다.

⁶ 특정 해에는 소형 아파트가 인기가 높아 가격이 많이 오르고, 다른 해에는 고급 아파트가 인기가 더 많은 등의 추세

이러한 연도별 통제 변수로는 면적 구간별(소형, 중형, 대형), 가격 구간별 추세가 반영된 연도별 특성 변수와 구별⁷ 추세 변수를 반영하여 [표 16]과 같이 회귀모형을 구성하였다. 이때 독립변수로 같이 좋은 학교의 정보 공개 여부(GS)와 연도별 특성 변수와 구별 추세 변수를 반영한 모형(1), 구별 추세변수 대신에 구별 아파트 가격지수를 반영한 모형(2), 좋은 학교에 대한 정보 공개 이후 시간 경과가 반영된 더미 변수(GS1, GS2, GS3)와 연도별 특성 변수와 구별 추세 변수를 반영한 모형(3), 구별 추세변수 대신에 구별 아파트 가격지수를 반영한 모형(4)에 대해서 회귀분석을 실시하였다.

그 결과 학업 성취도가 높은 초등학교의 통학범위 내 아파트가 다른 아파트에 비해 1.8 내지 2% 더 높이 가격이 상승한 것으로 나타났고, 시기 별 구분하여 효과를 살펴보면 우수 초등학교 정보가 공개된 직후에는 가격 상승률이 0.8 내지 1% 더 높았으나 시간이 지날수록 그 효과가 높아져 4~6년 기간에는 2.8 내지 3% 더 높이 가격이 상승하는 효과를 보이는 것으로 확인되었다.

⁷특정 해에는 강남, 서초, 송파 지역의 인기가 높을 수 있고 다른 해에는 정부의 부동산 규제 등의 영향으로 나머지 지역의 인기가 높을 수 있는 등 추세 반영

[표 16] 연도별 통제 변수가 반영된 모형의 실행 결과

	<i>Dependent variable:</i>			
	1_m_Uprice			
	(1)	(2)	(3)	(4)
GS	0.018*** (0.005)	0.020*** (0.005)		
GS1			0.008* (0.004)	0.010** (0.005)
GS2			0.018*** (0.005)	0.020*** (0.005)
GS3			0.028*** (0.006)	0.030*** (0.006)
District-Year dummy	Yes	No	Yes	No
District Price index	No	Yes	No	Yes
Years later Treatment	No	No	Yes	Yes
Year specific trends	Yes	Yes	Yes	Yes
Observations	11,496	11,496	11,496	11,496
R ²	0.985	0.982	0.985	0.982
Adjusted R ²	0.983	0.980	0.983	0.980
Residual Std. Error	0.047 (df = 10216)	0.051 (df = 10479)	0.047 (df = 10214)	0.051 (df = 10477)
<i>Note:</i>			*p<0.1; ** p<0.05; *** p<0.01	

V. 결론

지금까지 데이터 분석기술을 활용한 성과감사의 추진전략에 대하여 확인하여 보았다.

그동안 있어온 많은 성과감사 활성화 방안이나 성과감사 모델 개발 관련 연구에서 현재의 성과감사에 대한 전반적인 한계점을 확인하고 해외의 선진 사례를 반영하여 성과감사 개선방안을 제시하여 왔다. 이러한 연구는 주로 성과감사 자체가 왜 많이 실시되지 않고 있는지에 대한 연구를 통해 성과감사 확대를 저해하거나 우려하는 문제를 확인하고 개선책을 마련하는데 방점을 두고 있다.

공병천과 윤기웅(2017)은 성과감사를 활성화시키는데 문제점으로 감사 실무부서는 적발과 처벌 위주의 감사를 선호하고, 성과감사 결과에 대한 정치적 해석으로 감사의 독립성과 신뢰성 훼손이 우려되는 점 등의 문제를 제시하면서 한국형 성과감사 모델에 대해서는 우선 성과감사의 범위를 정의하는 것이 필요하며, 이에 따라 성과감사를 정형화하고, 방법론을 구성하면서 법적, 행정적 체계를 마련하는 것을 제안하였다.

또한 공병천, 윤기웅 및 유승현(2018)은 성과감사에 대한 생소함과, 어려움 및 실패 위험을 감내하고 감사를 실시할 수 있도록 성과감사에 대한 내부 감사실적의 평가 및 인사관리 시스템을 개선하고, 감사에 대한 독립성, 정치적 중립성에 대한 문제가 제기되지 않도록 감사계획 수립 단계부터 엄격하고 투명한 감사 수행 체계를 갖출 것을 제안하였다.

본 연구는 앞선 연구들과 다르게 성과감사의 활성화 방안을 다루기 보다는 성과 감사의 실효성을 높이는 측면에서 정책이나 사업의 효과를 평가하는 방안을 확인해보고자 하였고, 이런 관점에서 연구를 수행한 국가인 미국 감사원의 성과감사의 목적과 범위, 추진 사례를 연구해 보았고, 우리 나라 성과감사가 놓치고 있거나 나아가야 할 방향에 대하여 알아보았다.

그러나 미국 GAO 에서 계량 경제학에서 이용되는 데이터 분석 기술을 이용하여 정책의 효과성 평가를 실시하고 있고, 각종 행정학 및 경제학 논문에서 정책이나 사업의 효과를 평가하고 있지만 감사의 영역으로 가져온다면 성과 평가 부분에 대한 데이터의 검증, 모델의 신뢰성, 모델 구성에 따른 다양한 가정 사항의 적정성 등 검토하여야 할 사항은 많을 것이다. 아울러 효과성 평가에 많이 이용하는 이중차분 분석이나 회귀 단절설계 기법을 적용할 수 있는 사례도 제한될 수밖에 없다.

데이터 분석기술의 발전은 단순히 계량적 분석기법을 고도화 시킨 것 외에도 계량적 분석기법에 이용될 데이터를 생성하는데도 많은 도움을 줄 수 있다. GIS 를 이용할 경우 지리 정보를 분석하거나 텍스트 마이닝을 통해 추출한 데이터를 별도로 가공하여 나온 결과를 계량 분석에 적용할 연속변수 또는 더미변수 형태로 치환하여 적용할 수 있다.

따라서 빅데이터, 기계학습 등 데이터 분석기술의 발전 동향을 지속적으로 파악하고, 장기적인 관점에서 계량 경제학에 기반한 데이터

분석 기법을 활용하여 사업이나 정책의 효과를 평가할 수 있는 방안을 마련하는 것이 필요하다.

이를 위해서는 데이터 분석 방법을 이용한 평가 방법론에 대한 교육이 필요할 뿐 아니라 효과의 평가를 위해 필요한 평가 질문 설정에서, 데이터 수집 및 가공, 데이터 분석, 처리까지 필요한 기반 여건을 마련하고, 평가 결과에 대한 신뢰성 확보를 위한 방안을 마련할 필요가 있다.

[참고문헌]

- 공병천, 윤기웅. (2017). 한국형 성과감사 모델 개발을 위한 기초연구.
- 공병천·윤기웅·유승현. (2018), “한국 성과감사의 성공적 정착에 대한 논의:SWOT 분석을 중심으로,” 「감사논집」 30: 6-28
- 김경민, 이의준, 박대권. (2010). 초·중·고등학교 수요가 서울시 구별 아파트 가격에 미치는 영향: 거주이주시기를 중심으로. 국토연구, 54, 99-113.
- 김예지, 이영성. (2014). 서울시 초·중·고등학교 학업성과와 아파트 가격 사이의 상호 작용. 국토계획, 49(6), 51-64.
- 김형아. (2019). 정책의 효과성 평가에 중점을 둔 성과감사 발전방안 연구. , 29(4), 211-239.
- 김홍률. (2009). 미국 GAO 성과감사의 현황과 과제. 학술대회 발표논문집, (), 1-39.
- 남형권, 서원석. (2016). 교육시설 인접성이 아파트 가격에 미치는 공간적 영향분석. 부동산연구, (26). 97-109.
- 손호성, 이재훈 (2018). 행정학·정책학 연구에서의 이중차분 추정기법의 활용과 쟁점.
- 안문영, 추준석. (2017). 전국단위 학력평가 성적 차이가 아파트 가격에 미치는 영향 - 울산광역시 사례. 부동산연구, (27) 63-76.
- 윤기웅, 공병천. (2019). 한국 성과감사의 과제 및 방향 - 성과감사 결과보고서 비교 분석을 통하여. 한국행정학회 학술발표논문집, (), 1371-1387.

윤기웅, 유승현. (2018). 감사종류의 변화 과정에 대한 행정사적 고찰: 성과감사를 중심으로. , (43), 1-32.

유승현. (2018). 성과감사의 개념화. , 32(3), 69-99.

유재석.(2011). [2011 초·중·고 학업성취도 평가] 서울 강남, 처음으로 전국 1 위자리 내렸다. 조선일보,(2011. 12. 2.)

정길영. (2017), “성과감사 발전방안,” 「계간 감사」 신년호: 71-75

진영남, 손재영. (2005). 교육환경이 주택가격에 미치는 효과에 관한 실증분석 : 서울시 아파트시장을 중심으로.

최재해.(2017). 감사원 역할의 변화 추이에 관한 연구: 1949-2016 년 감사결과 분석을 중심으로.

한승훈, 안혜선 (2019). 행정학 및 정책학 연구에서의 회귀단절 설계의 활용 가능성 탐구.

Charpentier, Arthur, Emmanuel Flachaire, and Antoine Ly.

“Econometrics and Machine Learning.” *Economie & statistique* 505.505d (2019): 147-169. Web.

Zheng, Eric et al. “When Econometrics Meets Machine Learning.” *Data and Information Management* 1.2 (2017): 75-83. Web.

Sendhil Mullainathan, and Jann Spiess. “Machine Learning: An Applied Econometric Approach.” *The Journal of Economic Perspectives* 31.2 (2017): 87-106. Web.

WuQian et al. “The Influences of High-Quality Educational Resources on Housing Prices – Evidences from Beijing’ s Fangshan District.” *Real Estate Review* (2017): 55-70

Rischar, Maxime et al. “A Bayesian Nonparametric Approach to Geographic Regression Discontinuity Designs: Do School Districts Affect NYC House Prices?” [arXiv.org](https://arxiv.org/abs/2018.08.01) (2018)

Stevenson, Betsey, and Justin Wolfers. “Bargaining in the Shadow of the Law: Divorce Laws and Family Distress.” *Quarterly Journal of Economics* 121.1 (2006): 267–288.

Black, Sandra. “Do Better Schools Matter? Parental Valuation of Elementary Education.” *The Quarterly Journal of Economics* 114.2 (1999): 577–599.

Cheng, Cheng, and Mark Hoekstra. “Does Strengthening Self-Defense Law Deter Crime or Escalate Violence?: Evidence from Expansions to Castle Doctrine.” *The Journal of Human Resources* 48.3 (2013): 821–854.

Cornwell, Christopher, and Scott Cunningham. “The Long-Run Effect of Abortion on Sexually Transmitted Infections.” *American Law and Economics Review* 15.1 (2013): 381–407.

Dhar, Paramita, and Stephen L Ross. “School District Quality and Property Values: Examining Differences Along School District Boundaries.” *Journal of Urban Economics* 71.1 (2012): 18–25

Figlio, David, and Maurice Lucas. “What’s in a Grade? School Report Cards and House Prices.” *NBER Working Paper Series* (2000)

Goodman, Allen C. “Hedonic Prices, Price Indices and Housing Markets.” *Journal of Urban Economics* (1978): 471–484.

Haskins, R., & Margolis, G. (2015). Show me the evidence: Obama's fight for rigor and results in social policy. Washington, D.C.: The Brookings Institution

Hansen, Kirstine. "Moving House for Education in the Pre-school Years." *British Educational Research Journal* 40.3 (2014): 483–500.

GAO(2006), Modernizing Accountability Organizations in Times of Fiscal Constraint. Washington D. C.: General Accountability Office.

GAO(Government Accountability Office). "Government Auditing Standards 2018 Revision" (2018)

GAO. "Discipline Disparities for Black Students, Boys, and Students with Disabilities" (2018)

GAO. "Income and Wealth Disparities Continue through Old Age" (2019)

GAO. "State Approaches Taken to Control Access to Key Methamphetamine Ingredient Show Varied Impact on Domestic Drug Labs" (2013)

GAO. "Assessment Methodology for Economic Analysis" (2018)

GAO. "Designing Evaluation" (2012)

GAO. "Program Evaluation – A Variety of Rigorous Methods Can Help Identify Effective Interventions" (2009)

[첨부] “학업성적이 우수한 초등학교가 주택 가격에 미치는 영향 분석” 시 이용한 R code

Data collection

```
library(tidyverse)
library(openxlsx)

getwd()
setwd("C:/Users/khaka/Documents/UT/CausalInference(s20)/04_project/03_Data")

# define a function for constructing APT price - transaction volume data
APT_fn <- function(file) {
  # read Excel file
  APT_temp <- read.xlsx(paste0("raw_data/", file), 1, startRow = 17, colNames = TRUE, rowNames = FALSE)
  # change column names from Korean to English
  names(APT_temp) <- c("add1", "add2_1", "add2_2", "APT_name", "area", "C_yr_m", "C_d", "Price", "floor", "age")
  # add year, district variables
  APT_temp$year <- substr(APT_temp$C_yr_m, 1, 4)
  #APT_temp$dstrt <- substr(APT_temp$add1, 7, 9)
  #APT_temp$dong <- substr(APT_temp$add1, 11, 13)

  add_temp <- data.frame(do.call('rbind', strsplit(as.character(APT_temp$add1), split = " ")))
  APT_temp$dstrt <- add_temp$X2
  APT_temp$dong <- add_temp$X3
  # categorize area of apartments
  APT_temp$area_CT <- round(APT_temp[,5], digits = -1)
  # group by APTments and area with median price and transaction volume
  APT_temp %>%
    group_by(year, dstrt, dong, add2_1, APT_name, age, area_CT) %>%
    summarize(mid_price = median(Price), numb_contrat = n())
}

# for initial setting of data frame
```

```

#data_list <- list.files(pattern="APT*")
APT_complex <- APT_fn("APT06.xlsx")

# adding other years
data_list <- list.files(path = "raw_data", pattern = "APT*")[-1]

for (file in data_list) {
  temp <- APT_fn(file)
  APT_complex <- rbind(APT_complex, temp)
}
# apt id with district, name of APT and area
APT_complex$id <- paste0(APT_complex$dong,APT_complex$APT_name,APT_complex$area_CT)

# drop the insignificant data

## drop the apt complex which has less than 3 contracts

APT_1 <- subset(APT_complex, numb_contrat >= 3)

## yearly transaction
yr_price <- xtabs(mid_price ~ id + year, APT_1)
yr_price[yr_price == 0] <- NA
yr_price <- na.omit(yr_price)
full_yr <- row.names(yr_price)

chk <- data.frame(id=full_yr, chk = 1)

APT_2 <- merge(APT_1, chk, by = "id", all = TRUE) %>% na.omit()

length(full_yr)
summary(as_factor(APT_2$chk))

## school check
### The list of elementary schools with good performance in 2011 in Seoul
### 50 highest scoring schools and top 10% of each district("GU" in Korean) => top school
### Using school district of the top schools,
### make address lists of apartment complex withing the school districts -> school.xlsx
school_chk <- read.xlsx("school.xlsx", 1, colNames = TRUE, rowNames = FALSE) %>% unique()

# key variable to merge with school district data

```

```

APT_2$add <- paste(APT_2$dong,APT_2$add2_1)

# create dummy for information about good schools. The information was
  released in late 2011.
APT_3 <- merge(APT_2, school_chk, by = "add", all.x = TRUE)
APT_3$year <- as.numeric(APT_3$year)
APT_3[is.na(APT_3)] <- 0
# dummy for treatment
APT_3$GS <- ifelse(APT_3$year>=2012,
                  ifelse(APT_3$chk.y==1, 1, 0),0)

# other covariates
Xvars <- read.xlsx("covariates.xlsx", 1, colNames = TRUE, rowNames = TR
UE) %>%
  as.matrix() %>%
  as.data.frame.table(responseName = "price_index")

Xvars <- read.xlsx("covariates.xlsx", 2, colNames = TRUE, rowNames = TR
UE) %>%
  as.matrix() %>%
  as.data.frame.table(responseName = "volume") %>%
  merge(Xvars, by = c("Var1", "Var2"))

Xvars <- read.xlsx("covariates.xlsx", 3, colNames = TRUE, rowNames = TR
UE) %>%
  as.matrix() %>%
  as.data.frame.table(responseName = "households")%>%
  merge(Xvars, by = c("Var1", "Var2"))

Xvars$district_id <- as_factor(Xvars$Var1) %>% as.numeric()
names(Xvars)[1:2] <- c("dstrt","year")

# APT complex id and weight(total sales)
i_vars <- APT_3 %>%
  group_by(id) %>%
  summarise(Tot_Ctrt_wt = sum(numbr_contrat))
i_vars$APTid <- 1:nrow(i_vars)

# combine all variables
APT_4 <- merge(APT_3, i_vars, by = "id", all.x = TRUE)
APT_4 <- merge(APT_4, Xvars, by = c("dstrt","year"), all.x = TRUE)

# log values
APT_4 <- APT_4 %>% mutate(
  l_m_price = log(mid_price),
  l_households = log(households),

```

```

l_volume = log(volume),
l_price_index = log(price_index)
)
# write file
write.csv(APT_4, file = "APTandSchool.csv")

```

Regression

```

rm(list=ls())

library(tidyverse)
library(haven)
library(estimatr)
library(kableExtra)
library(lfe)
library(stargazer)

setwd("C:/Users/khaka/Documents/UT/CausalInference(s20)/04_project")

##### load data #####
APTprice <- read.csv("03_Data/APTandSchool.csv", header = TRUE)

##### add new variables #####
# detailed treatment 1~2 year, 3~4 year, 5~6 year
APTprice <- mutate(APTprice,
                    GS1 = ifelse(GS==1&(year==2012|year==2013),1,0),
                    GS2 = ifelse(GS==1&(year==2014|year==2015),1,0),
                    GS3 = ifelse(GS==1&(year==2016|year==2017),1,0)
)
# classification by area
summary(as_factor(APTprice$area_CT))
APTprice <- mutate(APTprice,
                    sqM = ifelse(area_CT <= 40,"XS", # ~ 45
                                extra small
                                ifelse(area_CT <= 60, "S", # 45 ~
                                        65 small
                                        ifelse(area_CT <= 80, "M",# 65 ~
                                                85 medium
                                                ifelse(area_CT <= 130, "L
", "XL") ) ) ) )
# 85 ~ 135 large / 135
~ extra large

```

```

# classification by price
APTprice <- mutate(APTprice,
                  priceC = ifelse(mid_Uprice <= 412.8, "low", # 1st qu
antile
                                ifelse(mid_Uprice <= 638.5, "mid", "h
igh")))
                                # 2nd&3rd quantile /4th quantil
e

##### modify dataset #####
# drop extra small & large apartments
APTpriceM <- subset(APTprice, APTprice$sqM != "XS" & APTprice$sqM != "XL")

#APTpriceM <- mutate(APTpriceM,
#                    age = 2020 - age,
#                    chk.y = as_factor(chk.y),
#                    sqM = as_factor(sqM),
#                    priceC = as_factor(priceC),
#                    year = as_factor(year),
#                    GS = as_factor(GS),
#                    GS1 = as_factor(GS1),
#                    GS2 = as_factor(GS2),
#                    GS3 = as_factor(GS3),
#)

##### data statistics #####
stargazer(APTpriceM[,c("year", "age", "area_CT", "mid_price", "mid_Uprice",
                      "numb_contrat")], type = "html", rownames = FALSE)

glimpse(APTpriceM[,c("year", "age", "area_CT", "mid_price", "mid_Uprice",
                    "numb_contrat")])

summary(APTpriceM[,c("year", "dstrt", "id", "chk.y", "sqM", "priceC")]) %>%
  kable() %>% kable_styling(fixed_thead = T, font_size = 11)

##### regression #####

# simple

reg1 <- felm(l_m_Uprice ~ GS | year + APTid|0|APTid ,

```

```

weights = APTprice$Tot_Ctrt_wt, data = APTprice)

reg2 <- felm(l_m_Uprice ~ GS + l_price_index | year + APTid|0|APTid ,
            weights = APTprice$Tot_Ctrt_wt, data = APTprice)

reg3 <- felm(l_m_Uprice ~ GS1 + GS2 + GS3 | year + APTid|0|APTid ,
            weights = APTprice$Tot_Ctrt_wt, data = APTprice)

reg4 <- felm(l_m_Uprice ~ GS1 + GS2 + GS3 + l_price_index | year + APTid|0|APTid ,
            weights = APTprice$Tot_Ctrt_wt, data = APTprice)

## table
stargazer(reg1, reg2, reg3, reg4, type = "html",
          omit = c("yd", "ysq", "yp", "l_price_index"),
          add.lines = list(c("District Price index", "No", "Yes", "No",
                             "Yes"),
                           c("Years later Treatment", "No", "No", "Yes"
                              , "Yes") ) )

## plot control group v.s. treatment group
C_T <- APTpriceM %>% group_by(year, chk.y) %>%
  summarise(price = weighted.mean(mid_Uprice, TOT_Ctrt_wt)) %>%
  as.data.frame() %>%
  mutate(TorC = ifelse(chk.y==1, "Treated", "Control"))

ggplot(data = C_T, aes(x = year, y = price, color = TorC)) +
  geom_line() + geom_point() +
  geom_vline(xintercept = 2012, linetype = "dashed", color = "blue") +
  theme_minimal() +
  xlab("Years before and after Good Schools information") + ylab("Unit
Price")

#####other plot #####
CTR <- APTpriceM %>%
  filter(chk.y == 0) %>%
  group_by(year) %>%
  summarise(CTR = weighted.mean(mid_Uprice, TOT_Ctrt_wt)) %>%
  as.data.frame()

Trt <- APTpriceM %>%
  filter(chk.y == 1) %>%

```

```

group_by(year) %>%
  summarise(TRT = weighted.mean(mid_Uprice, TOT_Ctrt_wt)) %>%
  as.data.frame()

C_T <- merge(CTR, Trt, by = "year") %>%
  mutate(DIFF = TRT - CTR)

ggplot(data = C_T, aes(x = year)) +
  geom_line(aes(y = CTR)) + geom_line( aes( y = TRT), color = "red") +
  geom_bar(aes(y = DIFF),stat='identity', fill = "skyblue1") +
  geom_vline(xintercept = 2012, linetype = "dashed", color = "red") +
  theme_minimal() +
  xlab("Years before and after Good Schools information") +
  ylab("Unit Price") + scale_x_continuous(breaks = c(2006:2017)) + them
e(legend.text = )

##### other fixed effect #####
yr_lp = unique(APTpriceM$year) %>% sort()

# create regin-by-year dummy
dstrt_lp = unique(APTpriceM$district_id) %>% sort()

for (yr in yr_lp){
  for (dt in dstrt_lp) {
    x <- paste0("yd",yr,dt)
    APTpriceM[[x]] <- ifelse(APTpriceM$district_id==dt&APTpriceM$year==
yr,1,0)
  } }

# create area - year fixed effect
sqM_lp = c("S","M","L")

for (yr in yr_lp){
  for (sq in sqM_lp) {
    x <- paste0("ysq",yr,sq)
    APTpriceM[[x]] <- ifelse(APTpriceM$sqM==sq & APTpriceM$year==yr,1,0
)
  } }

# create price - year fixed effect
price_lp = c("low","mid","high")

for (yr in yr_lp){
  for (p in price_lp) {

```

```

x <- paste0("yp",yr,p)
APTpriceM[[x]] <- ifelse(APTpriceM$priceC==p & APTpriceM$year==yr,1
,0)
} }

# formula
## get colnames
dstrtfixed <- colnames(
  APTpriceM %>%
  select(starts_with("yd20")) )

sqfixed <- colnames(
  APTpriceM %>%
  select(starts_with("ysq20")) )

pfixed <- colnames(
  APTpriceM %>%
  select(starts_with("yp20")) )

##
formula1 <- as.formula(
  paste("l_m_Uprice ~ GS +",
    paste(
      paste(dstrtfixed, collapse = " + "),
      paste(sqfixed, collapse = " + "),
      paste(pfixed, collapse = " + "), sep = " + "),
    "| year + APTid | 0 | APTid"
  ) )

formula2 <- as.formula(
  paste("l_m_Uprice ~ GS + l_price_index + ",
    paste(
      paste(sqfixed, collapse = " + "),
      paste(pfixed, collapse = " + "), sep = " + "),
    "| year + APTid | 0 | APTid"
  ) )

formula3 <- as.formula(
  paste("l_m_Uprice ~ GS1 + GS2 + GS3 + ",
    paste(
      paste(dstrtfixed, collapse = " + "),
      paste(sqfixed, collapse = " + "),
      paste(pfixed, collapse = " + "), sep = " + "),
    "| year + APTid | 0 | APTid"
  ) )

```

```

formula4 <- as.formula(
  paste("l_m_Uprice ~ GS1 + GS2 + GS3 + l_price_index + ",
        paste(
          paste(sqfixed, collapse = " + "),
          paste(pfixed, collapse = " + "), sep = " + "),
        "|year + APTid | 0 | APTid")
  )

## regression

regf1 <- felm(formula1, weights = APTpriceM$Tot_Ctrt_wt, data = APTpriceM)
#summary(reg1)
regf2 <- felm(formula2, weights = APTpriceM$Tot_Ctrt_wt, data = APTpriceM)
#summary(reg2)
regf3 <- felm(formula3, weights = APTpriceM$Tot_Ctrt_wt, data = APTpriceM)
#summary(reg3)
regf4 <- felm(formula4, weights = APTpriceM$Tot_Ctrt_wt, data = APTpriceM)
#summary(reg4)

## table
stargazer(regf1, regf2, regf3, regf4, type = "html",
           omit = c("yd", "ysq", "yp", "l_price_index"),
           add.lines = list(c("District-
Year dummy", "Yes", "No", "Yes", "No"),
                           c("District Price index", "No", "Yes", "No",
"Yes"),
                           c("Years later Treatment", "No", "No", "Yes"
, "Yes"),
                           c("Year specific trends", "Yes", "Yes", "Yes"
, "Yes"))) )

```