

“ 재정사업사전검증 체계 강화를 위한 연구 - RCT 도입방안을 중심으로* ”

I 서론: 배경 및 문제점

새로운 재정사업이 충분한 사전 점검 없이 정치적인 요구에 의해 급격하게 도입되는 사례가 증가하면서, 시행착오로 인한 사회적 비용이 증가하고 있다. 자원조달 가능성이나 소요액의 부정확한 추정에 근거한 문제도 부각되고 있으며, 사업 계획의 부적정성으로 인해 예기치 않은 부작용이 노정되는 사례도 발생하고 있다. 대형 국책사업을 중심으로 대형 신규 사업의 적정성을 심사하는 예비타당성 조사제도가 운영되고 있으나, 노동이나 복지 등 새롭게 확대되고 있는 사업영역에 대한 점검 기능은 만족할 만한 수준이 아니라고 판단된다. 고용이나 복지 관련 사업까지 예비타당성 조사제도를 확대하는 시도가 이루어지고 있으나, 기존의 예비타당성 조사제도의 한계 및 부작용이 부각되는 상황에서, 예비타당성 조사제도의 확대 시행의 실효성에 대한 의문도 존재하고 있다.

현행 예비타당성조사 제도는 정형화된 사업이나 정책에 대해서는 비교적 의미 있는 점검이 가능하지만, 새롭게 추진되는 사업이나 정책의 쟁점이나 효과성에 대한 점검은 한계가 있다. 따라서 비교적 정형화된 사업과 사업의 효과성을 추정하는 주요 패러미터에 대한 신뢰성 있는 추정치가 존재할 때만 적용 가능한 방법이라고 볼 수 있다. 현재 복지사업에 대해서도 예비타당성 조사가 확대되고 있으나, 사업 계획의 적정성 점검수단으로는 의미가 있으나 실제 사업의 효과성이나 사업 운영의 주요 쟁점에 대한 점검 수단으로서는 한계가 있을 수밖에 없다.

* 본 보고서는 저자의 2015년도 기본연구과제 “재정사업사전검증 체계 강화를 위한 연구 - RCT 도입방안을 중심으로”의 내용을 요약·정리한 것임

예를 들어, 최근 시행 예정인 저소득층 영아 기저귀·분유값 지원 사업¹⁾의 경우 예비타당성 조사로는 정책의 효과를 정확히 예측하기 어렵다. 이 사업은 저소득층의 경제적 부담이 출산율을 낮추는 원인이기 때문에 양육에 필요한 기저귀와 분유값을 지원하여 출산율을 높이려는 국정과제이다. 보건복지부는 사업의 정책 효과를 시범사업을 통하여 검증하기 위하여 50억원의 시범사업예산을 요청했지만 중기재정소요 500억원 이상인 복지사업의 경우 예비타당성 조사를 받아야 하는 「국가재정법」 규정으로 인하여 시범사업은 보류되고, 복지사업에 대한 예비타당성 조사 후 사업 시행이 결정되었다.

그러나, 위 사업에 대한 예비타당성 조사의 결과로 얼마나 신뢰성 있는 정책효과의 추정이 가능할지 의문이다. 편익에 대한 명확한 패러미터가 존재하는 SOC사업과 달리 사업 시행의 결과로 얻어지는 출산율 증가의 편익을 예측하기가 불가능하기 때문이다. 사업의 시행의 결과로 저소득층 가정에서 출산율이 증가하는지를 정확히 예측할 수 있는 방법은 소규모의 경우라도 실제 정책을 시행하는 것이 유일하다. 따라서 Randomized Control Trial(RCT)²⁾를 통한 진실험적 평가방법은 정책효과를 예측하는 유용한 평가방법이 될 수 있다. 시범사업에서 일정 소득기준 이하의 정책대상자를 신청을 통해 받은 후 추첨을 통하여 처치집단과 비교집단으로 구분하고 처치집단에 기저귀와 분유를 지원한 후 실제 처치집단에서 출산율이 비교집단에 비해 얼마나 증가하는지를 비교함으로써 정확한 정책효과를 추정할 수 있기 때문이다.

위 사업의 경우 예비타당성 조사에서 통과되면 향후 장기간에 걸쳐 막대한 재원이 소요되는 현실을 감안할 때, 정확한 평가방법론에 의한 재정사업 사전검증체계의 제도화가 시급한 정책과제라 할 수 있다. 또한 과거 진행되었던 청년층 취업지원 사업, 대학 연구지원 사업, 중소기업 지원사업 등 수많은 복지, 고용관련 사업들의 경우도 사업의 효과에 대한 정확한 사전 검증 없이 시작되어 큰 정책효과를 거두지 못하고 자원만 낭비되고 있는 실정이다.

따라서, 우리나라에서 최근 급속하게 확대되고 있는 복지나 고용관련 사업의 경우, 사회실험 방식의 RCT 정책평가를 통해서 사업의 쟁점과 효과성을 파악하여 정책을 추진하는 것이 바람직하다고 판단된다. 외국에서도 고용이나 복지사업에 대해서는 RCT 평가방식이 적용되고 있으며, 정형화된 예비타당성 조사 방식의 적용은 찾아보기 어렵다. RCT가 적절하게 적용될 수 있는 재정사업에 대해서는 재정사업의 도입 이전에 사회실험을 거치도록 하는 방안 마련이 필요하다.

1) 미국의 경우 'WIC(Woman, Infant and Child) Program'이란 명칭으로 활성화되어 있다.

2) 무작위 추출을 통해 동질적인 처치(실험)집단과 비교(통제)집단을 구성하여 정책의 효과를 평가하는 방법은 Randomized Trial Control(RCT) 외에, 처치-비교 집단 무작위 추출 실험, 사회실험(Social Experiment), 현장실험(Field Experiment), 정책실험(Policy Experiment) 등 다양한 용어로 사용되고 있다.

II 조사 및 분석결과

1. RCT (Randomized Control Trial) 개념

프로그램, 특히 정부가 수행하는 재정사업 프로그램은 평가(evaluation)를 통해 그 성과를 측정하고 더 나아가 프로그램 설계를 개선할 수 있다. 본 보고서에서 의미하는 평가란 프로그램(또는, 정책 혹은 프로젝트)이 효과적이었는지, 더 나아가 그러한 효과의 크기가 어느 정도인지 파악하는 일련의 과정을 의미한다. 프로그램은 그 도입단계에서부터 달성해야 할 목적을 지니고 있기 마련인데, 효과성 평가는 이러한 목적을 달성했는지 여부를 점검하는 것으로 프로그램 평가의 주요 목적이다. 더 나아가 그 효과의 크기를 추정하는 것은 추후 효율성 평가 등에 중요한 근거를 제시한다.

예를 들어, 국가가 수학과목 학업성취도가 낮은 학생들을 대상으로 운영하는 보충수업 프로그램을 생각해 보자. 실제로 교육부가 2015년 2월에 발표한 '제2차 수학교육 종합계획(2015~2019)'과 2014년 12월에 발표한 '사교육 경감 및 공교육 정상화 대책'에 따르면 학습부진 학생들의 효율적 자기주도 학습을 지원하기 위해 맞춤형 수학멘토링이 운영될 계획이며, 이 프로그램을 통해 수학에 대한 흥미와 자신감을 부여하고 학습 결손을 보정할 수 있기를 기대하고 있다. 이러한 보충수업 프로그램을 담당하고 있는 공무원은 해당 프로그램이 학습부진 학생들의 학업성취도(또는 시험성적) 향상이라는 목표를 달성했는지, 또한 그렇다면 그 효과의 크기는 (투입된 비용 대비) 어느 정도인지 알고 싶을 것이다. 이러한 질문이 중요한 이유는 그 결과에 따라 프로그램을 개선할 수 있으며 프로그램 연장 등으로 인해 추가 예산이 필요한 경우 중요한 근거자료가 되기 때문이다.

재정사업 프로그램을 운영하는 공무원 입장에서 이렇게 당연한 질문에 답하는 것은 일반적으로 쉬운 작업이 아니다. 위의 예에서 보충수업 프로그램의 효과성을 판단하기 위해서는 해당 프로그램의 수혜를 받은 집단과 그렇지 않은 집단의 성과를 비교해야만 한다. 하지만 보충수업에 참여하는 것은 전적으로 개인의 선택에 달려 있기 때문에 프로그램의 수혜를 받은 집단과 그렇지 않은 집단 사이의 비교는 주의를 필요로 한다. 만약 자기 주도 학습능력이 상대적으로 뛰어난 학생들이 주로 이 프로그램에 지원했다면 우리가 관찰할 수 있는 두 집단 간의 성과의 차이는 프로그램 본연의 성과와 프로그램에 지원하기 이전의 개인의 역량의 차이가 결합된 것이기 때문이다. 만약 후자의 크기가 크다면 이러한 집단 간의 단순 비교는 프로그램의 진정한 성과를 왜곡시킬 수 있다.

위의 예에서 살펴본 바와 같이 프로그램의 효과성을 측정하는 것은 프로그램의 성과를 이해하는 데 중요한 작업이지만 그 해답을 찾는 과정은 쉽지 않다. 프로그램의 수혜를 받은 집단과 그렇지 않은 집단을 단순히

비교한 결과는 프로그램 성과와 두 집단 사이의 본연적인 성과의 차이가 포함되었기 때문이다. 더군다나 프로그램 설계를 위해 효과성을 프로그램 시행 이전에 측정하는 일은 더욱 어려운 작업이다.

이러한 측면에서 RCT(randomized controlled trials)는 프로그램 시행 이전에 프로그램의 성과를 엄밀히 평가할 수 있는 방안 중 하나로서 최근 해외 여러 국가에서 각광받고 있다. RCT란 프로그램의 처치집단(수혜집단)과 비교집단을 무작위 방식을 통해 구성함으로써 두 집단 간의 성과를 비교할 때에 집단 사이의 본연적인 성과의 차이가 없도록 설계한 프로그램 평가방법이다.

2. 재정사업 성과평가에서 RCT의 적용가능성

가. 사전평가

사회실험을 가장 효과적으로 적용할 수 있는 평가단계는 사전평가이다. 본 사업 시행 전에 실험을 통하여 미래의 효과를 평가하고 평가결과에 근거하여 사업의 시행 여부를 결정할 수 있기 때문이다. 우리나라에서 사전평가는 예비타당성 조사와 시범사업의 형태로 실시되어 왔다. 예비타당성 조사는 총사업비가 500억원 이상이면서 국가재정지원 규모가 300억원 이상인 건설사업, 정보화사업, 국가연구개발사업과 중기재정지출이 500억원 이상인 사회복지, 보건, 교육, 노동, 문화 및 관광, 환경보호, 농림해양수산, 산업·중소기업 분야의 사업에 걸쳐 실시한다.³⁾ 예비타당성 조사가 대규모 재정사업을 효과를 계량적으로 추정하여 무분별한 재정사업을 신설을 막는 데 기여해 왔다는 평가를 받고 있지만, 추정의 정확성은 여전히 많은 논란이 되고 있다. 특히, 건설이나 SOC사업 외에, 교육, 복지, 노동, 문화 사업 등은 B/C나 B/E분석과 같은 계량기법으로 편익(Benefit)이나 효과(Effectiveness)를 정확히 추정하기 어렵다.

이러한 한계를 보완하기 위하여 예비타당성 조사는 경제적 효과 외에 정책적 효과, 지역균형발전, 기술적 효과를 함께 평가한 후 각 평가항목에 대한 가중치를 부여하여 최종 대안을 선택하는 AHP 의사결정기법에 의해 사업의 시행 여부를 결정한다. 비교적 편익이 명확히 측정되는 건설사업과 달리 기타 재정사업의 경제성 효과추정의 정확성은 추정의 어려움으로 인하여 떨어진다. 게다가 <표 1>에서 보듯 기타 재정사업의 경우 정책적 분석의 비율이 50~70%를 넘는다. 비계량적으로 측정되는 정책적 분석의 가중치가 50%를 넘기 때문에 기타 재정사업의 예비타당성 조사 예측 분석의 정확성은 현저하게 떨어질 가능성이 높다. 이러한 문제에도 불구하고 500억원 미만의 고용, 복지, 문화, 예술 등 소규모 비 건축사업에 간이 예비타당성 도입을 검토하고 있어 정확한 사전평가방법론의 정립이 시급한 실정이다.

3) 「국가재정법」 제38조-법제처 국가법령정보센터

〈표 1〉 평가항목의 가중치 범위

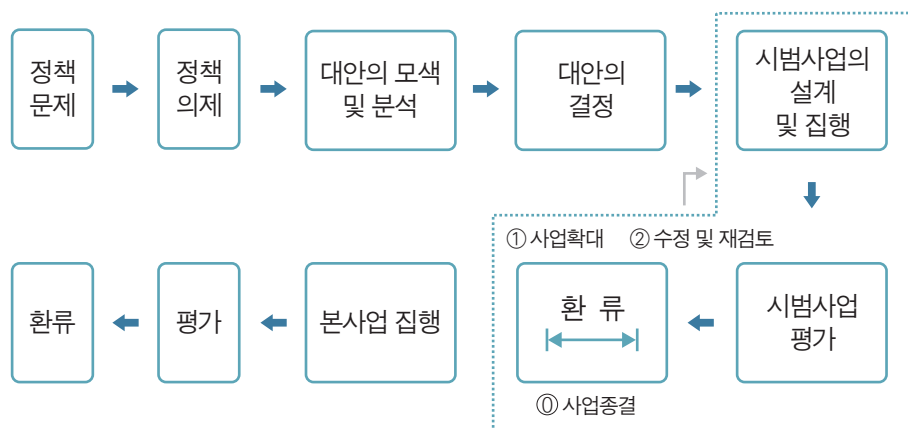
구분		경제성 분석	정책적 분석	지역균형 발전	기술성 분석
건설사업		40~50%	25~35%	20~30%	-
기타 재정사업	B/C 분석 시	40~50%	50~60%	-	-
	E/C 분석 시	30~40%	60~70%	-	-
R&D 정보화	B/C 분석 시	40~50%	20~30%	-	30~40%
	E/C 분석 시	30~40%		-	40~50%

자료: 「2013 KDI 예비타당성 조사 지침」

이러한 예비타당성 조사의 한계로 인하여 개별 부처에서 많이 사용되고 있는 시범사업(Pilot Projects)이다. 시범사업은 이삼열 외(2009, p.13)는 시범사업의 정의를 “본 정책의 전면적인 집행에 앞서 정책의 효과성을 검증하고 공공자원의 낭비 등을 피해 정책의 효율성을 높이는 것을 목적으로, 특정 정책의 효과나 작동기제를 사전에 측정 또는 관찰하기 위한 엄격한 사전설계를 바탕으로 집행되는 비교적 작은 규모의 사업”이라고 정의한다.

시범사업은 엄격한 사전설계에 의하여 본 정책을 실시하기 전에 작은 규모로 사업을 실시하고 그 효과에 기반하여 본사업을 시행 여부를 결정한다는 점에서 가장 효과적인 증거기반 사업평가(Evidence based policy) 수단이 될 수 있다. 특히, 시험사업은 사회실험을 위한 좋은 환경을 제공한다. 시범사업 내에서 처치 집단과 비교집단을 구분하거나 시범사업의 정책 수혜자와 시범사업의 수혜자가 아닌 일반 정책대상자를 처치집단과 비교집단으로 설정하여 효과를 비교할 수 있기 때문이다.

[그림 1] 시범사업이 고려된 정책과정



자료: 이삼열 외(2009), p. 16

시범사업에서의 사회실험의 적용 가능성에도 불구하고 시범사업에서 엄밀한 사회실험을 통한 평가한 경우는 매우 드문 편이다. 예를 들어 이삼열 외(2009)는 시범사업의 시행이 법으로 명문화⁴⁾된 보건복지부에서 실시된 2002~2009년까지 시범사업의 수와 평가방법을 분석하였다. 분석결과 보건복지부는 7년 동안 62개의 시범사업을 실시했는데 그 중 21개의 시범사업의 평가서가 존재하였다.

이러한 시범사업의 평가에 RCT와 같이 엄밀하게 설계된 처치-비교집단 사전사후 비교설계방식으로 평가한 시범사업 평가⁵⁾는 1개⁶⁾만 존재(전체 2%) 하였다. 나머지 대다수는 단일집단 사후설계(81%)와 단일집단 사전사후 비교설계(17%)로 평가였다. 이는 많은 경우 시범사업의 평가가 제대로 실시되지 않고 있으며, 평가를 진행하더라도 RCT와 같은 엄밀한 정책평가가 실시되지 않고 있음을 보여준다. 이런 점에서 우리나라의 시범사업은 사전검증을 통한 사업의 효율성 제고라는 시범사업 본래의 취지를 살리지 못하고 있는 실정이다.

〈표 2〉 보건복지부 시범사업 시행횟수

(단위: 개)

연도	시범사업의 개수	시범사업 평가서 수
2002	4	0
2003	8	1
2004	4	2
2005	13	7
2006	6	2
2007	7	4
2008	20	5
총계	62	21

자료: 이삼열 외(2009), p. 32

따라서, 예비타당성 조사와 시범사업의 한계를 극복하기 위해서는 재정사업 사전평가에 있어 RCT에 의한 실험평가를 전제로 한 시범사업을 제도화 하는 것을 검토할 필요가 있다. 특히, 대규모 예산이 들어가는 비건설분야 재정사업의 경우 정확한 편익의 예측이 어렵기 때문에 RCT 평가 기반 시범사업을 실시한 후 정확한 정책의 효과가 있을 때만 본사업을 실시하는 것을 검토할 수 있다. 소규모 사업의 경우에도 기존의 활용

4) 「보건의료기본법」 제44조(보건의료 시범사업) 1. 국가 및 지방자치단체는 새로운 보건의료 제도를 시행하기 위하여 필요한 경우에는 시범사업을 실시할 수 있다.
2. 국가 및 지방자치단체는 제1항의 규정에 의한 시범사업을 실시한 때에는 그 결과를 평가하여 새로이 시행될 보건의료제도에 반영해야 한다(법제처 국가법령센터)

5) 이삼열 외(2009)

6) 진실형 방식에 의해 평가는 과학적인 실험설계가 가능한 병원에서의 금연프로그램의 효과에 대해 행해졌다(이삼열 외, 2009)

가능한 데이터가 있을 경우에는 시범사업의 저비용 RCT 평가 설계를 통해 정책의 효과를 정확히 검증할 수 있을 것이다.

나. 정책분야

RCT와 같은 실험평가가 적용될 수 있는 정책분야는 넓다. 미국 OMB⁷⁾는 사회실험의 평가가 적용될 수 있는 전제조건을 세 가지로 제시하고 다음의 조건만 충족된다면 대부분의 정책분야에서 RCT와 같은 실험평가의 실시가능성을 밝히고 있다.

- (1) 사업의 참가자와 비참가자가 무작위 추출에 의해 두 집단 또는 다중 집단으로 구분될 수 있으며, 이런 무작위 추출에 의해 구성될 샘플의 크기가 충분히 큰 사업
- (2) 각 집단 간에 다른 사업추진방식을 적용할 수 있는 사업
- (3) 각 집단별로 다른 사업추진방식의 적용으로 인하여 발생한 사업의 결과(Outcomes)를 측정할 수 있는 사업

이론적으로 위와 같은 전제조건이 충족된다면 모든 정부 정책사업에서 사회실험을 통한 정책평가가 가능하지만 현실적으로 사회실험은 교육이나 복지사업에서 주로 시행되어 왔다. <표 3>과 같이 OMB는 미국에서 RCT가 시행되어왔던 대표적인 정책분야를 소개하였는데, 주로 교육, 보건, 복지, 치안분야에서 사용되어 오고 있음을 알 수 있다.

〈표 3〉 RCT 기반 평가 정책분야

(단위: 회)

정책분야	교육	보건	복지	치안	재정	기타
실행횟수	3	5	4	3	1	2

자료: OMB, "What Constitutes Strong Evidence of a Program's Effectiveness?," 2004, pp. 9~10

우리나라의 경우에도 교육, 보건복지, 고용·노동, 문화·체육, 재정, 해외원조 사업 등에서 사회실험의 적용이 가능할 것으로 판단된다. 이들 분야의 재정사업 평가에서 처치집단과 비교집단의 구분이 가능하고 객관적 데이터를 통해 정책의 효과를 명확하게 파악할 수 있기 때문이다. 먼저, 교육분야의 경우, 학교정책, 교육 방법, 학생상담 프로그램에서 실험적 평가방법론의 적용이 가능할 것으로 보인다. 구체적으로 최근 시행된 자유학기제나 방과 후 돌봄서비스와 같은 사업의 경우 시범사업을 실시하여 정책 시행 학교와 기존 학교를 처치그룹과 비교그룹으로 구분한 후 효과가 있을 경우 전체 학교로 확대하는 방법을 검토할 수 있을 것이다.

7) OMB, "What Constitutes Strong Evidence of a Program's Effectiveness?," 2004, p. 8

보건·복지 분야의 경우도 RCT와 같은 실험평가가 활발히 적용될 수 있는 분야이다. 각종 보건 제도설계, 보건서비스 및 교육 프로그램, 저소득층 공적부조 및 자활 관련 사업은 무작위 추출에 의한 처치-비교집단의 실험평가 설계가 가능하다. 또한 실제로 해외국가에서는 이 분야에서 가장 활발하게 정책평가가 활성화되고 있다. 적용 가능한 예로서 최근 정책의 효과에 의문이 발생하고 있는 무상보육 서비스의 경우, 보편적 무상보육 외에 차등보육지원 프로그램을 개발하여 차별화된 지원 프로그램의 효과를 비교하는 실험설계를 통해 정책의 효과를 정확히 검증할 수 있을 것이다.

고용·노동 분야 또한 RCT와 같은 사회실험 방식의 정책평가가 유용하게 활용될 수 있는 분야이다. 현재 우리나라에는 다양한 고용 및 노동 지원 프로그램이 운영되고 있으며 이 분야의 공공지출은 전체 GDP 대비 0.77%⁸⁾에 이르고 있다. 이러한 높은 지출에도 불구하고 이 분야에 대한 평가는 엄밀하게 이루어지지 않고 있는 편이다. 예를 들어 사업의 효과에 대해 많은 논란이 있는 고용보조 청년인턴사업의 효과는 RCT와 같은 실험평가를 정확하게 측정될 수 있을 것이다. 청년인턴을 고용하길 원하는 기업의 신청을 받은 후 최종 지원 대상 기업을 무작위 추첨으로 선정한 후 선정기업과 미선정기업의 인턴채용비율과 고용증가율을 비교하는 방법으로 사업의 효과를 정확하게 검증할 수 있을 것이다.

우리나라의 경우 문화·체육분야도 RCT에 의한 사회실험평가가 적절히 활용될 수 있는 분야이다. 해외와는 달리 최근 정부는 국민의 문화생활 향유와 이 분야의 고용확대를 위해 다양한 문화 및 체육 육성 지원사업을 진행하고 있기 때문이다. 국민의 문화 및 체육활동 참가율을 높이기 위한 문화 바우처 사업이나 생활체육 프로그램이 대표적 예다. 예를 들어 이와 같은 문화체육분야 정책사업의 효과성을 검증하기 위하여 문화 바우처의 정책수혜자를 무작위 추첨방식으로 선정하여 바우처를 지급받은 사람들과 지급받지 않은 사람들의 문화이용률 수준을 비교하는 방식으로 문화바우처의 효과를 검증할 수 있다.

마지막으로 정책금융, 세제, 해외원조 또한 RCT에 의한 진실실험평가가 적용될 수 있는 사업분야이다. 이미 해외에서는 이 분야에서 실험을 통한 정책평가가 활발히 진행되었다. 체납자의 납부율을 높이기 위한 체납 고지서 메시지 및 디자인 실험은 영국과 미국에서 실시된 이 분야의 대표적인 실험의 예이다. 특히, 다양한 해외원조사업을 수행하고 있는 UN이나 월드비전과 같은 국제기구들은 개발도상국에 대한 원조의 효과성을 정확하게 측정하기 위하여 RCT에 의한 실험평가를 실시하고 있다. 우리나라의 경우도 한국국제협력단(KOICA)를 통하여 다양한 해외원조사업을 실시하고 있기 때문에 원조 수원국의 동의를 받아 해외 원조 프로그램 수원국에서 무작위 추출 처치-비교집단 실험평가⁹⁾를 통해 특정 원조사업의 효과를 정확히 검증할 수 있을 것이다.

8) 홍승현·원종학(2013, p. 30)

9) RCT에 의한 실험평가 가능성을 탐색하기 위해 KOICA는 MIT대학의 J-Poverty Action Lab를 통해 RCT관련 단기 교육연수를 진행하였다.

〈표 4〉 RCT에 의한 사회실험이 적용 가능한 정책분야

분야	적용가능 사업 분야	사업의 예
교육	학교관련 정책 교수법 및 교육 방식 학생 상담 프로그램	방과 후 돌봄서비스 자유학기제 원어민 교사
보건·복지	건강관련 제도설계, 건강서비스 및 교육프로그램 저소득층 공적 부조 자활지원 사업	금연 교육 프로그램 무상보육사업
고용·노동	고용 및 실업 보조사업, 직업훈련 사업 취업알선 사업	실업자 훈련 프로그램 청년 인턴프로그램
문화·체육	문화프로그램 지원 사업 국민체육 활성화 사업	여행, 공연 바우처 사업 생활체육 지원 사업
재정지원	투융자 등 정책금융 세금고지 및 체납관련 사업	중소기업 창업자금 지원 세금고지서 차등구성
해외원조	해외원조프로그램	ODA 수원국내의 실험평가

자료: 오영민 외(2015)

Ⅲ 정책제언

이상에서, RCT 도입의 필요성, 개념, 우리나라 재정사업 사전평가에서의 도입 가능성에 대하여 살펴보았다. RCT를 통한 재정사업 사전검증체계 강화를 위해서는 다음과 같은 정책적 수단이 필요하다.

1. 제도적 토대 구축

재정사업 사전검증체계 강화를 위해 RCT와 같은 과학적인 평가방법을 적용하기 위해서는 원활한 평가 수행을 위한 법적·제도적 토대가 필요하다. 제도적 안정성 없이 비교적 새로운 RCT와 같은 평가기법이 정착되기 어렵기 때문이다. 특히, RCT를 시행하는 데 있어서는 관련공무원의 평가 자체에 대한 이해가 낮아 실행이 어려울 뿐만 아니라 윤리적인 문제나 정책대상자의 심리적 거부감 등 평가의 실행이 어려운 경우가 많다. 따라서 과학적인 평가방법에 대한 법적·제도적 절차를 명확히 하여 제도의 지속 가능성을 높일 필요가 있다.

가. 법적인 근거 마련

우선, 재정사업 사전 검증평가를 위해 예비타당성 조사가 「국가재정법」과 「국가재정법 시행령」에 규정되어 있는 것처럼 RCT에 대한 규정을 관련 법규에 명문화할 필요가 있다. 현재 「국가재정법」 제38조에는 재정사업 예비타당성 조사규정이 명문화되어 있으며 「국가재정법 시행령」 제13조에는 예비타당성 조사의 실시와 면제 요건이 규정되어 있다. 따라서 대규모 예산이 들어가면서 기존의 예비타당성 조사의 평가기법으로 정확한 사업효과의 예측이 어려운 노동, 복지, 교육 사업에 대하여 RCT 기반 정책평가를 의무화하는 규정을 「국가재정법」과 시행령에 신설할 필요가 있다. 미국의 경우 2002년에 'Education Science Reform Act'과 같은 법률 제정을 통해 교육프로그램의 효과성을 평가하는 데 있어 RCT와 같은 Scientifically Based Research를 실시할 것을 명문화하였다. 이 법은 모든 교육 정책 및 프로그램의 평가를 과학적 방법론에 근거하여 실시하도록 규정하고 있다.

우리의 경우도 RCT기반 정책평가가 안정적으로 실시되기 위해서는 「국가재정법」 등에 관련규정을 명문화할 필요가 있다. 이러한 규정에는 RCT 평가를 위한 예비사업, 평가대상 사업의 범위와 기본절차, 예비사업의 효과성을 과학적으로 검증할 수 있는 평가 설계의 유형에 대한 개괄적인 규정이 포함되어야 한다. <표 5>는 RCT기반 정책평가 기본이 되는 증거기반 사전타당성 평가가 「국가재정법」에 규정된 예를 보여준다.

<표 5> RCT 기반 사전 타당성 평가규정의 예

제38조(예비타당성조사) 기획재정부장관은 총사업비가 500억원 이상이고 국가의 재정지원 규모가 300억원 이상인 신규 사업으로서 다음 각 호의 어느 하나에 해당하는 대규모사업에 대한 예산을 편성하기 위하여 미리 예비타당성조사를 실시하고, 그 결과를 요약하여 국회 소관 상임위원회와 예산결산특별위원회에 제출하여야 한다. 다만, 제4호의 사업은 제28조에 따라 제출된 중기사업계획서에 의한 재정지출이 500억원 이상 수반되는 신규 사업으로 한다.

신설: 제39조 (예비사업 타당성 평가) 기획재정부장관은 제38조의 예비타당성조사 대상사업으로서 예비타당성조사에 의해 사업의 효과성 검증이 곤란하거나 불가능한 사업에 대하여 사업의 실시 전에 예비사업을 실시하고 예비사업의 효과성을 입증하는 타당성 평가를 실시하여야 한다. 다만, 사업의 시행부서에서 예비사업 타당성 평가에 준하는 평가를 실시하였을 경우 기획재정부의 승인을 거쳐 평가를 면제할 수 있다.

사업의 효과성에 대한 인과관계를 과학적으로 검증할 수 있는 평가는 다음과 같다.

1. 무작위 추출에 의한 처치-비교집단 효과성 비교 평가
2. 1항에 준하는 사업효과의 인과성을 확보할 수 있는 평가
3. 1항과 2항의 예비사업 타당성 평가의 대상사업, 수행체계, 분석방법, 평가환류에 관한 구체적인 절차는 대통령령과 예비사업 타당성 평가지침에 따른다.

자료: 법제처 국가법령정보센터 국가재정법 규정을 재구성

나. 행정자료공유 제도화

RCT와 같은 실험평가를 활성화하기 위해 필요한 제도적 방안으로 자료공유시스템을 구축하는 것도 중요한 과제이다. 개별 실험이 효과적으로 진행되기 위해서는 비용문제를 고려하지 않을 수 없는데, 다양한 기존 행정 자료를 활용할 수 있다면 비용을 크게 절감할 수 있기 때문이다. 미국의 경우 RCT를 비롯한 정책평가를 활성화하기 위하여 자료공유시스템을 구축하고 각 부처가 보유하고 있는 자료들을 서로 공유하여 평가의 비용을 최소화하고 있다. 우리나라의 경우 조세지출이나 재정사업의 심층평가에서 부처의 자료 제공 미협조로 인해 어려움을 겪고 있는 경우가 많은데, 이는 RCT를 통한 실험평가에도 유사하게 나타날 확률이 높다. 따라서 국가재정법 시행령에 각 부처의 자료 제출 공유의무를 명문화하고 행정자료를 통합관리할 수 있는 시스템을 평가전담기관에 설치해야 한다. 또한 제도 이행의 실효성을 높이기 위하여 부처의 자료공유 실태를 정기적으로 점검하고 이를 공개하거나 자료공유 노력을 정부 업무평가 또는 행정관리 역량평가에 반영할 필요가 있다.

다. 전담조직 설치

다음으로, RCT 기반 정책평가가 재정사업의 평가에 제도화되기 위해서는 해외 국가에서와 같이 전담조직 설치가 필요하다. 이러한 전담조직은 현재 재정사업 사전사후 평가를 담당하고 있는 국책연구원에 설치할 수 있다. 다만, 기존 예비타당성 조사를 수행하고 있는 KDI 경우 기존 예비타당성 조사와 상충될 수 있고 RCT 기반 평가가 재정사업의 사전평가 외에 심층평가와 같은 사후평가에서도 적용될 수 있기 때문에 다른 국책 기관이나 해외 국가와 같이 별도의 평가전문기관 설립을 통해서 실시하는 것이 바람직하다. 별도의 전문 기관은 내각 소속의 정부조직에서 독립기관으로 독립한 영국의 BIT의 형태, 미국 대통령 직속하에 공무원 조직인 Social and Behavioral Science Team, MIT의 J-PAL과 같은 순수 민간기관의 형태가 검토될 수 있을 것이다. 이러한 전문기관은 (1) RCT 정책평가 실행 (2) RCT 방법론 교육 (3) RCT 홍보 및 확산 (4) 자료관리의 네 가지 핵심 기능을 수행해야 한다.

2. RCT 평가실행을 위한 수행체계와 절차의 명확화

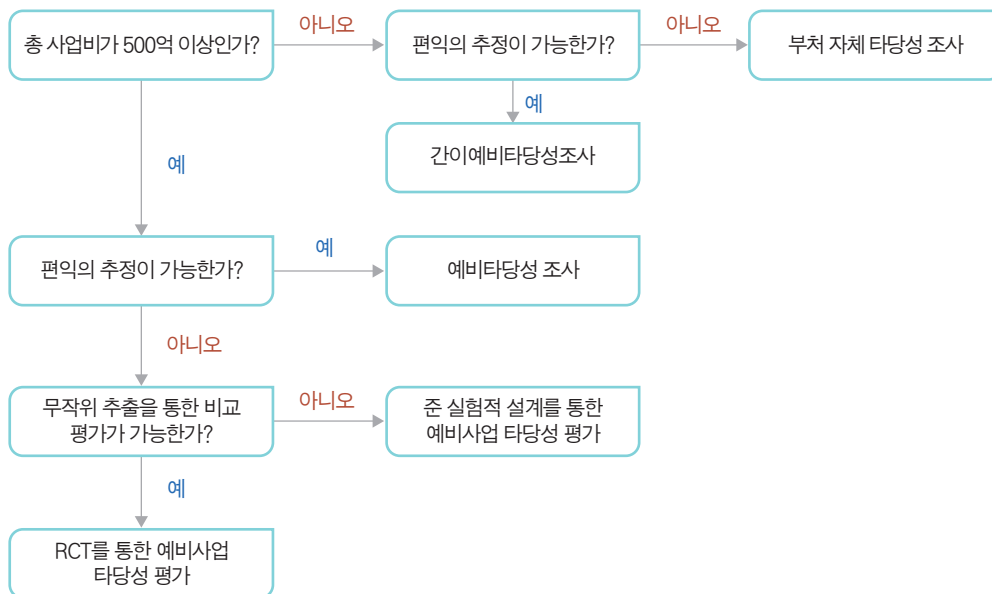
RCT 평가를 위한 법적 근거와 전담조직이 설치되었다면 RCT 평가 실행을 위한 구체적인 수행체계와 절차가 마련되어야 한다. 예비타당성 조사의 경우도 별도의 지침을 통해 평가 실행을 위한 구체적인 수행체계와 절차를 규정하고 있다. RCT를 통한 재정사업 평가의 경우도 실행 과정에 장애요인이 많기 때문에 명확하고 구체적인 절차를 작성할 필요가 있다. 이러한 절차에는 평가대상 사업 선정, 평가수행체계, 평가분석방법, 평가결과 공개와 활용에 대한 내용이 포함되어야 한다.

가. 평가대상사업 선정기준 명확화

먼저, RCT기반 재정사업 평가대상이 되는 사업의 범위를 명확하게 선정할 필요가 있다. 이는 예비타당성 조사와 같이 RCT에 의한 엄밀한 평가 설계가 불가능하거나 시행에 있어 정책적인 고려가 필요한 사업영역이 존재하기 때문이다. 이러한 특성을 고려하여 RCT를 통한 재정사업 사전검증은 예비타당성 조사와 같이 총 사업비가 500억원 이상(국가재정지원 300억원 이상)이면서 예비타당성 조사에 의해 정확한 편익의 도출이 어렵거나 불가능한 사업에 대해서 실시하는 것을 검토할 수 있다. 이러한 사업들은 주로 정확한 편익의 추정이 어려워 비용/편익분석 대신 비용/효과분석을 실시하고, AHP분석에서 정책적 효과에 가중치가 높게 부여되는 고용, 복지, 노동, 교육 등의 사회서비스들이다.

RCT기반 사전평가 대상사업 선정에서 마지막으로 고려해야 할 것은 선정기준의 객관성을 확보하는 것이다. 어떤 사업을 예비타당성 조사로 할 것인지 혹은 RCT 평가로 할 것인지, 아니면 정성적인 방법으로 사전평가를 실시할 것인지를 결정하는 것은 개별 사업의 성격과 실행 환경에 따라 매우 다르기 때문이다. 이를 위해 우선 RCT 평가대상 선정에 대한 구체적인 기준을 제시하고, 이러한 기준에 따라 평가대상 사업을 결정할 독립적인 평가사업 선정위원회를 설치할 필요가 있다. 독립적인 평가대상사업 선정위원회의 선정기준은 세부지침을 통해 구체화되어야 하지만 [그림 2]과 같이 사업규모, 편익추정 여부, 무작위 추출 여부를 통해 RCT기반 평가를 위한 최종사업을 선정할 수 있다.

[그림 2] RCT기반 예비사업 타당성 평가대상 사업 선정절차 예시



자료: 오영민 외(2015)

나. 평가수행체계의 구체화

평가대상사업 선정기준을 확정하면 평가를 수행하는 체계에 대한 절차를 마련해야 한다. 평가수행체계는 평가의 기획 및 감독기관, 평가시행기관, 연구진 선정, 연구기간을 통하여 구체화되어야 한다. RCT 정책평가는 기획재정부, 평가실행기관, 개별 부처 협업을 통해서만 수행될 수 있다. 기획재정부에서 RCT 평가대상사업과 평가에 대한 예산을 확정하면 평가실행기관은 사업부서와 사전에 면밀한 평가계획을 수립하여 예비사업을 실시하고 사전에 계획된 평가계획에 의거하여 예비사업의 효과성을 평가한다.

예를 들어, 평가실행기관은 연구진을 구성하고 연구진은 개별 부처 담당자와 공동으로 평가의 장애요인, 처치/비교 실행방법, 사업 효과성 측정방법을 포함한 구체적인 평가계획서를 확정한다. 평가계획서가 확정 되면 실제 예비사업을 처치집단에 시행하고 그 효과를 비교집단과 비교하여 사업의 최종 효과성을 판단한다. 평가 수행기간은 예산주기와 사업효과성 발생시간을 고려하여 1년으로 설정하되, 사업의 성격과 유형에 따라 6개월에서 2년까지 탄력적으로 조정할 수 있도록 한다. 다만 기획재정부에 의한 RCT 사전평가가 개별 부처의 평가에 대한 관심과 역량을 낮출 수 있기 때문에, 부처 자체적으로 시범사업 실시 후 사업의 효과성을 RCT와 그에 준하는 평가절차를 통해 입증하였을 경우 기획재정부에 의한 평가는 생략한다. 개별 부처가 실시한 평가절차의 정확성에 대한 승인은 전술한 평가대상사업 선정위원회에서 실시한다.

다. 평가방법의 확정

RCT기반 예비사업 평가의 분석방법도 엄밀한 RCT 평가를 위해 매우 중요한 요소이다. 일단, RCT 평가 기법을 적용할 대상사업과 연구진, 사업부서가 정해지면 어떤 평가방법을 사용할 것인가를 확정해야 한다. 무작위 추출을 통한 엄밀한 RCT 평가를 수행하기 위해서는 (1) 처치집단과 비교집단의 구성 (2) 집단 간 동질성의 확보 (3) 비교집단 간 다른 처치의 세 가지 전제조건이 충족되어야 한다. [그림 3]은 RCT에 의한 사전 평가에서 고려될 수 있는 평가방법의 유형과 조건을 보여준다.

[그림 3] RCT기반 재정사업평가 방법론의 선택과 요건



자료: 오영민 외(2015)

(1) 처치집단과 비교집단의 구성

RCT의 첫 번째 단계는 서로 다른 집단 간에 정책효과를 비교할 수 있는 비교집단을 구성하는 것이다. 그러나 비교집단 설정이 불가능한 경우가 많다. 주민의 반발이나 법 규정에 의해 자격 기준이 넘는 모든 정책수혜자들에게 정책이 시행되어야 하는 경우가 발생하기 때문이다. 이 경우 RCT에 의한 정책평가가 불가능하며, 사업의 효과를 검증하는 준실험적 또는 비실험적 평가방법을 사용해야 한다. 정책시행 전에 정책대상집단의 성과를 알 수 있다면 정책 시행 후에 성과를 측정하여 비교하는 단일집단 사전사후 비교법을 사용할 수 있다.

다만, 이 경우 비교집단이 설정되지 않았기 때문에 다른 요인에 의한 정책효과 발생을 통제할 수 없게 된다. 정책대상집단의 사전성과를 사전에 확보할 수 없다면 사업의 최종 효과성을 회귀분석과 같은 비실험적 방법으로 검증할 수 있는 회귀불연속 설계(Regression Discontinuity Design)나 도구변수법(Instrumental Variable Design)을 사용할 수 있다.

(2) 두 집단 간 동질성 확보

처치집단과 비교집단이 구성되었다고 할지라도 두 집단 간에 동질성이 확보되어야 한다. 처치집단과 비교 집단 간에 동질성이 확보되지 않는다면 사업의 효과가 아니라 특정 집단의 특성에 의해 정책효과가 발생할 수 있기 때문이다. 집단 간 동질성은 추첨을 통한 무작위 추출로 확보될 수 있지만 주민의 선호나 형평성 문제 때문에 수혜대상자 결정에 있어 선착순 배정이 선호되고 있다. 이 경우, 선착순 신청자를 처치집단, 신청하지 않은 사람들을 비교집단으로 하여 정책효과를 검증할 수 있지만 선착순으로 신청한 사람들은 원래부터 정책 선호도가 높은 집단이기 때문에 사업 시행의 효과로 정책효과가 발생했는지를 파악하기 어렵다. 두 집단 간의 동질성이 확보되지 않았을 경우 사용가능한 평가 설계로는 통계적 매칭법¹⁰⁾(Statistical Matching)과 이중 차분법(Difference in Difference: DID)이 있다. 통계적 매칭법은 정책 처치에 참가하지 않은 사람들을 중에서 처치집단으로 뽑힌 사람들의 특성과 유사한 사람들을 인위적으로 선발하여 비교집단을 구성하는 것이다. 매칭법은 처치집단과 유사한 비교집단을 선발할지라도 인위적으로 비교집단을 설정한다는 점에서 정확한 인과관계검증에 한계가 있다. 매칭법 외에 고려할 수 있는 평가기법이 이중차분법이다. 이중차분법에서는 정책의 참가집단과 비참가집단의 정책 처치 시행 전과 시행 후 정책효과와의 차이를 각각 측정하고 후 두 집단 간의 차이를 다시 한번 차감하여 정책의 효과를 산출한다. DID는 사업 전후의 정책 참가자와 비참가자의 성과 변화를 차감함으로써 보다 비교적 정확하게 정책효과를 측정할 수 있다. 그러나 정책 참가자의 특성이 비참가자보다 사업 전과 사업 후의 성과 변화에 더 큰 영향을 끼친다면 정책효과와의 정확한 인과관계 검증이 어려운 한계가 있다.

(3) 대상 집단 간 상이한 정책 처치의 실행

처치집단과 비교집단이 구성된 후 집단 간 동질성이 확보되었다 하더라도 처치집단에만 특정 정책을 시행할 수 없는 상황이 발생할 수도 있다. 처치집단의 구성원들이 정책 처치를 거부하거나 비교집단의 구성원들이 정책수혜를 요구할 때 발생한다. 자연과학과는 달리 정책의 수혜를 받는 사회과학의 RCT 평가에서 정책 처치를 거부하는 집단은 드문 편이지만, 비교집단이 동일한 처치를 통한 정책수혜를 요구하는 경우는 빈번하게 발생한다. 이렇듯 대상 집단 간 상이한 정책 처치의 실행이 어려울 경우 고려할 수 있는 방법이 교대실험(Rotation Design)이나 단계실험 설계(Phase-in)이다. 이러한 실험설계를 통하여 처치집단에 대한 정책 처치 후에 교대로 또는 단계적으로 비교집단에도 처치집단과 똑같은 정책 처치를 실시한다.

10) 통계적 매칭법에는 정책대상자와 통계적으로 특성이 정확히 일치하는 사람만을 선발하여 비교그룹에 매칭하는 'Exact Matching'과 성향이 유사한 사람을 선발하여 매칭하는 'Propensity Score Matching'이 있다.

(4) RCT에 의한 정책평가

전술한 세 가지 조건이 확보된다면 RCT에 의한 예비사업 평가를 실시할 수 있다. 여기서 중요한 것은 정책 효과의 충분한 통계적 파워를 확보할 수 있는 정책대상 집단의 충분한 표본 수와 사전사후 비교를 위한 대상 집단의 사전성과 데이터를 확보하는 것이다. 표본의 크기는 통계적 파워의 수준과 정책 처치로 달성되는 정책효과 크기에 의하여 결정된다. 예를 들어 학생들의 성적을 향상시키기 위해 제안된 예비사업의 참가자 수는 통계적 파워의 수준을 높일수록 향상시키고자 하는 성적의 폭이 클수록 증가한다. RCT에 의한 평가 설계로 진행되는 예비사업에서 참가자의 수를 정하는 것은 평가예산을 결정하는 데 중요한 변수가 된다. 연구진과 사업담당자는 적절한 통계적 파워¹¹⁾와 정책효과 크기를 사전에 결정한 후 참가자의 수를 정하고 그에 맞추어 예산을 편성할 수 있다.

라. 평가결과의 공개와 본 사업 시행 결정

일단 예비사업의 종료되면 처치집단과 비교집단 간의 정책효과 차이를 분석한 후 평가결과를 공개해야 한다. 이러한 공개절차는 예비타당성 조사의 절차에 따른다. 평가결과 처치를 받은 정책대상자 집단의 성과가 비교집단의 성과보다 높게 나타났다면 본 사업의 시행을 결정할 수 있다. 다만, RCT 평가결과의 외적 타당성을 확보하기 위해, 연구진은 회귀분석과 같은 비실험적 평가분석이나 참가자 인터뷰를 통한 질적 분석결과와 최종 평가결과를 비교해야 한다. 비교결과 결과가 유사하다면 기획재정부는 본사업의 시행을 최종 승인하고 예산을 배정한다.

3. RCT 활성화를 위한 대내외 환경조성

RCT와 같은 전문적인 정책평가를 원활히 운영하기 위해서는 행정부 내부의 사업담당자와 조직의 역량이 요구된다. 또한 RCT와 관련된 윤리적인 문제와 실행의 어려움으로 인하여 국회 및 국민에 대한 홍보를 통해 외부의 지지를 이끌어내야 한다. 이렇게 행정부 내부와 국민으로부터 RCT 정책평가에 대한 이해와 지지를 얻은 후에, RCT 평가에 대한 전문가들 풀(Pool)을 구축하여 지속적인 증거기반 정책평가에 대한 지식과 경험을 공유해야 한다.

11) 일반적으로 사용되는 적절한 통계적 파워의 수준은 80%이다.

가. 행정부 내부의 평가 역량 강화

전술한 바와 같이 RCT 정책평가를 운영하는 데는 여러 장애요인과 전문성이 요구된다. 이는 RCT 평가가 내생성을 줄여 정확한 정책효과를 측정할 수 있는 장점에도 불구하고 실무에서 크게 활성화되지 않은 원인이 되었다. 더구나 연구진에 의해 수행되는 예비타당성 조사와는 달리, RCT에서는 실제 정책 처치가 실행되기 때문에 사업담당자의 RCT에 대한 깊은 이해를 바탕으로 한 적극적인 참여가 성공적인 정책평가를 위해 필수적이다. 이를 위해 사업담당공무원의 RCT 정책평가에 대한 이해도를 높이기 위한 역량 강화가 요구된다.

역량 강화를 위해서는 MIT의 J-PAL의 훈련과정을 참고할 필요가 있다. J-PAL에서는 5일 전일제 강의가 사례토론, 실제 적용실습을 통하여 RCT에 대한 강도 높은 교육프로그램을 제공하고 있다. 우리나라의 경우 RCT가 재정사업의 사전평가로 제도화되면 전담기관에서 실제 RCT평가와 더불어 담당자에 대한 교육훈련을 실시해야 한다. 또한 RCT가 가능하지 않은 다양한 사업들이 존재하기 때문에 RCT 평가 교육 외에 다양한 정책 평가 기법에 대한 교육과정 프로그램을 개발하여 사업담당자의 평가 관련 역량을 높여야 한다.

특히, 평가의 역량 강화를 위해서는 교육훈련의 강화와 함께 희망리본의 사례에서 언급되었듯이 현재 순환보직으로 인해 1~2년에 머무르고 있는 사업담당자의 임기를 적어도 예비사업이나 시범사업의 평가가 종료되는 기간까지 연장할 필요가 있다. 정책평가에 대한 이해도를 높이고 정책효과의 산출을 위해서는 일정시간이 필요하기 때문이다. 따라서 대규모 예산이 들어가는 신규 사업의 경우, RCT에 대한 이해와 사업타당성의 입증을 위해 사업책임자의 임기를 보장해주는 유연한 인사제도가 필요하다.

나. 국회 및 대국민 홍보

RCT 정책평가 보급에 가장 큰 장애요인은 정책대상자에게 차별적으로 정책 처치를 하는 윤리적인 문제이다. 이를 위해 RCT가 사람을 대상으로 하는 실험이 아니라 정책의 효과를 과학적으로 검증하는 평가기법이라는 사실을 국회 및 국민들에게 적극적으로 홍보할 필요가 있다. 영국의 BIT는 RCT를 통해 도출된 평가결과를 적극적으로 언론을 통해 홍보하고 있다. 우리나라의 경우, 청년고용, 무상보육과 급식, BK 21사업 등 사회적으로 논란이 되었던 대규모 정책사업이 있었지만, 정책효과에 대한 과학적 증거를 통해 진지하게 토론된 적은 거의 없었다. 따라서, 실증기반 정책결정과 평가를 확산시키기 위해 대규모 예산이 들어가지만 정책효과에 논란이 있는 사업의 경우, 예비사업단계에서부터 지속적으로 국회와 국민들에게 홍보 및 설득작업을 할 필요가 있다.

이런 과정을 거쳐 RCT 정책평가에 대한 국민들의 이해도가 높아지면 장기적으로 ‘PAYGO’와 더불어 대규모 재정을 수반하면서 정책효과가 의심되는 의원입법에 대하여 법안제출 전에 반드시 RCT 평가결과를 제출하는 것을 의무화하는 것을 검토할 수 있을 것이다. 더불어, 정책 실험평가에 대한 대국민 수용도를 높이기 위해 실제 RCT 평가과정에서 발생할 수 있는 법적, 윤리적인 문제를 사전에 심사할 수 있는 IRB¹²⁾ (Institutional Review Board)를 정부 내에 설치하는 것을 검토할 필요가 있다.

다. RCT 평가관련 전문가 네트워크 구축을 통한 지식공유

마지막으로, RCT 평가의 보급과 확산을 위해서는 관련 전문가 네트워크를 구축하여 관련 지식과 경험을 공유하는 것이 필요하다. 사실 정책평가에 있어 RCT와 같은 실험설계방법의 기초지식은 석사과정의 연구 설계 수업에서 다루고 있는 내용이다. 실무에서도 재정사업 심층평가 지침을 통하여 RCT의 방법론이 이론적으로 설명되고 있다. 그러나 이론적인 지식 외에 RCT 평가과정에서 발생할 수 있는 장애요인을 해결하고 실행 방법을 구체적으로 적용할 수 있는 평가전문가는 국내에 매우 부족한 실정이다. 이러한 전문가의 부족은 RCT 평가가 미국이나 영국에 비해 국내에서 활성화되지 않고 있는 이유 중의 하나이다. 따라서 RCT를 통한 재정사업 평가의 보급을 위해 평가관련 전문가들의 네트워크 구축과 지식공유가 필요하다. 이러한 전문가 그룹은 연구논문 작성을 위한 사회실험 자료 분석뿐만 아니라 실제 정책과 사업을 분석하여 RCT에 기반한 연구 및 평가 계획을 수립하고 이를 현실 정책평가에 활용할 수 있는 방법과 절차를 제시할 수 있다. 또한 평가 관련 연구 및 실제 평가과정에서 얻어진 지식과 경험을 전문가 그룹에서 적극적으로 공유하여 RCT기반 정책 평가 확산에 기여할 수 있다.

IV 결론: 기대효과

새로운 재정사업이 충분한 사전점검 없이 정치적인 요구에 의해 급격하게 도입되는 사례가 증가하면서, 시행착오로 인한 사회적 비용이 증가하고 다. 이러한 부작용을 막기 위해 500억원 이상 대형사업의 경우 타당성을 조사하는 예비타당성 조사가 운영되고 있으나 정책적 효과를 정확히 계량화할 수 없는 교육이나 복지사업에 대한 정확한 사전평가는 매우 어렵다. 예비타당성 조사가 나름대로 재정사업의 무분별한 시행을

12) 연구에 참여하는 연구대상자의 윤리적인, 과학적 문제발생을 연구 시작 전에 심사하여 연구계획을 승인하는 기구로서, 해외 연구기관에 설치되어 있다. 우리나라의 경우 임상실험과정에서 윤리적인 문제가 발생할 수 있는 생명공학 및 병원연구소에 설치되어 있다.

막는 완충작용을 해왔을지라도 사업의 정확한 편익을 추정할 수 없는 사업에는 적용하기 어려운 한계가 존재하기 때문이다.

이런 의미에서 우리나라에서 최근 급속하게 확대되고 있는 복지나 고용관련 사업을 사회실험 방식의 RCT 평가를 도입하여 효과가 있는 사업만을 시행한다면 정확한 효과 검증 없이 무분별한 재정사업의 시행으로 인한 재정의 누수를 크게 줄일 것으로 판단된다. 특히 외국에서도 최근 교육이나 복지사업에 대해서는 RCT 평가방식이 적용되고 있기 때문에 최근 재정의 어려움을 겪고 있는 우리나라에서 RCT의 제도화는 매우 중요한 과제이다. 이 글이 아직 RCT가 크게 활용되고 있지 않은 우리나라의 적용가능성에 대한 이해도를 높여 실험 방식의 RCT 평가가 우리나라 재정사업평가에서 활발하게 사용되기를 기대한다. 또한 본고의 내용이 실제 정책평가를 담당하는 실무자에게 많은 도움이 되기를 희망하며 글을 맺는다.

| 작성자 |

오영민 / 한국조세재정연구원 부연구위원
(044-414-2231)

박노옥 / 한국조세재정연구원 선임연구위원
(044-414-2267)

강희우 / 한국조세재정연구원 부연구위원
(044-414-2224)

| 참고문헌 |

- 법제처, 국가법령정보센터-<http://www.law.go.kr/main.html>.
- 이삼열·정의룡·이은하, 「시범사업에 관한 탐색적 연구: 보건복지가족부 사업을 중심으로」, Working Paper, 2009.
- 오영민·박노옥·강희우, 『재정사업사전검증체계 강화를 위한 연구: RCT 도입방안을 중심으로』, 연구보고서 15-17, 한국조세재정연구원, 2015.
- 한국개발원, 「2013 KDI 예비타당성 조사지침」, 2013.
- 홍승현·원종학, 「적극적 노동정책의 재정효율성 평가방법에 관한 연구」, 한국조세재정연구원, 2013.
- OMB “What Constitutes Strong Evidence of a Program’s Effectiveness,” 2004.

