

스위스의 평가비용

슈테판 리더(Stefan Rieder)*

번역 : 정 미 경

-
- I. 서
 - II. 평가비용: 개념 정의
 - 1. 직접비용: 경영학적 관점
 - 2. 간접비용: 경제학적 관점
 - III. 평가비용: 실증적 결과
 - 1. 평가비용은 얼마인가?
 - 2. 평가비용은 어디까지 허용될 수 있나?
 - IV. 토 론
 - 1. 평가비용은 얼마인가?
 - 2. 평가비용은 어디까지 허용될 수 있나?
-

슈테판 리더(Stefan Rieder)가 평가비용은 평가를 의뢰한 측이나 평가를 위임받아 평가를 수행하는 측이나 다 같이 관심을 갖는 사안이다. 스위스에서 평가비용은 실제 어느 정도인가? 그리고 적절한 평가예산 책정을 위해 어떤 권장기준을 찾을 수 있는가? 연방의 평가업무에 대한 기존의 문헌과 데이터를 바탕으로 위의 두 가지 질문에 대한 답을 추적했다.

* 박사, 인터페이스 정치조사연구상담 부문대표, 루체른대학 강사

** 경제학 박사, 프랑크푸르트대학 강사

I. 서

무엇을 [전문적으로]¹⁾ 평가한다는 것은 많은 경우 꽤 복잡한 사안으로 간주된다. 따라서 평가비용과 관련하여 평가의뢰자 또는 평가수임자 할 것 없이 드물지 않게 다음 두 가지 질문을 던진다. “평가비용은 얼마인가?” 또 “평가비용은 얼마까지 허용될 수 있는가?” 첫 번째 질문은 지출규모 파악을 목적으로 하며 실증적인 성격을 갖는다. 그것은 평가를 위해 지출한 경험적 가격에 관한 것이다. 반면 두 번째 질문은 평가계획에 관련되어 있으며 평가의 «합리적인 비용수준»을 묻는 것으로 규범적인 내용을 가지고 있다. 이들 질문은 그에 대한 답을 찾는 연구의 주제와 범위에 따라 변동이 심할 것이기 때문에 쉽게 답하기가 어렵다. 그럼에도 불구하고 각 조사 마다 평가비용이 차지하는 역할 때문에 이에 대해 심도 있게 고찰할 필요성이 있다.

위의 두 가지 질문에 대해 논의를 구체화하기 위해 이 두 가지 질문은 더욱 세분화되어야 한다. 평가를 위한 지출규모에 대한 문항은 다음과 같이 나눌 수 있다. 평가비용에 포함되는 것이 무엇인가? 무엇까지 비용으로 포함시킬 수 있는가? 여러 가지 토론의 주제가 되고 있는 간접적 평가비용은 무엇인가? 연방정부는 평가를 위해 얼마를 실제로 지출하는가? 평가의 종류에 따라 비용편차가 있는가?

평가비용의 허용치에 대해서는 종종 퍼센티지로 대답되었다. 이는 대체로 평가를 위한 지출규모가 정책프로그램 또는 정책조치의 총비용을 기준으로 얼마나 되어야 할지 보여 준다. 우리는 이 질문을 다음과 같이 세분화하였다: 평가 지출규모와 관련된 기준치들이 실제 있는가? 이 기준치들은 서로 얼마나 큰 차이가 나는가? 이 기준치들을 연방행정부[Bundesverwaltung] 실제 평가지출규모와 비교하면 어떤 차이가 있는가?

1) []안의 글은 번역자가 내용의 이해를 돕기 위해 삽입한 것이다. 원문에서 평가는 수식어 없이 Evaluation으로만 쓰여 졌으나 이글의 경우 일반적 평가가 아닌 평가전문가에 의한 공기관의 프로그램 평가를 뜻함으로 글의 이해를 돕기 위해 삽입하였다.

평가에 대한 핸드북이나 평가에 대한 기존의 연구에서 평가비용의 문제는 간략하게만 다루어져 왔다. 이 글은 스위스와 외국의 저서에서 평가비용에 대한 연구내용을 모아 소개하고 그에 대한 해설을 덧붙이는 것을 목표로 한다. 이를 위해 우선 제 2장에서는 먼저 평가비용에 속한 몇몇 개념을 정의한다. 3장 1절에서는 활용 가능한 스위스의 데이터로 평가비용의 규모를 간략한 도식을 만들어 소개한다. 그 다음으로 평가비용규모의 권고를 개괄하고 이와 연방행정부에서 실제로 이루어진 지출규모를 비교한다 (3장 2절). 4장에서는 연구결과를 요약하고 토론한다.

II. 평가비용: 개념 정의

평가에 드는 비용은 얼마인가? 이 문제에 답을 하기위해서 먼저 비용이 어떻게 구성되는지 살펴본다. [평가비용은 먼저] 직접비용과 간접비용이 구분된다. 직접비용은 경영학의 비용이론에 따라 상당히 명료하게 정의되고 산정된다. 쉽지 않는 문제는 간접비용을 정의하고 파악하는 것이다. 우리는 이 두 가지 [비용 개념]을 이어지는 절에서 좀 더 자세하게 고찰하기로 한다.

1. 직접비용: 경영학적 관점

우리는 비용을 정의하기 위해 평가를 스스로 실행했거나 평가를 부분적이거나 전체적으로 외부[전문인력]에 의뢰하여 진행한 바 있는 공공행정기관의 관점에서 비용고찰에 임한다. 직접비용은 평가의뢰자, 평가자 (내부 및 외부의 평가자) 그리고 피 평가자에 의해서 발생한다.

직접비용은 경영학의 비용이론에 의거 개별비용과 공통비용으로 나눌 수 있다. 개별비용은 명확히 구분이 가능한 일체의 인건비와 경비이다. 대개 평가를 기획, 집행하고 평가결과를 알리는 과정에서 발생하는 인건비가 대부분을 차지한다. 공통비용은 기본적으로 평가업무의 구분방식에 따라 부과된다. 전형적으로

로 부서지휘비용 또는 그 밖에 평가수임부서에 (직접) 해당되는 않는 행정 (인사관리, 전산서비스) 비용이 이에 속한다.

다음의 표는 직접비용의 가장 중요한 항목들을 소개한다.

〈표 1〉 비용항목에 따른 직접적 평가비용

비용항목	내역
개별비용	
인적개별비용	<ul style="list-style-type: none"> - 평가의뢰자의 기안서 작성을 위한 시간소모 - 평가수임자의 평가 제안서작성을 위한 시간소모 - 평가의뢰자들의 제안서선정을 위한 시간소모 - 평가수임자의 평가집행 (표본조사, 보고서 작성)을 위한 시간소모 - 평가의뢰자의 평가작업지원과 자체표본추출, 자료조사 등에서 발생하는 시간소모 - 평가의뢰인과 평가수임자의 평가결과에 대한 프레젠테이션과 토론을 위한 시간소모 - 준비 및 데이터조사 시간소모 (예를 들면 인터뷰시간, 통계분석)
물적개별비용	<ul style="list-style-type: none"> - 평가집행자의 평가집행을 위한 여행경비 - 평가의뢰자, 평가를 수임받은 평가집행자, 피 평가자에게서 발생하는 임대료, 장비 및 소프트웨어 구입비 - 외주업무의 비용 (예를 들면 번역비)
공통비용	
인적공통비용	<ul style="list-style-type: none"> - 평가의뢰자의 의뢰업무처리를 위한 행정비용 (계약체결, 관리) - 평가의뢰자, 평가집행자, 피 평가자의 인적관리 행정비용
물적공통비용	<ul style="list-style-type: none"> - 평가의뢰자, 평가집행자, 피 평가자에게서 발생하는 임대료, 장비 및 소프트웨어 구입비

실제로 외주평가를 할 경우에 비용의 큰 몫을 차지하는 부분이 외주업체의 보수로 지급되기 때문에 개별비용과 공통비용의 구분이 간단해진다. [이 경우] 비용을 정확하게 구분하는 것은 공공행정주체의 입장에서 그다지 중요하지 않다. 이와 반대로 행정조직 내부에서 자체평가가 이루어지는 경우 비용을 개별비용과 공통비용으로 구분하기위해 [비용의 성격이] 파악되어야 한다. 이것은 수익과

지출의 결산이 [결산회기를 근거로] 시간단위로 나뉘어 시행되고 [수지결산에는] 공통비용이라는 비용항목이 존재하지 않기 때문에 종종 어려움에 봉착한다. 이 경우 비용은 개별비용과 공통비용에 대한 계산비율을 정하여 산출해야 한다. [또 다른 한 가지 직접비용 산출과 관련된 것은] 평가사업을 수입하고자하는 평가사업의 수행업체들이 사업제안서를 작성하는데 투입된 비용들이 평가비용을 산출하는데 포함된 경우는 전혀 없었다.

2. 간접비용: 경제학적 관점

직접비용보다 훨씬 파악하기 어려운 것이 간접비용이다. 여기서 우리는 매우 상이한 두 개의 간접비용 개념정의를 소개한다. 첫 번째는 기회비용으로 간접비용을 정의하는 것이다 (Alkin/Ruskus, 1984, 3). 기회비용은 종종 대안비용, 포기비용 또는 잠재가격으로 표기되며 비용-효용분석에 활용된다 (Mühlenkamp, 2009). 기회비용은 평가에 사용할 예산을 다른 사업에 투입하였을 경우 발생하는 효용에 상응한다. 예를 들면 [평가 사업이 진행되지 않았더라면] 평가사업에 들어갈 비용은 일반계획재정에 추가적으로 지출할 수 있었을 것이고 일반계획 사업은 이를 통해 해당사업의 효과와 효용을 높일 수 있을 것이다.

간접비용에 대한 두 번째 개념에는 평가에 수반되는 잠재적인 부정적 효과에 초점이 맞춰져 있다. [즉 평가가 잘못된 경우 잠재적 손실도 간접비용으로 고려해야 한다.] 무엇보다 사업의 시간소모 또는 사업의 방법에 대해 평가가 그와 유사한 종류의 정책조치나 사업계획을 집행하는 책임자를 잘못된 방향으로 이끌 때 발생할 수 있다. 이것은 오로지 오랜 시간이 지난 다음에야 그 예측하기 어려운 결과를 불러일으킬 전략을 시행하지 않게 될 위험을 지닌다 (Alkin/Ruskus 1984, 3; Frey 2007, 209-210, 216). 플레이에 따르면 평가를 통해 연구영역에서 내적 동기가 감소되거나 성과를 평가하는 기준이 조작될 위험도 존재한다고 밝힌다 (Frey 2007, 210-212). 이 세 가지 효과들이 평가의 간접비용이라고 할 수 있는 비용을 발생시킨다.

Ⅲ. 평가비용: 실증적 결과

우리는 이 장에서 위의 개념고찰을 마치고 실증적 결과로 방향을 돌린다. 이를 위해서 우리는 먼저 평가비용이 얼마나 되는지에 대한 의문을 제기한다. 그 답을 찾기 위해 우리는 미국의 연구와 연방행정부의 평가관련 데이터를 활용했다.

1. 평가비용은 얼마인가?

평가비용에 대한 연구는 많지 않다. 소수의 평가비용에 대한 연구 중의 활용할 만한 연구가 1983년 알킨과 스테처(Alkin/Stecher)에 의해서 이루어졌고 미국에서 출판되었다. 이 연구는 25명 미국의 평가자들에게 평가비용 산정을 요청했다. 이 방식으로 파악된 평가는 지역 및 지방차원에서 교육, 건강, 사회보장 분야에서 실행된 것이었다. 이 연구에는 집행과 절차평가의 형식적 특성 뿐 아니라 효과연구에서 총량적 특성도 포함되어 있다 (Alkin/Stecher, 1983, 122). 여기에서 비용은 평가집행과 관련하여 자체 전문인력비용과 외주 전문인력비용, 비서업무비용, 데이터처리비용, 물적 비용 (여행경비, 장비와 시설의 임대료 등)으로 구분되어 있다. 평가의 직접비용은 고려되지 않았다. 그 밖에도 우리는 기술하는 방법의 어려움을 이유로 이 비용이 공통비용 또한 고려되지 않는다는 것을 전제로 출발한다. 그 결과 가장 큰 비용은 인건비에 할당되어졌다 (비용은 전문인력에 70%, 보조 인력에 16% 할당되었다). 알킨과 스테처가 조사한 평가는 매우 소수만 외부에 발주되었다. 즉 (인건비에 포함되지 않는) 상담비용은 2% 정도만 할당되었고 나머지 비용은 12%로 비교적 낮은 편이었는데 그 중 가장 큰 몫은 약 그 1/3에 해당되는 자재비 및 전화비였다 (Alkin/Stecher, 1983, 125).

알킨과 스테처의 조사결과에서 한 가지 흥미로운 것은 평가비용의 규모에 의거한 분류방식이다. 이에 따르면 평가비용이 4,000 달러인 경우 소규모평가로 분류되어있다. 종종 당초의 구상이 실제 사업집행 과정에 적절하게 적용되었는

지 여부를 검토하는데 초점이 맞추어진 새로운 혁신계획과 관련된 사업평가가 여기에 해당된다. 중간규모의 평가는 비용 10,000 달러 규모의 평가이다. 이 규모 평가는 형식적 특성(기존에 진행되는 프로그램의 재검토)과 총량적 특성(재정조달개편에 대한 결정과 실제 재정지출의 재검토)을 함께 보여준다. 25,000 달러 규모의 평가는 대규모의 포괄적인 평가로 분류된다. 이 규모는 대체로 기존의 프로그램을 평가하며 총량적 형식적 두 가지 특징을 다 포함하고 있다 (Alkin/Stecher, 1983, 25). (1980년~2010년 미국 소비자물가인상률) 약 165% 물가상승을 고려하면 현재 10,600 달러 규모는 소규모, 26,500 달러는 중간규모 그리고 66,200 달러는 대규모 평가사업에 속한다. (주석 1)

스위스에서 위의 연구와 비교할 만한 데이터를 찾으려면 먼저 평가재정수요 지침을 아래와 같이 제공한 부스만 (Bussmann, 1995, 56-57)의 연구를 발견하게 된다. 부스만은 소규모조사연구(예를 들면 집행평가)를 50,000 프랑켄으로 산정한다. 부스만에 따르면 100,000 프랑켄까지는 명백히 한정된 부분에 초점을 맞춘 간단한 조처에 대한 조사연구로 추계된다. 프로그램의 효과와 효율에 대한 상세한 평가는 대략 50,000에서 300,000 프랑켄까지 비용이 지출된다. 이 산정 방식에 개별비용만 아니라 공통비용도 포함시켰는지는 그의 지침서를 통해 파악할 수가 없다.

평가 비용에 대한 실증데이터는 2001년 출판된 비드메르와 그 동료들의 연구에서도 찾아 볼 수 있다 (Widmer et al., 2001). 저자는 8개 연방관청에서 실제로 행해진 평가를 연구했고 다음 표에서 볼 수 있는 개별평가비용에 대한 데이터를 수집했다. 우리는 여기 주로 평가의뢰자와 평가집행자의 개별비용이 수집되었다고 본다.

이 비용 중 피 평가자에게 귀속될 부분을 어느 정도인지 모든 평가관련자에게 해당되는 공통비용이 얼마인지는 자료를 통해서 알 수는 없었다.

〈표 2〉 여러 스위스 연방관청들의 평가비용

업무부처	비 용	금 액 (단위: 프랑켄)
개발협력국 (DEZA)	프로젝트평가 평균비용	< 50,000
	주정부프로그램 평가평균비용	100,000
	주정부에 전체에 해당되는 문제에 대한 횡단면평가에서 평균비용	150,000
연방사회보장청(BSV)	평가평균비용	40,000 bis 120,000
연방사법청(BJ)	피해자구조법분야 3개 평가의 평균비용	60,000 bis 70,000
연방에너지청(BFE)	에너지2000 조치의 평가평균비용	109. 000

최신데이터는 발타자르(Balthasar, 2007, 392-393)의 연구에서 찾을 수 있다. 그는 1999년에서 2002년 사이 연방행정부에서 실시된 278개 평가의 평가비용을 활용했다. 그의 평가비용에는 평가의뢰자와 평가집행자에게서 발생한 개별비용이 포함된 것으로 보이지만 피 평가자의 개별비용과 전체 평가관련자에게서 발생한 공통비용은 평가비용에 포함되지 않은 것으로 보인다. 결과는 다음과 같다.

〈표 3〉 1999년~2002년 연방행정부에서 실시한 278개 평가의 비용

	비용 (단위: 프랑켄)	총 278개 조사된 평가에서 해당 평가비율
소규모 평가	<50,000	30%
중간규모 평가	55,000~99,000	25%
규모가 큰 평가	100,000~250,000	20%
규모가 매우 큰 평가	>250,000	15%
데이터 부재	-	10%
278개 평가의 평균비용	115,000	
1999년~2002년 연방행정부가 평가에 지출한 전체 지출액	8,000,000	

출처 : Balthasar, 2007, S. 392.

발타자르가 수집한 평가비용의 분포는 업무부처마다 매우 큰 차이를 보인다. 29개 조사 행정부처 중 5개 부처에서는 평가의 절반이 50,000 프랑켄 이하였다. 3개의 연방부처에서는 정반대로 평가비용의 절반이 250,000 프랑켄 이상이었다. 30년이 지난 알킨과 스테처의 연구데이터를 스위스 데이터와 (가능한 선까지) 비교해보자. 스위스의 경우 특히 소규모평가의 비용이 확연히 높다는 것을 확인할 수 있다. 스위스에서 소규모 평가의 최하비용은 약 40,000 프랑켄이다.

한발 더 나아가서 부스만의 지침서에 제시된 1990년대 중반 평가비용의 변동폭은 2000년 전후에 데이터를 연구한 비드메르(Widmer et al. 2001)와 발타자르(2007)의 실증결과와 일치함을 확인할 수 있다. 물론 이것이 변동폭의 상한선과 하한선을 벗어난 가외치가 없다는 것을 의미하지는 않는다. 전문가 단기평가를 보면 부분적으로 평가비용이 10,000 프랑켄 이하인 것도 있었고 종단비교를 수행하는 상세한 평가는 비용이 500,000 프랑켄 이상인 것도 있었다.

그럼 어떤 것이 평가비용을 결정하는 요소인가? 이것은 평가의 종류, 평가를 위해 선택된 방법, 평가대상의 복잡성, 평가대상의 시간적 범위까지 다양하고 풍부하다. 이와 관련된 스위스의 실증자료는 매우 드물다. 비드메르(1996, 4-5)는 8개의 평가(수행기간 12개월 미만)와 7개의 장기평가에 대해 평가비용과 평가기간을 비교했다. 이것은 조사연구의 질적인 수준이 비슷한 경우 평가프로젝트 수행기간이 비용에 결정적 영향을 미친다는 것을 보여주었다.

발타자르(2007)의 데이터는 평가비용과 평가종류의 상관관계를 총량적으로 검토할 수 있도록 해준다. 그의 278개의 평가에서 19개는 분명하게 프로그램의 집행평가로 분류할 수 있다. 이것들은 유일하게 어떤 대상에 대한 조치가 어떻게 적용되었는가에 대한 평가이다. 47개 평가는 전적으로 효과에 대한 평가였는데 이 효과평가는 오직 해당그룹이나 목표그룹에 나타난 효과를 연구하는 것이었다. 그 밖의 평가는 상이한 중점들이 혼합된 것들이다. 알킨과 스테처(1993, 125)의 연구와 부스만(Bussmann, 1995) 지침서를 근거로 우리는 집행평가는 예산이 낮은 편이고, 효과평가에서는 높은 예산을 기대할 수 있다는 가설을 설정할 수 있었다. 그러나 발타자르의 연구결과와 이와 다르다. 그의 연구에서는 집행평가에서 비용 50,000 프랑켄 이하의 연구가 많지 않음을 발견할 수 있었다.

〈표 4〉 비용범주에 따른 1999년~2002년 연방행정부의 평가수

평가비용 (단위: 프랑켄)	순수한 집행평가 수	순수한 효과평가 수
< 50,000	3	13
55,000~99,000	8	12
100,000~250,000	7	13
> 250,000	1	8
Total	19	47

출처: 발타자르(2007)의 연구데이터분석

2. 평가비용은 어디까지 허용될 수 있나?

우리가 관찰한 평가비용에 대한 자료들은 공공기관들이 각 조치와 프로그램의 평가에 얼마나 지출할 준비가 되어있는지 대체로 정확하게 보고를 하고 있다. 그러나 이 자료의 수치들이 평가비용을 지출해야 하는 것인지 아닌지 또 얼마나 지출되어야 하는지를 말해주지는 않는다. 이와 관련하여 프로그램의 책임자들은 종종 평가비용을 책정할 근거로 삼을 기준을 세울 수 없는지 문의한다.

이에 대해 단순한 (이론적) 지출기준치는 다음과 같이 정의될 수 있다. 즉 평가효용이 최소한 평가비용에 필적해야 한다는 것이다. 이 명제는 예를 들면 이에 맞는 평가표준을 정하도록 과제를 부여한다 (Widmer et al., 2000, 7; Widmer, 2005, 5; DeGEval, 2008, 27). 개발협력국은 이 명제에 덧붙여서 평가비용은 평가를 통해 획득되는 학습효과까지 포함시켜 기준치를 정해야한다고 주장한다 (DEZA, 2008, 5). 이 경우 평가계획을 세우면서 비용과 효용의 상관관계를 알고자 하면 그것은 기껏해야 -가능할 경우에- 실증적으로, 아니면 사후평가를 한 다음에야 비용이 산출될 근거를 마련할 수 있다는 현실적 문제에 봉착하게 된다. 기회비용을 산정하는 것이나 부정적인 부수효과를 파악하는 것도 어려움이 크다. 따라서 이론적 기준치 또는 이를 근거로 평가비용의 표준을 정하는 것은 실제로 전혀 검토되지 않았다.

그렇다면 대안적 기준치가 있는가? 각각 분명히 규정된 평가프로젝트에 대해서는 평가문항의 작성 그리고 적절한 평가방법론을 찾으면서 예산이 책정된다.

평가대상과 방법이 구체성을 더할수록 비용산정에 정확성이 더해진다. 이러한 비용 산출절차는 평가대상에 대해 상대적으로 많은 정보가 있는 경우 그리고 특정 내용에 대한 효용테스트가 잘 이루어진 방법론이 투입되는 경우 실용적이다. 만약 이런 두 가지 전제가 충족되면 세 번째 평가프로그램의 예산책정비율의 정의가 가능하다. 그러면 어떤 방법으로 비율을 정할 수 있는가? 이 글에서는 이 문제에 대한 1차적 답으로 평가비용에 대한 권고와 실증자료를 찾았다. 그것들은 다음의 표에 개괄되었다.

〈표 5〉 프로그램 전체예산 중 평가지출 비율에 대한 권고와 실증데이터

출처	명시	평가비용의 비율
평가예산규모 권고		
Alkin/Stecher (1983, 120)	프로젝트비용에서 자체평가지출의 기준치 (단위: 퍼센티지)	1-10%
연방보건청(BAG)-지침 (BAG 1997, 24)	프로젝트비용에서 자체평가지출의 기준치 (단위: 퍼센티지)	5%
연방보건청(BAG)-지침 (BAG 1997, 24)	프로젝트비용에서 외주평가지출의 기준치 (단위: 퍼센티지)	10%~15%
Widmer et al. (2001, 37)	프로젝트비용에서 술, 담배, 불법마약 예방예산에 대한 평가지출 기준치 (단위: 퍼센티지)	10%
유럽위원회 재정총국 (European Commission DG Budget) (2004, 32)	지출계획 전체비용에서 평가예산의 비율	5%
유럽위원회 지역정책 총국 (European Commission DG Regional Policy) (2008, 47)	혁신프로그램 또는 시범프로그램의 전체비용에서 평가예산 비율	10%까지
평가예산 규모에 대한 실증데이터		
Alkin/Stecher (1983, 120)	교육관련 여러 공공기관과 프로그램의 평가예산 비율	0,2~5.5%
Widmer et al. (2001, 65)	에너지2000 프로그램의 전체지출 중 평가비용 (단위: 퍼센티지)	1.1%
Balthasar (2007, 392)	1999년~2002년 연방 전체지출에서 연방의 평가비용 (단위: 퍼센티지)	0.02%
Balthasar (2007, 393)	1999년~2002년 각 부처 전체 지출에서 부처의 평가비용 (단위: 퍼센티지)	최고 0.05% 최저 0.005%
Balthasar (2007, 392)	유럽연합 (EU) 전체지출에서 EU의 평가비용 (단위: 퍼센티지)	0.09~0.14%

〈표 5〉의 실증적 수치와 권고된 평가지출규모를 어디까지 서로 비교할 수 있는가? 〈표 5〉에서 권고는 대부분 프로그램이나 조처의 [전체 예산을] 기준으로 하지만 실증적인 수치는 대체로 [해당 조직의] 전체지출규모를 기준으로 하기 때문에 [직접적인 비교에] 어려움이 있다. 그럼에도 여기에서 조심스럽게 몇몇 사실을 확인할 수 있다. 〈표 5〉에 제시된 수치들 중 두 개가 정확하게 프로그램 전체비용을 기준으로 하고 있다 (1982년 알킨과 스테처의 권고와 그들의 실증데이터의 수치; 에너지2000 프로그램과 2001 비드메르 실증데이터의 수치). 이 두 결과는 기준치에서 저비용평가에 속한다. 발타자르 (2007)의 조사에서 부처 전체지출에서 평가비용지출의 비율을 관찰하면 대체로 실제 평가지출은 권고기준치에 훨씬 못 미친다 (부처예산기준 최고치 0.05%). 물론 부처의 전체지출에 분명히 여러 가지 이론적 또 실질적 이유로 인해 평가로 항목정리가 되지 않은 부분 또 개별적으로 의미가 크지 않은 평가지출 부분이 상당 포함되어있다 (예를 들면 [평가업무가 포함된]입법절차 지원업무). 그러나 우리가 프로그램과 조처의 평가를 위한 할당을 총 부처지출의 (최저) 10%로 정할지라도 평가예산은 최고 부서예산의 0.5% 정도이다. [표에 제시된] 지출은 이에 훨씬 못 미친다. 달리 표현하면 부처 총지출에서 평가지출의 몫은 실제 각 저서의 권고와 달리 저조하고 그 격차도 크다. 그 이유는 다음 두 가지로 생각해 볼 수 있다. 그 이유는 평가지출규모에 대한 권고가 너무 높거나 또는 그 반대로 평가지출이 전체예산에 너무 낮게 책정되어 있기 때문이다. 이중 어느 것이 실제 이유인지는 여기서 제시된 데이터만으로는 판단할 수가 없다.

IV. 토 론

글을 마감하며 이글의 주요한 두 가지 질문에 대해 다시 한 번 토론해보자.

1. 평가비용은 얼마인가?

앞의 여러 가지 비용 정의를 통해 평가비용을 산정하는데 대체로 개별적으로 산정하기가 쉬운 개별비용의 내역은 잘 알 수 있음이 파악되었다. 이것은 특히 비용의 상한선이 정해진 외주평가의 경우 이에 해당되었다. 그 밖에 비용을 정확하게 명시하는데 어려움을 주는 세 가지 제한성을 다음과 같이 정리할 수 있다.

- 첫째, 대부분의 행정업무에 내부비용을 도출할 수 있는 방식으로 정리된 수입지출결산이 결여되어있다. 이에 달리 방법이 없을 경우 시간단위비용(총비용)을 활용하여 내부행정업무의 개별비용과 공통비용을 계산하는 것이 대안이다. 그 전제는 대부분의 업무부처에서 실행되지 않지만 (대략적이라도) 업무의 시간소모내역이 최소 하나라도 존재하는 것이다. 이를 근거로 기 실행된 평가에서 부분적으로만 또는 전혀 알려지지 않은 내부행정비용과 외주평가에 대한 부수비용을 [간접적으로] 산출할 수 있다.
- 둘째, 피 평가자의 비용은 전혀 고려되지 않았다. 비용이 얼마인지에 대한 보고나 산출은 현재 - 적어도 스위스에는 - 존재하지 않는다.
- 셋째, 입찰과정에서 외부평가 수임자의 제안서작성을 위한 비용은 비용고찰에 전혀 포함되지 않았다. 그 이유는 수임자들의 제안서작성비용이 그들이 입찰에서 탈락되었다는 이유로 피해보상을 할 의무가 거의 또는 전혀 없기 때문이다. 이로 인해 이것은 행정업무에서 고려할 사항이 아니다.

마지막 세 번째 점은 소규모 평가의 경우 중요하다. 소규모 평가가 입찰될 경우 보통 조사된 서너 개 공급업체는 입찰이 되기까지 이미 약 20,000 프랑켄 정도 비용을 지출하게 된다. 약 50,000 프랑켄 평가예산에서 이것은 상당한 액수이다. 여기 우리는 외부평가수임자의 제안서작성을 위한 비용이 가격에 포함

되었는지 의문을 제기할 수 있다. 이것은 숙고될 만한 사항이고 장기적으로 이 비용을 평가비용에 포함시키는 것이 타당할 수 있다. 그런 장기적 변화를 기대하는 것 말고도 이 문제는 단기적으로 평가자수임자들과 평가집행자들에게 부정적인 효과를 가져다 줄 수 있는 문제이다.

그 이유로 다음 두 가지를 언급한다.

- 비드메르(1996, 7)는 그가 조사한 단기평가(1년 이하) 중에서 비용이 초과된 경우가 반이 넘는다고 지적한다. 초과비용은 대체로 평가집행자가 부담을 하게 된다 (주석 2). 달리 표현하면 소규모 평가의 경우 (입찰)경쟁으로 인해 계산상 현실성이 없는 낮은 평가예산으로 입찰을 하게 되고 이후 평가수행자에게 추가비용이 부메랑으로 되돌아오는 경우가 많다.
- 비드메르(1996, 9)에서는 그 밖에도 장기평가(1년 이상)보다 단기평가에서 평가의 질적인 불균형이 큼을 보여줬다. 이것은 예산부족이 평가의 질적 저하로 이어질 수 있음을 의미한다. 공급자의 제안서가 할당된 예산으로 실현되기 어렵다는 말이 된다. 이러한 고려로부터 소규모 평가의 경우 흔히 [입찰과정이 없는] 직접계약체결이 합리적이다. 의뢰자 입장에서나 수임자 입장에서나 입찰과정의 제안서작성 및 제안서선정을 위한 비용이 줄거나 없어지고 그 금액을 본격적인 평가지출에 사용되도록 할 수 있기 때문이다.

2. 평가비용은 어디까지 허용될 수 있나?

평가비용의 범주는 스위스의 경우 상당히 잘 평가될 수 있었다. 평가비용은 압도적 다수가 약 40,000에서 50,000 프랑켄 사이에 분포되어 있다. 비용규모가 더 큰 것은 드물고 연방차원에서는 - 발타자르(2007)의 수치에 따르면 - 전체평가의 10% 정도만 앞의 비용규모를 넘어서는 평가를 할 수 있도록 되어있다. 평가예산의 규모를 결정하는 각각의 이유에 대해서는 주어진 데이터를 통해서 알 수 있는 것이 거의 없다. 단지 이와 관련하여 다음 두 가지 결론으로 도출할 수 있었다.

- 단기로 수행되는 평가가 - 예측하지 못했던 것이 아니지만 - 상대적으로 저

렵하게 진행된다. 그 원인은 단기평가가 평가대상, 평가방법의 복잡성이 낮고 조사 대상에 대해 상대적으로 풍부한 사전정보가 제공하며 보고서작성에서 요구수준이 낮기 때문이다 (Widmer, 1996, 6-7).

- 평가의 종류 (집행평가 또는 효과평가)는 최소한 연방차원에서 조사한 데이터 상에서는 평가비용규모에 영향이 없다.

평가비용이 평가컨셉트에 따라 전체 프로그램비용에서 얼마나 차지하는 것이 적절한지 질문이 던져진다. 이 글에서 제시한 권고와 연방행정부의 실질적인 평가비용규모는 권고된 예산과 실제 지출사이에 큰 차이를 드러낸다. 비록 현존 데이터가 개별 프로그램과 조처를 기초로 하지 않았고 대부분 전체지출을 기준으로 하였다고 해도 그 차이는 눈에 띄게 크다. 그 이유가 권고수준이 너무 높은 탓인지 아니면 지출이 너무 낮기 때문인지 여기서 판단을 할 수 없었다. 하지만 이 글의 결론으로 최소한 평가지출에 대한 권고가 상이한 근거를 찾을 필요를 제기한다. 이와 함께 평가지출이 정규프로그램, 분배프로그램 (촉진프로그램이나 지원금) 또는 순수한 설득프로그램(정보와 상담)인가에 따라 상이하게 산출되는지 조사할 필요가 제기된다. 더 나아가 평가지출의 규모가 프로그램 성숙도에 따라 변동할 것이라는 지적도 역시 중요한 지적이라고 보인다. 즉 새로운 혁신적 조처 (전형적인 시범사업)은 아마도 이미 실험된 조처들보다 지출이 클 수 있다. 정확하게 어떤 차이가 나는지는 더 연구가 필요한 지점이다. 이 글에서는 평가의 비용-효용관계에 대해서 논하지 않았다. 만약 비용과 효용을 알 수 있다면 평가예산을 한계비용이 한계효용과 일치하도록 조정할 수 있을 것이다. 비드메르(1996, 5)는 평가의 한계효용체감을 가정한다. 하지만 이에 상응하는 한계효용곡선이 실제로 어떠한 것인지는 다음 두 가지 이유로 조사하기가 쉽지 않다.

- 첫째, 필요한 일정 비용들은 측정하기가 쉽지 않기 때문이다. 이러한 비용의 산출문제는 이미 직접비용에서 나타난다. 그리고 이 문제는 간접비용의 경우 더욱 큰 문제가 된다.
- 둘째, 평가의 효용을 환산하는 것은 한층 더 어려운 것이기 때문이다. 만약 평가의 효용을 기술하는 개념을 고찰하고자 하면 우리는 바로 그를 측정할

척도의 문제에 부딪히게 된다. 예를 들어 발테자르(2008, 12)는 평가의 효용과 관련 수단적, 개념적, 상징적과 절차적 사용을 구분했다. 슈토크만과 마이어(Stockmann/Meyer, 2010 20, 23)는 평가의 3개 목적을 다음과 같이 구분했다. 관리, 정당성확보 그리고 대사회적 계몽이다. 관리적 효용은 - 발타자르의 수단적 사용에 상응 - 그래도 측정 가능한 반면 다른 평가결과의 사용법을 금액으로 환산하는 것은 그 일이 극복 불가능한 것은 아니더라도 광범위한 작업이다. 이런 이유로 비용-효용 고려에서 구체적인 평가지출규모를 도출하는 것이 가능할 지 의문시된다. 이와 관련 발타자르(2007)의 데이터에 근거한 분석결과는 회의적이다. 그의 278개 평가에 대한 연구는 평가비용규모, 평가결과 활용규모 그리고 비용과 효용관계에 대해 활용 가능한 데이터를 제공한다. 마지막 두 가지에 대한 보고는 평가의뢰자에 의해 산출되었다. 그러나 이에 대한 연구결과는 평가비용규모와 결과활용규모 사이에 상관관계가 없다는 것을 보여주었다. 비용규모와 비용-효용관계도 상관관계가 없는 것으로 나타났다. 저가의 평가와 마찬가지로 고가의 평가도 그 결과가 집중적으로 사용되지 않을 수 있다. 규모가 큰 평가가 한계효용 체감으로 인해 더 위험할 수 있다는 논리, 즉 낮은 비용-효용 상관관계를 보일 것이라는 가설 또한 그의 데이터에서 입증되지 않았다.

〈주석〉

1. 미국 노동통계청 인플레이션산출 소비자물가지수 (CPI inflation calculator des U.S. Bureau of Labor Statistics)를 활용한 계산:
<http://www.bls.gov/bls/inflation.htm>.
2. 참고서적에서는 이 효과를 «승자 저주 (Fluch des Gewinners)»로 표현한다. 입찰경쟁의 승자가 입찰가로 제공할 수 없는 과도한 서비스 제공을 약속한 경우를 말한다.(Borrmann/Finsinger 1999, 318-331).

참 고 자 료

- Alkin, Marvin C./Stecher, Brian, 1983, A Study of Evaluation Costs[평가비용연구], in: Alkin, Marvin C./Solmon, Lewis C. [발행인], The Costs of Evaluation,[평가비용] Sage Publication, Beverly Hills/London/New Delhi.
- Alkin, Marvin C./Ruskus, Joan A., 1984, Reflections on Evaluation Costs: Direct and indirect[평가비용 설명: 직접과 간접], CSE Report No. 239, Center for the Study of Evaluation[평가연구센터], Los Angeles.
- Balthasar, Andreas, 2007, Institutionelle Verankerung und Verwendung von Evaluationen [제도적 고리와 평가의 활용], Rügger, Zürich/Chur.
- Balthasar, Andreas, 2008, Der Einfluss unterschiedlicher Institutionalisiertungsmodelle auf die Verwendung von Evaluationen: Erfahrungen aus der Schweiz [평가사용에 대한 제도화모델의 상이한 영향: 스위스의 경험], *Zeitschrift für Evaluation*, 2008/2, S. 243-271.
- Bormann, Jörg/Finsinger, Jörg, 1999, Markt und Regulierung [시장과 규제], Verlag Wahlen, München.
- Bundesamt für Gesundheit [연방보건청] (BAG), 1997, Leitfaden für die Planung von Projekt- und Programmevaluation[프로젝트 및 프로그램 평가 계획을 위한 지침], Bern
- Bussmann Werner, 1995: Evaluationen staatlicher Massnahmen erfolgreich begleiten und nutzen: Ein Leitfaden[국가적 조치의 평가를 성공적으로 안내하는 것과 이용하는 것: 지침서], Rügger, Zürich/Chur.
- DeGEval Gesellschaft für Evaluation e.V., 2008, Standards für Evaluation[평가 스탠다드], 4. unveränderte Aufl.(제4판), Mainz.
- Direktion für Entwicklung und Zusammenarbeit[개발협력국] (DEZA), 2008, Evaluationspolitik der Direktion für Entwicklung und Zusammenarbeit (DEZA) [개발협력국(DEZA)의 평가정책], Bern.
- European Commission DG Budget, 2004, Evaluating EU Activities, A Practical

- Guide for the Comission Services[유럽연합 활동평가, 위원회활동에 대한 실무안내], Brüssel.
- European Commission DG Regional Policy, 2008, EVALSED: The resource for the evaluation of Socio-Economic Development [사회경제개발 평가를 위한 자원], Brüssel.
- Frey, Bruno S., 2007, Evaluierungen, Evaluierungen ... Evaluitis[평가, 평가 평가병], *Perspektiven der Wirtschaftspolitik [경제정책의 전망]*, 2007/3, S. 207-220.
- Mülenkamp, Holger, 2009, Kosten-Nutzen-Analyse[비용-효용-분석], Oldenbourg, München.
- Stockmann, Reinhard/Meyer, Wolfgang, 2010, Evaluation, Eine Einführung[평가, 안내] Verlag Barbara Budrich, Opladen/Farmington Hills.
- Widmer, Thomas, 1996, Kurzevaluationen: Eine Alternative zu umfangreicheren Evaluationsstudien? Eine meta-evaluative Untersuchung, Nationales Forschungsprogramm 27[단기평가: 광범한 평가연구의 대안? 메타-평가 연구, 내셔널 연구프로그램 27] Reihe Kurzberichte[단기보고 시리즈], Bern.
- Widmer, Thomas, 2005, Leitfaden für Wirksamkeitsüberprüfungen beim Bund [연방의 효과평가를 위한 지침서], Bundesamt für Justiz/ (BJ), Bern.
- Widmer, Thomas/Landert, Charles/Bachmann, Nicole, 2000, Evaluations-Standards der schweizerischen Evaluationsgesellschaft[평가-스위스 평가사의 평가스탠다드] (SEVAL-Standards), <http://www.seval.ch/de/standards/index.cfm> (abgerufen am 04.02.2011).
- Widmer, Thomas/Rügg, Erwin/Neuenschwander, Peter, 2001, Stand und Aussicht der Evaluation beim Bund[연방의 평가 실태와 전망], Schlussbericht zuhanden der Schweizerischen Bundeskanzlei [스위스 내 각사무국 귀중 최종보고서], Bern.