

공공기관 경영실적 평가제도 실증 연구

2016. 12.

최 한 수 · 홍 우 형

서 언

2015년 기준으로 총 316개의 기관이 「공공기관의 운영에 관한 법률」에 의해 ‘공공기관’으로 지정되어 있다. 총자산 약 778조원, 총부채는 520조원, 당기 순이익 11조 4천억원이라는 수치에서 알 수 있듯이 이들 기관이 국민 경제에 차지하는 비중은 상당히 크다. 특히 이들 316개의 기관은 27만 9천여 명의 인력을 고용하고 있는데 이는 국가행정부 공무원 정원의 약 45% 수준에 해당하는 규모이다. 이처럼 공공기관이 국민경제에서 차지하는 경제적 비중과 고용 규모를 살펴보았을 때 이들 기관의 효율적 운영 문제는 비단 연구자뿐만 아니라 의회와 행정부의 정책결정자들의 주요한 관심 대상이 될 수밖에 없음을 알 수 있다.

정부는 공공기관의 경영효율성을 높이고 기관 내의 자율적인 책임경영 관행을 수립하기 위해 지난 1983년부터 공공기관의 경영실적평가제도를 도입하여 기관의 성과를 측정해왔다. 특히 평가결과에 따라 기관의 예산이나 기관 구성원의 보수가 달라질 수 있다는 점에서 경영평가제도가 공공기관에 갖는 영향력은 매우 크다. 정부 역시 이 평가제도를 통해 공공기관의 경영관행과 그 구성원들의 노력 수준에 변화를 주고 싶어한다. 이러한 면에서 경영평가제도는 인센티브 제도의 성격을 갖는다고 할 수 있다. 그동안 경영실적평가제도에 대한 여러 연구가 있었지만 이처럼 경영평가가 갖는 유인계약로서의 특징과 효과를 살펴본 연구는 부족했다. 본원의 최한수 박사와 홍우형 박사가 공동으로 진행한 『공공기관 경영실적 평가제도 실증 연구』는 공공기관 경영실적평가제도를 계약이론, 그중에서도 다중 임무하의 주인-대리인 간의 최적 계약이론의 관점에서 분석하고 있다.

본 연구에서는 최적 계약이론에 비추어 경영평가실적제도를 어떻게 설계하고 운영해 나가는 것이 사회적으로 바람직한 결과를 가져오는지, 이를 위

해 현행 제도를 어떠한 방향으로 개선해 나가는 것이 타당한지 등의 주제에 대한 이론적 분석과 실증 근거들이 제시되어 있다. 이러한 작업들이 정부의 공기업 관리정책의 효율성을 높이는 데 기여할 수 있기를 희망한다. 저자들은 본 연구의 제안 단계에서부터 원내 발표회에 이르기까지 유익한 의견을 주신 본원의 여러 박사들과 익명의 외부 논평자에게 감사의 말씀을 전한다. 그리고 데이터 관리와 경영평가제도의 연혁과 주요 내용을 정리해준 권선정 연구원에게도 감사드린다.

마지막으로 본 연구는 저자들 개인의 견해이며 반드시 기관의 의견을 반영한 것이 아님을 밝혀둔다.

2016년 12월

한국조세재정연구원

원장 박 형 수

요약 및 정책적 시사점

본 연구에서는 Holmström and Milgrom(1991)의 다중 임무하의 주인-대리인 간의 최적 계약이론을 이용하여 현행 공공기관 경평제도 설계방식의 적정성을 검토하고 이의 개선방향을 제시할 것이다. 규범적 관점에서 경평제도 설계의 적정성을 판단할 기준은 다음과 같다.

- (1) 도덕적 해이의 문제가 심각한 기관(예컨대 부채비율이나 복리후생비 지출 수준이 높은 기관)이 경영평가에서 불이익을 받아야 한다.
- (2) 이번 기에 경영평가 점수가 좋지 않은 공공기관의 경우 다음 기에 도덕적 해이 문제를 해결하기 위해 더 많은 노력(예컨대 부채비율이나 복리후생비 개선)을 해야 한다.
- (3) 경평결과가 효율성이나 공공성 제고를 위한 기관의 노력과 무관한 요인(기관의 크기나 업력 등)에 의해 결정되어서는 안 된다.
- (4) 경평결과가 효율성이나 공공성을 개선하기 위한 노력 수준에 대한 정확한 정보를 담고 있다고 보기 어려운 항목(리더십 · 책임경영 범주)에 의해 크게 영향을 받아서는 안 된다.

이러한 명제를 실증적으로 검증하기 위해 지난 2008~2013년도 사이에 약 100여 개의 공공기관을 상대로 진행된 경평결과를 이용하여 분석을 진행해 보았다. 그 주요 결과는 다음과 같다.

첫째, 현행 경평제도는 공공기관의 도덕적 해이를 규율하는 데 큰 역할을 하지 못하고 있다. 부채비율이 높은 기관의 경평등급이 그렇지 않은 기관에 비해 낮다는 증거는 확인되지 않았다. 또한 이번 기에 경평에서 좋지 않은 등급을 받은 공기업의 경우 다음 기에 부채비율이 하락하는 현상도 발견하지 못했다.

둘째, 평가결과가 해당 기관의 공공성이나 효율성 제고를 위한 노력과 상관없는 비본질적 요인에 의해 영향을 받는다. 자산의 크기가 평가 결과에 영향을 미쳤다. 즉 자산의 규모가 큰 공공기관이 중소규모의 자산을 보유하고 있는 공공기관에 비해 경쟁에서 보다 유리한 결과를 얻는 것이 확인되었다. 또한 공공기관의 업력도 경쟁결과에 영향을 미쳤다.

마지막으로, 경쟁결과가 리더십이나 책임경영처럼 본질적으로 측정오차를 클 뿐만 아니라 공공기관의 효율성이나 공공성을 위한 노력 수준에 대한 부정확한 정보를 담고 있는 요인에 의해 좌우되었다. 이는 해당 기관의 경쟁성과급의 크기가 사실상 리더십과 책임경영 항목에 대한 평가 결과에 의해 크게 영향을 받고 있다는 것을 의미한다.

본 연구의 정책적 시사점은 다음과 같다. 첫째, 현행 경쟁제도는 공공기관으로 하여금 변화하는 경영환경에 대한 신속한 대처를 어렵게 만드는 방식으로 설계되어 있다. 이러한 대처 노력에 원칙적으로 어떠한 가중치도 부여하고 있지 않기 때문이다. 이러한 평가지표의 경직성에서 비롯된 비효율성으로 인해 경쟁제도의 유인계약적 역할은 크게 약화될 수밖에 없다. 따라서 정부가 현행 경쟁제도의 평가들을 그대로 유지하고 싶다면 경쟁제도의 목적을 공공기관의 도덕적 해이에 대한 통제 장치로 제한하고 운영해나가야 한다.

둘째, 경쟁결과가 비본질적 요인에 의해 영향받는 상황을 개선하기 위해 평가 지표를 대폭 축소하고 평가내용도 최대한 단순해지도록 재설계해야 한다. 실제 기관의 노력 수준을 반영한다고 보기 어렵거나 기관의 고유 목적 사업을 통해 달성하고자 하는 공공성과 직접 관련이 없는 지표들은 평가항목에서 과감하게 제외할 필요가 있다. 고객만족도, 리더십, 정부 권장정책 지표 등이 그 대표적 예이다. 아울러 재무정보가 아닌 ‘조정된’ 재무정보나 ‘목표치 결과’에 근거한 평가방식은 기관의 장기적 재무 건전성에 대한 부정확한 정보를 제공해주는 결과를 가져온다는 점에서 시급히 개선할 필요가 있다.

셋째, 현재처럼 경쟁성과급의 크기가 리더십이나 책임경영처럼 공공기관의 노력 수준에 대한 부정확한 정보를 제공해주는 지표에 의해 결정되는 것이 계속 지속된다면 경쟁성과급의 크기를 줄여나가거나 극단적인 경우 아예 이를 폐지하는 것이 사회적으로 최적일 수 있다.

목 차

I. 서론	11
1. 연구의 배경	11
2. 연구의 주요 결론	16
3. 연구의 정책적 시사점	18
II. 경영평가의 제도적 배경 및 선행 연구 정리	22
1. 경영평가제도의 내용과 특징	22
2. 선행연구 정리	25
III. 경영평가제도에 대한 이론적 고찰	27
1. 경영평가제도의 경제학적 근거	27
2. 계약(contract)으로서의 경영평가제도	34
3. 복수의 과제를 대리인에게 맡기는 경우의 최적 계약이론	38
IV. 경영평가제도 설계의 적절성과 관련된 명제	43
1. 도덕적 해이의 통제장치로서의 경영평가제도	43
2. 평가지표설계의 적절성	45
3. 경영평가성과급 설계의 적절성	47
V. 실증분석 결과	51
1. 공공기관 경영평가 결과에 대한 기초통계 분석	51
2. 검증 명제 실증분석	65

VI. 결론 및 정책적 시사점	81
1. 결론 요약	81
2. 정책적 시사점 및 제도 개선방향	82
참고문헌	87

표목차

〈표 II-1〉 2010년 공공기관 경영평가제도 개편 내용	23
〈표 III-1〉 2015년 평가지표 개선 요청 및 반영 결과	30
〈표 V-1〉 표준화 순위 변동의 기초통계(2007 ~ 2013년)	54
〈표 V-2〉 연도별 표준화 순위 변동의 기초통계량	54
〈표 V-3〉 표준화 순위 전이확률행렬(transition probability matrix)	59
〈표 V-4〉 표준화 순위 변동의 연도별 전이확률행렬	59
〈표 V-5〉 경영평가 범주 및 평가지표 체계	60
〈표 V-6〉 경영평가 주요지표의 기초통계(2008 ~ 2010년)	64
〈표 V-7〉 경영평가 주요지표의 기초통계(2011 ~ 2013년)	64
〈표 V-8〉 경영평가 순위의 결정요인 분석	66
〈표 V-9〉 연도별 경영평가 순위의 결정요인 분석	68
〈표 V-10〉 기관유형별 분석대상 공공기관	70
〈표 V-11〉 주요 재무정보의 기초통계량	71
〈표 V-12〉 공공기관의 재무정보가 경영평가 결과에 미치는 효과	73
〈표 V-13〉 경영평가 결과가 주요 재무정보에 미치는 효과(전체 표본)	74
〈표 V-14〉 경영평가 결과가 주요 재무정보에 미치는 효과(기금관리형 제외 표본) ..	75

그림목차

[그림 II-1] 공기업·준정부기관(중소형 제외)의 지표 및 가중치 기준	24
[그림 II-2] 준정부기관 중 중소형 유형의 지표 및 가중치 기준	25
[그림 V-1] 표준화 순위 변동의 분포	55
[그림 V-2] 연도별 표준화 순위 변동의 분포	56
[그림 V-3] 연도별 경영평가 계량지표 분포	63
[그림 V-4] 연도별 경영평가 비계량지표 분포	63

I. 서론

1. 연구의 배경

공공기관의 경영실적 평가제도(이하 ‘경평제도’로 약칭)는 공공기관의 경영실적을 객관적으로 평가하여 그 결과를 기관장의 임기 및 직원의 성과급과 연동시키는 제도이다(기획재정부, 2016a). 이는 공공기관의 경영효율성을 향상시키고 기관 내에 자율적이고 책임있는 경영 관행을 수립하기 위해 만들어진 것이다(기획재정부, 2016a).

공공기관 경평제도의 연원은 1983년 제정된 「정부투자기관관리기본법」상의 정부투자기관 경영평가제도로까지 거슬러 올라갈 수 있다(이원희·라영재, 2015). 그러나 당시의 경평제도는 현재의 제도와 여러 면에서 차이가 있다. 우선 평가를 받는 기관의 수가 십 여 개에 불과했고 평가항목도 제한적이었다. 따라서 현재 사용되는 분석방법과 평가틀을 갖춘 경평제도는 2007년 「공공기관의 운영에 관한 법률」(이하 ‘공운법’으로 약칭) 시행 이후 시작되었다고 보는 것이 옳다.

경평제도의 도입 효과와 그 유지 필요성에 대해서는 상반된 견해가 존재한다. 현행 경평제도는 공공기관에 현대적 경영 관리 기법을 적용시켰다는 점에서 매우 선진적인 제도이며(장지인 외, 2013) 이로 인해 다른 개발도상국에서 참고의 대상(benchmarking)이 되고 있다는 견해가 있는 반면 기관들 간의 차이를 고려하지 않은 평가기준의 획일적인 적용으로 인해 오히려 기관운영에 있어 비효율성을 발생시키고 단지 관료의 공공기관 통제수단으로서만 기능한다는 반대의 견해(라영재·윤테범, 2013)도 존재한다.

다음의 인천공항공사 일화는 현재의 경평제도가 직면한 미묘한 상황을 잘 보여준다. 2016년 인천공항에서 항공기 160여 대의 출발을 지연시킨 수하물

대란과 밀입국 사태가 연달아 발생하면서 인천공항공사의 보안 및 안전 관리에 문제가 있다는 내외의 지적이 잇따랐다.

이에 정일영 신임 인천공항공사 사장은 한 언론과의 인터뷰에서 공기업 경영평가 점수를 일부 포기하더라도 보안 및 안전 관련 비용을 충분히 쓰겠다는 계획을 밝혔다.¹⁾ 일반인의 관점에서 보면 이 발언은 이상하게 들릴 수 밖에 없다. 보안 및 안전 영역은 국제공항을 운영하는 인천공항공사의 핵심 사업 영역 중 하나라고 할 수 있다. 그럼에도 불구하고 인천공항공사가 주요 사업에 대한 투자를 진행하였다는 사실이 기관 평가에 있어 불이익을 주는 요인으로 작용한다면 과연 경영이 추구하는 목적이 무엇인가라는 의문이 드는 것은 지극히 당연하다.

앞에서 예시한 인천공항공사의 문제는 일면 변화된 환경에서 긴급하게 요구되는 과제에 대한 대처 정도를 평가에 반영할 수 없는 평가지표의 경직성 문제를 보여주는 일화인 것처럼 보인다. 그러나 다른 한편으로 이는 경평제도가 오히려 경영의 자율성과 대응력을 떨어뜨리는 족쇄로 작용하고 있음을 보여주는 사례로 해석될 수도 있다.

본 연구는 바로 이러한 문제에 대한 답을 찾는 노력의 일환이다. 본 연구에서는 경평제도를 정보(information)와 유인(incentive)이라는 경제학의 기본 분석틀 안에서 이해하고 그 기능과 역할을 분석해보려 한다. 경제학적 관점에서 경평제도의 존재 이유를 어떻게 정당화할 수 있을까? 유인계약(incentive contract)의 시각에서 경평제도를 어떻게 설계하는 것이 사회적으로 최적의 결과를 발생시킬까? 공공기관 임원과 직원의 성과급을 경평결과에 연동되도록 설계하는 것이 과연 공공기관의 효율성과 공공성을 높여주는 결과를 가져올까? 경평제도를 설계함에 있어 효율성과 공공성을 동시에 추구해야 하는 공공기관의 수임목적(mandate)의 이중성을 어떻게 고려해야 하는 것인가?

1) 「공기업 경영평가 점수 포기하더라도 인천공항 보안 비용 충분히 쓸 것」, 『조선일보』, (2016. 2. 3.) http://news.chosun.com/site/data/html_dir/2016/02/03/2016020300349.html(접속일: 2016. 10. 17).

본 연구의 일차적 목적은 이러한 질문을 새로운 시각으로 재해석하여 올바른 해답을 도출하는 데 도움이 되는 적절한 이론적 분석틀을 찾고 이를 바탕으로 제도의 개선안을 도출해내는 데 있다. 본 연구의 방법론과 분석틀을 구체적으로 설명한다면 다음과 같다. 본 연구에서는 공공기관 임직원의 보수가 경평결과에 영향을 받는다는 점에 착안하여 경평제도를 주인과 대리인 사이에 체결되는 유인계약의 하나로 이해할 것이다. 특히 본 연구에서는 공공기관 경평제도가 갖고 있는 특성들을 가장 잘 분석해주는 이론적 틀로써, Holmström과 Milgrom이 지난 1991년 *Journal of Law, Economics, & Organization*에 발표한 논문("Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design")에서 처음 제시된 다중(多重)임무하의 주인-대리인 사이의 최적 계약이론을 사용할 것이다. 여기에 계약이론(contract theory)의 일반적 통찰을 더하여 최종적으로 현행 경평제도의 설계의 적정성을 판단하는 데 도움을 주는 몇 가지 규범적 명제들을 도출할 것이다. 그리고 실제 경평 데이터를 이용하여 현행 제도가 이러한 규범적 명제에 비추어 얼마나 효율적으로 설계되고 운영되고 있는지를 살펴볼 것이다.

그 구체적인 내용은 다음과 같다. 먼저 현재의 경평제도가 공공기관의 도덕적 해이(moral hazard)의 통제장치로서 그 기능을 제대로 수행하고 있는가를 살펴볼 것이다. 왜 도덕적 해이인가? 그것은 현행 공공기관을 상대로 한 경평제도가 다음의 두 가지 측면에서 노동시장에서 흔히 관찰되는 CEO의 스톡옵션계약(stock option contract)보다 보험시장에서 흔히 접하는 약관(예컨대 화재보험)과 구조적으로 보다 유사하기 때문이다.

먼저 민간기업의 경우 기업의 성과는 흔히 주가가치라는 '결과'에 따라 측정된다. 반면에 공공기관의 경우 그 성과는 결과 외에도 그 '투입'과 '과정'에 의해서도 평가되고 있다. 대표적 유인계약이라고 할 수 있는 스톡옵션계약이 이런 식으로 설계되는 경우는 거의 찾아볼 수 없다.

또한 민간기업의 성과는 주가(株價)에 의해 평가받는다. 주가는 현재 경영진이 기업을 운영하는 과정에서 빈번히 발생하는 우발성(contingency)과 불확실성(uncertainty)에 대해 적절히 반응했는가에 의해 실시간으로 변한다.

이런 측면에서 주가는 매우 유연한 평가지표이다. 반면에 공공기관의 성과는 정부가 판단한 몇 개의 성과지표에 따라 평가받는다. 그 지표는 사전에 고지되고 원칙적으로 수정되지 않는다. 따라서 공공기관에 대한 경평결과는 지표 공표 이후 발생한 우발성과 불확실성에 대한 대응의 적정성을 보여주는 정보가 원칙적으로 결여되어 있다. 오히려 정부는 평가의 공정성이나 일관성을 지키기 위해 경직된 평가지표를 선호한다.

즉 경평제도는 사전에 고지된 기준에 해당하는 피평가기관의 특정 행위만을 평가하는 구조를 가지고 있다. 이러한 구조 역시 민간의 스톡옵션계약에서는 찾아볼 수 없다. 대신 이러한 행태의 계약은 보험계약에서 자주 목격된다. 보험 약관에서는 보험계약자의 도덕적 해이와 밀접하게 관련된 행위 유형들을 사전에 열거한 뒤 그 행위의 준수 여부에 따라 보험료를 할인(혹은 할증)하는 방식을 채택하고 있다. 이러한 형태의 계약은 보험계약자의 도덕적 해이를 감시함으로써 사회적으로 보다 바람직한 결과를 발생시킨다는 점에서 그 필요성을 인정받아 왔다. 따라서 마찬가지로 공공기관의 도덕적 해이를 보여주는 평가지표를 사전에 공지하고 이의 이행 여부에 따라 공공기관에게 성과급 형식의 금전적 이득(혹은 손해)을 제공해주는 현행 경평제도 역시 만약 그 목적이 도덕적 해이를 억제하기 위한 것이라면 어느 정도의 합리성을 인정받을 수 있다. 따라서 본 연구에서는 현재의 경평제도가 실제 공공기관의 도덕적 해이와 관련된 행위를 통제하는 역할을 해왔는지를 데이터를 통해 검증할 것이다.

경평은 도덕적 해이에 대한 통제장치로서의 역할과 별개로 기관의 구성원으로 하여금 기관의 설립 목적에 부합하는 역할을 하도록 노력하게 만드는 유인계약의 역할 역시 요구받는다. 이처럼 경평제도가 유인계약으로 역할을 제대로 하기 위해서는 무엇보다도 경영성과에 대한 객관적 측정이 가능하다는 것이 전제되어야 한다. 대리인이 좋은 성과를 내기 위해 충분한 노력을 기울였으나 성과측정이 제대로 이루어지지 않아 노력과 성과 사이의 대응관계가 무너지고 오히려 비본질적 요인이 기관의 성과를 결정한다면 경평의 유인계약적 기능은 제대로 작동하지 않을 것이다.

따라서 기존 경평제도의 역할에 대한 두 번째 검증 지점은 경평결과가 해당 기관의 효율성이나 공공성을 제고하기 위한 노력과 별 상관없는 요인에 의해 어느 정도 영향을 받고 있는가를 살펴보는 것이다. 본 연구에서는 해당 기관의 규모나 업력(業歷)을 노력과 무관한 비본질적 요인의 사례로 보고 이것이 평가에 어떤 영향을 미치는지를 살펴볼 것이다.

기관의 노력과 무관한 비본질적 요인이 경평결과에 영향을 미치고 있는지 여부를 추론하는 또 하나의 방법은 경평순위의 변동성을 살펴보는 것이다. 그 이유를 직관적으로 설명한다면 다음과 같다. 매년 실시되는 경평의 제도적 특성을 감안했을 때, 만약 경평결과가 그 기관의 효율성과 같은 본질적 요인에 의해 결정된다고 가정한다면 단기 변동이 크지 않을 것이라고 생각할 수 있다. 반대로 그 결과가 우연이나 비본질적 요인에 영향을 받는다고 생각한다면(예컨대 우연히도 평가 서류를 잘 준비하거나 특정 기관에 호의를 갖고 있는 인물이 평가단에 들어있는 경우) 순위 변동이 연도별로 빈번하게 발생할 것이고 그 폭도 비교적 클 것이다.

마지막 검증 명제는 Holmström and Milgrom(1991) 모형의 가장 중요한 예측과 관련이 있다. 이는 공공기관의 성과평가가 경영효율성(혹은 사업성)과 공익성처럼 각기 독립적이면서 측정의 난이도가 상이한 과제에 대해 이루어질 때 유인계약을 어떻게 설계하는 것이 사회후생의 측면에서 최적의 결과를 가져오는가의 문제이다.

대리인이 취할 수 있는 행동의 항목이 복수인 경우 그중 일부는 투입한 노력의 정도를 관측하기가 용이한 반면 다른 일부는 그렇지 않은 경우가 있다. Holmström and Milgrom(1991)의 주장에 따르면 이 경우의 최적 계약은 대리인의 보상체계를 두 가지의 성과 척도와 매우 약하게 연동시키는 것이다(김빛마로, 2016). 이는 경평 성과급을 설계함에 있어 성과를 객관적으로 측정하는 것이 어려운 항목에 연동시켜서는 안 된다는 것을 의미한다. 경평 성과급이 리더십 및 책임경영 지표처럼 투입 노력의 수준을 파악하기 어려운 항목에 의해 크게 영향을 받을 경우 대리인의 투입 노력 배분에 왜곡이 발생하기 때문이다. 본 연구에서는 경평성과급 계약의 최적 설계와 관련하

여 기관별 성과급의 격차가 리더십 및 책임경영 지표에 의해 어느 정도 영향을 받는가를 살펴볼 것이다.

결론적으로 본 연구에서는 다음의 4가지 명제를 통해 경평제도가 공공기관의 경영효율성과 공공성을 제고하는 수단으로써 잘 작동할 수 있도록 설계되어 운영되고 있는지를 살펴볼 것이다.

1. 도덕적 해이의 문제가 심각한 기관(예컨대 부채비율이나 복리후생비 지출 수준이 높은 기관)이 경영평가에서 불이익을 받아야 한다.
2. 이번 기에 경영평가 점수가 좋지 않은 공공기관의 경우 다음 기에 도덕적 해이 문제를 해결하기 위해 더 많은 노력(예컨대 부채비율이나 복리후생비 개선)을 해야 한다.
3. 경평결과가 효율성이나 공공성 제고를 위한 기관의 노력과 무관한 요인(기관의 크기나 업력 등)에 의해 결정되어서는 안 된다.
4. 경평결과가 효율성이나 공공성을 개선하기 위한 노력 수준에 대한 정확한 정보를 담고 있다고 보기 어려운 항목(리더십·책임경영 범주)에 의해 크게 영향을 받아서는 안 된다.

2. 연구의 주요 결론

2008년부터 2013년 사이 매년 100여 개의 공공기관을 대상으로 한 경평에 대한 분석 결과는 다음과 같다. 첫째, 현행 경평제도는 공공기관의 도덕적 해이의 정도를 낮추는 역할을 제대로 수행하지 못하고 있었다. 부채비율이 높은 기관의 경평등급이 그렇지 않은 기관에 비해 낮다는 증거는 확인되지 않았다. 또한 이번 기에 경평에서 좋지 않은 등급을 받은 공기업의 경우 다음 기에 부채비율이 하락하는 현상도 발견하지 못했다. 도덕적 해이의 또 다른 지표로 복리후생비를 사용한 분석에서도 동일한 결과가 나왔다. 공공기관의 복리후생비 지출 수준이 경평결과에 유의미한 차이를 발생시키지 않았다. 즉 복리후생비의 지출 수준이 높다고 해서 경평결과에서 불이익을 받

는 현상은 확인되지 않았다.

둘째, 경평결과가 해당 기관의 효율성이나 공공성을 높이기 위한 노력과 무관한 요인에 영향을 받는 현상이 목격되었다. 자산의 크기가(10%의 유의 수준에서) 평가 결과에 영향을 미쳤다. 즉 자산의 규모가 큰 공공기관이 중소규모의 자산을 보유하고 있는 공공기관에 비해 경평에서 보다 유리한 결과를 얻는 것이 확인되었다. 또한 공공기관의 업력(age)도 경평결과에 영향을 미쳤다. 즉 해당 기관의 설립년도가 오래되었을수록 평가에서 좋은 결과를 받았다. 현재와 같은 틀을 갖춘 경평의 시행 역사가 최소 10년 이상 되었다는 점을 감안하면 이는 업력이 오래된 기관일수록 경영평가의 경험이 내부에 축적되어 있을 가능성이 높고 이것이 경평결과에 영향을 미친 것으로 해석할 수 있다.

경평순위의 변동성을 통해 경평결과가 투입노력의 수준과 무관하게 결정되는지를 살펴본 결과는 다음과 같다. 전이확률행렬(transition probability matrix)을 통해 이러한 현상이 실제 관찰되는지를 살펴본 결과 특정 기관이 일 년 전에 경평결과 기준으로 상·중·하 그룹 중 어디 하나에 속해 있고 그 다음 해에도 동일한 그룹에 머무를 확률은 45%에 불과했다. 다시 말하면 이는 개별 기관들이 과거 자신들의 순위 그룹으로부터 이탈하여 다른 그룹으로 옮겨갈 확률이 55%나 된다는 것을 의미한다. 이는 상당히 높은 수준이다.

특히 전년도에 상위 그룹에 속해 있던 기관이 다음 연도에 하위 그룹으로 떨어지거나 반대로 하위 그룹에 속해 있던 기관이 상위 그룹으로 올라갈 확률도 13%나 되었다. 즉 100개 공공기관 중에 13개 기관은 1년이라는 짧은 기간 사이에 큰 폭의 순위 변동을 경험하고 있었다. 이처럼 경평결과에 따른 순위 변동이 연도별로 빈번히 발생하고 그 폭 또한 비교적 크다는 사실로부터 우리는 현재의 평가제도가 비본질적 요인 혹은 지표 개선을 위한 기관의 단기적 노력에 상당한 영향을 받고 있다고 추론할 수 있다.

마지막으로, 경평결과가 리더십이나 책임경영 지표처럼 본질적으로 측정 오차(measurement error)가 클 뿐 아니라 기관의 노력 수준에 대한 신뢰할

만한 정보를 제공해주지 못하는 항목에 의해 어느 정도 좌우되는지 살펴보았다. 분석 결과 해당 공공기관의 경평 순위는 리더십이나 책임경영 지표의 순위에 가장 크게 영향을 받는 것으로 밝혀졌다. 경평순위가 결정되는 과정에서 주요사업의 성과나 경영효율성 개선 노력보다는 리더십이나 책임경영 지표의 중요도가 상대적으로 매우 크게 작동한다는 사실을 알 수 있었다.

이를 연도별로 살펴보아도 그 결과는 크게 달라지지 않았다. 2008년부터 2013년 사이에 2011년 한 해를 제외하고 해당 기관의 경평순위를 결정함에 있어 비계량적인 방식으로 측정된 리더십과 책임경영 지표의 중요성이 다른 지표에 비해 상대적으로 크다는 사실을 확인할 수 있었다. 2011년을 기점으로 평가항목에 큰 변화가 있었던 점을 감안하여 분석의 시계(time horizon)를 2011년 이전과 이후 두 기간으로 나누어 살펴보아도 그 결과는 크게 달라지지 않았다. 이는 해당 기관의 경평성과급의 크기가 사실상 리더십과 책임경영 항목에 대한 평가 결과에 의해 크게 영향을 받고 있다는 것을 의미한다.

3. 연구의 정책적 시사점

앞에서 제시한 실증 분석결과를 바탕으로 본 연구에서는 다음과 같은 제도개선방안을 제안하고자 한다. 첫 번째 제안은 경평제도의 기능(혹은 목적)을 명확하게 해야 한다는 것이다. 현행 경평제도는 공공기관으로 하여금 변화하는 경영환경에 대한 신속한 대처를 어렵게 만드는 방식으로 설계되어 있다. 이러한 대처 노력에 원칙적으로 어떠한 가중치도 부여하고 있지 않기 때문이다. 이러한 평가지표의 경직성에서 비롯된 비효율성으로 인해 경평제도의 유인계약적 역할은 크게 약화될 수밖에 없다.

따라서 만약 정부가 현행 경평제도의 평가틀을 그대로 유지하고 싶다면 이를 공공기관의 도덕적 해이에 대한 통제 장치의 하나로 보고 운영해나가야 한다. 경평제도의 평가대상을 부채비율이나 복리후생비와 같은 몇 가지 객관적 재무지표들로 한정된 뒤 그 결과에 따라 기관에게 인센티브(혹은 경

우에 따라서는 페널티)를 부과하는 것이다. 부채비율이나 복리후생비 지표는 실제 공공기관의 도덕적 해이 수준을 비교적 정확히 보여주는 지표이다. 또한 공공기관의 도덕적 해이로 인해 기관의 부채비율이나 복리후생비 같은 운영비용이 상승하는 것을 방지할 경우 이후 국가재정의 상당한 부담 요인이 될 수 있다는 점에서 이의 규제 필요성도 충분히 인정된다.

두 번째 제안은 지표의 설계와 관련된 이슈이다. 만약 정책결정자가 경평을 통해 기관의 성과 향상 노력을 제고하려 한다면 현행 제도를 대폭적으로 손질해야 한다. 계약이론에 따르면 유인계약이 효과적으로 작동하기 위해 필요한 가장 중요한 전제조건 중 하나는 성과를 측정하는 지표들이 대리인의 노력 수준을 가능한 한 정확히 반영해야 한다는 것이다. 즉 경평결과가 공공기관의 노력과 무관한 요인에 의해 결정된다면 경평의 유인계약으로서의 기능은 크게 반감된다.

이를 위해서는 무엇보다도 평가 기준을 간소화하여 기관의 평가 부담을 덜어줄 필요가 있다. 본 연구의 발견처럼 기관의 크기나 업력이 평가결과에 영향을 준다는 것은 현행 제도가 피평가기관에 필요 이상의 평가부담을 주고 있다는 것을 의미한다. 특히 공공기관들 사이에 비계량 점수를 잘 받기 위해 많은 예산과 인력을 투자하여 작성된 보기 좋은 보고서가 요구된다는 인식(김주찬·이민창, 2014)이 만연해 있는데 이러한 잘못된 인식이 가져올 비효율성을 제거하기 위해 경평지표를 대폭 축소하고 그 내용도 최대한 단순하게 정해지도록 설계할 필요가 있다.

또한 실제 재무정보가 아닌 ‘조정된’ 재무정보나 ‘목표치 결과’에 근거한 평가방식은 기관의 장기적 재무 건전성에 대한 부정확한 정보를 제공해주는 결과를 가져온다는 점에서 시급히 개선되어야 한다. 일부에서는 기관이 통제할 수 없는 위험을 고려하기 위해 이러한 평가방식이 불가피하다고 주장한다. 그러나 평가과정에서 기관이 통제할 수 없는 위험으로 인해 발생하는 경평결과의 객관성과 공정성 시비는 지표의 자의적 조정이 아니라 경평성과급의 크기와 그 민감도를 조정하는 방식으로 해소되어야 한다.

아울러 실제 기관의 노력 수준을 반영한다고 보기 어렵거나 기관의 고유

목적 사업을 통해 달성하고자 하는 공공성과 직접 관련이 없는 지표들은 평가항목에서 과감하게 제외할 필요가 있다. 고객만족도, 리더십(2013년 기준 지표), 전략기획 및 기관혁신(2015년), 정부 권장정책 지표 등이 그 대표적 예이다. 경평결과가 이러한 지표들에 크게 영향을 받을수록 경평의 유인계약적 기능은 무력화되어 버린다.

마지막으로 경평성과급 설계의 적정성에 대해 보다 꼼꼼한 검토가 필요하다. 현행 경평제도는 공공기관으로 하여금 효율성과 공익성이란 과제를 균형있게 추구할 것을 요구하고 있다. 그리고 이를 위해 경평성과급을 두 목표의 성과지표의 결과에 연동시켰다. 문제는 두 지표의 정확성의 정도가 매우 상이하다는 것이다. Holmström and Milgrom(1991)의 이론이 의미하는 바는 이 경우 측정이 어려운 과제의 성과에 따라 보수를 지급하는 것을 피해야 하며 상대적으로 성과 측정이 용이한 나머지 과제에 대한 성과급 또한 낮추는 것이 공공기관으로 하여금 두 목표를 균형 있게 수행하도록 만든다는 것이다. 이는 현행 경평성과급의 비중을 점차 줄여나가거나 경우에 따라서는 아예 폐지하는 것이 사회적으로 최적일 수 있다는 것을 의미한다.

그동안 주로 경영학자와 행정학자를 중심으로 경평제도에 대한 많은 연구들이 진행되어 왔다. 그러나 본 연구처럼 경평의 본질을 일종의 주인-대리인 간의 유인계약으로 이해하고 제도 설계의 적정성을 분석한 연구는 필자들이 아는 범위 내에서 이 연구가 처음이다. 이러한 접근 방법의 가장 큰 장점은 지난 40년 동안 이론과 실증 경제학자에 의해 축적되어 온 계약이론의 분석틀과 그 정책적 함의를 경평의 분석과 개선 작업에 활용할 수 있다는 점이다. 이러한 과정을 통해 우리는 현행 제도 설계 문제를 보다 과학적으로 바라보게 되었으며 또한 제도를 구성하는 여러 지표들 간의 내적 긴장관계(trade-offs)를 정확히 인식하게 됨으로써 이후 경평제도를 개선해나가는 데 여러 시사점을 얻을 수 있게 되었다. 특히 공공기관 경평제도가 지방공기업이나 금융형 공공기관의 성과평가의 준거점이 되고 있는 현실을 고려한다면 본 연구는 공공영역의 성과평가의 지평을 확장하는 데 기여할 수 있다.

본 보고서의 주요 구성과 내용은 다음과 같다. 제Ⅱ장에서는 공공기관 경

평제도의 특징과 그 선행 연구를 간략하게 소개할 것이다. 제Ⅲ장에서는 이론적 관점에서 경평제도가 갖는 특징들을 살펴볼 것이다. 특히 Holmström and Milgrom(1991)의 모델을 이용하여 경평성과급이 유인계약의 하나로 제대로 작동하는 데 필요한 전제조건들과 구성요소들이 무엇인지 살펴볼 것이다. 제Ⅳ장에서는 규범적인 관점에서 현행 경평제도의 적정성을 진단하는데 필요한 지표들을 명제의 형식으로 제시할 것이다. 제Ⅴ장에서는 앞에서 제시된 명제들을 실제 경평 데이터를 이용하여 검증해볼 것이다. 제Ⅵ장에서는 경평의 개선방향을 제시할 것이다.

Ⅱ. 경영평가의 제도적 배경 및 선행 연구 정리

1. 경영평가제도의 내용과 특징

한국에서 공공기관 경영성과에 대한 평가는 1968년에 처음 시도된 것으로 알려지고 있다. 그러나 당시만 하더라도 대부분의 공공기관이 정부의 강한 경영통제 아래 놓여 있었다. 당연히 이들 공공기관의 예산과 사업계획 그리고 인사의 자율성은 지극히 제한된 수준에서 가능했다. 이와 같은 상황에서 공공기관의 경영성과를 평가한다는 것은 여러 한계를 드러낼 수밖에 없었다.

따라서 본격적인 의미의 공공기관 경평제도의 도입과 운영은 1984년 「정부투자기관관리기본법」 제정 이후부터로 보는 것이 옳다. 1983년부터 2003년까지는 「정부투자기관관리기본법」하에 규정된 평가제도가 운영되었다. 평가대상기관의 수는 연도별로 일정하지 않았다. 상황에 따라 특정 기관이 제외되거나 추가되는 경우가 많았기 때문이다. 대략 1983년부터 2003년 사이 그 수는 적게는 12개에서 많게는 25개 정도를 유지하였다. 1999년에는 기존의 기관평가 외에 사장경영계약 및 사장 평가가 도입되면서 기관장에 대한 평가도 시작되었다.

2004년부터 2006년까지는 평가대상기관에 88개의 정부산하기관이 추가되었다. 「정부산하기관관리기본법」이 제정되면서 기존의 「정부투자기관관리기본법」 외에 「정부산하기관관리기본법」을 통해서도 공공기관 경영평가제도가 운영되기 시작했다. 평가유형 또한 기존 2~5개에서 정부투자기관 3개, 정부산하기관 8개로 구체화되고 그 수도 늘어났다(한국조세재정연구원, 2010).

2007년에 공운법이 제정되면서 공공기관에 대한 관리가 보다 체계화되었다. 공기업과 준정부기관으로 평가를 받는 기관의 범위가 명확해졌고 2008년부터는 기관과 기관장을 분리하여 별도의 기준으로 이들의 성과를 평가하기

시작하였다. 2010년에 들어와서 공공기관 경영평가제도에 대폭적인 개선이 있었다. 그 주요 내용은 <표 II-1>과 같다.

〈표 II-1〉 2010년 공공기관 경영평가제도 개편 내용

	목적	개편내용
1	공기업의 경쟁력 강화	당해기관의 전년도 실적과 비교가 아닌 글로벌 기업과의 실적 비교
2	사회적 책임 강화	사회공헌지표(일자리 창출, 동반성장 등) 신설
3	일자리 창출	일자리 창출에 대한 실질적 성과 평가 강화
4	공공기관의 평가부담 완화	지표 축소 및 간소화
5	기관별 맞춤형 평가	기관이 평가지표를 선택할 수 있게 개편
6	공정한 평가	평가단 인사의 다양화 (민간기업 CEO 등 다양한 인사를 포함)
7	방만경영 방지	방만경영 적발 시 처벌 강화
8	시대상황 반영	사회적 이슈가 되는 과제에 대한 평가 강화
9	국민 만족도 향상	국민들의 평가를 반영

자료: 기획재정부 보도자료, 「공공기관 경영평가제도 대폭 개편」 2010. 12. 29.

2011년부터는 공기업과 준정부기관을 공기업 I, 공기업 II, 기금관리형, 위탁집행형, 강소형으로 좀 더 세분화하여 분류한 뒤 평가하고 있다. 평가의 범주는 크게 경영관리와 주요사업으로 나눌 수 있다. 경영관리에는 ▲ 경영 전략 및 사회공헌 ▲ 업무효율 ▲ 조직·인적자원 및 성과관리 ▲ 재무예산 관리 및 성과, 보수 및 복리후생관리의 하부 범주가 있다. 주요사업은 기관 고유 주요사업의 계획 수립 및 활동, 결과로 나누어 각각의 항목에 대한 평가를 진행하고 있다(기획재정부, 2016a).

기관의 평가결과는 S에서 E까지 6개의 등급이 있는데 그 결과에 따라서 성과급 차등지급, 인사조치, 다음 연도 기관의 예산반영 등의 후속조치가 뒤 따른다. 경평제도가 갖고 있는 유인계약의 특징은 사실은 평가 결과에 따라 해당 공공기관의 기관장과 그 직원들의 성과급이 차등 지급된다는 사실에서 분명히 드러난다. ‘2016년 공기업·준정부기관 예산편성 지침’에 따르면 공

기업은 월 기본급의 250%, 준정부기관은 기준월봉의 100%, 종전 정부투자기관은 월 기본급의 300% 이내에서 경영평가 성과급을 받는다(기획재정부, 2016b).

[그림 II-1]과 [그림 II-2]는 『2012년도 공공기관 경영평가 편람』에 나온 공기업과 준정부기관의 지표 및 가중치 기준이다.

[그림 II-1] 공기업·준정부기관(중소형 제외)의 지표 및 가중치 기준

평가 범주	평가지표	공기업		준정부기관							
				위 탁 집행형	기금관리형						
					연금형		정 책 금용형		사 업 기금형		
		비 계 량	계 량	비 계 량	계 량	비 계 량	계 량	비 계 량	계 량	비 계 량	계 량
리더십· 책 임 경 영	1. 리더십	5		5		5		5		5	
	2. 책임경영	3		3		3		3		3	
	3. 국민평가		5		8		8		8		8
	4. 사회적기여 － 사회공헌 － 정부권장정책	2		2		2		2		2	
			5		5		5		5		5
	소 계	10	10	10	13	10	13	10	13	10	13
경영 효율 ²⁾	1. 업무효율		8		6		6		6		6
	2. 조직 및 인적자원관리	4		4		4		4		4	
	3. 재무예산관리 및 성과 ¹⁾ － 재무예산관리 － 재무예산성과 － 계량관리업무비	4		4		6		5		5	
			6		2		4		2		1
			2		5		4		4		4
	4. 보수 및 성과관리 － 보수 및 성과관리 － 총인건비 인상률	4		4		4		4		4	
			4		4		4		4		4
	5. 노사관리	3		3		3		3		3	
소 계	15	20	15	17	17	18	16	16	16	15	
주요 사업 ²⁾	소 계	15	30	25	20	22	20	25	20	26	20
전체 합계		40	60	50	50	49	51	51	49	52	48

주: 1) 기금관리형의 경우 '재무예산관리 및 성과(재무예산관리, 재무예산성과, 계량관리업무비)'는 '기금운용관리 및 성과(기금운용관리, 기금운용성과, 계량관리업무비)'로 변경됨.

2) 개별기관의 현안과제, 성과지표 도출 한계 등의 특성에 따라 세부지표별 가중치 조정.

자료: 기획재정부, 『2012년도 공공기관 경영평가 편람』, 2011. 12, p.13.

[그림 11-2] 준정부기관 중 중소형 유형의 지표 및 가중치 기준

평가범주	평가지표	중소형	
		비계량	계량
리더십·책임경영	1. 국민평가		8
	2. 정부권장정책		5
	소 계		13
경영효율	1. 업무효율		6
	2. 재무예산성과 ¹⁾		
	- 재무예산성과 ¹⁾		2
	- 계량관리업무비		5
	3. 총인건비 인상률		4
	소 계		17
주요사업 ²⁾	1. 주요사업 성과관리의 적정성	10	
	2. 주요사업 성과		20
	소 계	10	20
전체 합계		10	50

주: 1) 중소형으로 구분된 기금관리형 준정부기관의 경우 '재무예산성과'는 '기금운용 관리 및 성과'로 변경됨.

2) 개별기관의 성과지표 도출의 한계 등의 특성을 고려하여 계량, 비계량지표 가중치 조정.

자료: 기획재정부, 『2012년도 공공기관 경영평가 편람』, 2011. 12, p.14.

2. 선행연구 정리

공공기관 경평제도에 대한 연구는 경평제도의 의의와 필요성, 평가지표의 적절성과 평가결과에 영향을 주는 요인에 대한 분석, 제도 개선과제의 제시에 이르기까지 다각적인 측면에서 이루어져 왔다.

한국조세재정연구원(2010)과 장지인 외(2013)는 공공기관 경평제도의 도입 역사와 제도 변천 과정에 대한 연구를 진행하였다. 이원희·라영재(2015)는 경평제도의 설계 및 운영과정에 참여한 주무부처 담당 공무원과 경평단 단장과 간사들의 회고를 통해 제도의 도입과 설계 배경 및 목적 등에 대한 다양한 정보를 제공해주고 있다.

경평제도의 설계에 있어 이석원 (2005)은 단일한 평가기준을 만들어 이를 모든 공공기관들에게 획일적으로 적용할 경우 정작 그 기관의 고유적 특징이 평가지표에 제대로 반영되지 않는 문제점이 있다는 점을 지적한다. 이창길·최성락(2009)은 명목상 절대 평가로 이루어지는 비계량 평가가 실질적으로 상대평가로 이루어지고 있다고 말한다. 김완희 (2010)는 조직 내적 역량과 무관한 정책환경의 변화라는 외생적 사유가 전체 평가결과를 바꿀 수 있을 정도의 수준으로 남용되는 것은 문제가 있다고 지적하였다. 특히 이로 인해 경영결과가 해당 공공기관의 재무적 건전성에 대한 정확한 정보를 담지 못하게 되고 있다는 점을 지적하고 있다.

제도 개선과 관련하여 박미정 (2010)은 유형별 평가지표 및 가중치를 상이하게 설정하는 것이 바람직하며 평가결과의 서열화 역시 배제하는 것이 바람직하다고 말한다. 또한 주관적 평가에 있어 기관 이미지의 영향을 받을 수 있는 후광효과에 대한 우려를 지적한다. 이오·유승현(2010)은 평가를 받는 공공기관이 느끼는 과중한 평가부담을 문제점으로 지적하며 평가 우수기관에 대해서 소수 핵심성과에 대한 계량지표만 점검해야한다고 제안하였다.

윤태범·라영재 (2013)는 공공기관 직원들을 대상으로 진행한 설문조사의 결과 분석을 통해 경영평가 준비에 투입되는 조직의 인적·물적 자원의 배치가 적정수준을 넘어선 것은 아닌지 의문을 제기하였다. 이를 통해 평가부담을 덜어줄 여러 개선안을 제시하고 있다.

Ⅲ. 경영평가제도에 대한 이론적 고찰

1. 경영평가제도 경제학적 근거

가. 최고경영자의 스톡옵션계약과 경평제도 비교

경평제도를 경제학적인 관점에서 분석하기 위해 가장 먼저 진행해야 할 작업은 경평제도의 필요성을 정당화해주는 논거가 무엇인가라는 질문을 던지는 것이다. 여기서 한 가지 주의해야 할 것이 있다. 이러한 질문을 던지는 의도는 공공기관의 성과측정을 위한 시도 자체가 불필요하다고 주장하는 것은 아니라는 점이다. 오히려 이러한 질문의 의도는 공공기관의 성과측정의 필요성을 인정하는 것을 전제로, 다만 현행 경평제도와 같은 특정한 방식의 장단점을 따져보자는 데 있다.

앞에서 설명한 바 있듯이 여기서 말하는 ‘특정한 방식’은 구체적으로 민간기업의 성과 평가에서 찾아볼 수 없는 다음의 세 가지 요소를 의미한다.

첫 번째, 공공기관 평가의 경우 많은 사람들이 효율성 외에 공익성이란 별도의 기준에 따라 그 성과를 측정하고 평가하려 한다는 것이다. 미국 연방정부와 민간기업의 조직적 차이를 연구한 James Wilson(1989)의 지적처럼 조직의 수입과제(mission)와 설립 근거의 모호함(fuzziness)이 공공기관과 사기업이 구별되는 중요한 특징이다.

두 번째, 효율성과 공익성이란 복수의 과제를 달성하기 위한 공공기관의 노력은 ‘결과(outcome)’라는 단 한 개의 지표에 따라 평가되지 않는다. 실제 공공기관 경평제도의 프로세스는 먼저 조직의 목적을 달성하기 위한 공공기관의 노력을 ‘계획-수행-결과’의 3단계로 분리하는 것으로 시작하다. 다음으로 각각의 단계에 대응하는 평가범주(리더십·경영시스템·경영성과)를 설정한 뒤 각각의 범주에 다시 몇 개의 하위 평가지표를 만든다. 그리고 그

평가지표마다 자의적으로 부여된 가중치가 있다. 최종적으로 이런 평가지표의 점수를 합산하여 기관의 평가 점수(와 순위)가 나온다.

뒤에 가서 좀 더 자세히 살펴보겠지만 이런 평가프로세스에는 민간기업에서는 좀처럼 찾아보기 힘든 요소들이 많다. 예를 들어 국민평가나 사회공헌, 그리고 업무효율처럼 본질적으로 계량화하여 측정하기 쉽지 않아 평가의 객관성 논란을 일으킬 수밖에 없는 항목들이 다수 평가지표에 포함되어 있다.

또한 정평은 기관의 성과를 평가함에 있어 노력의 결과물뿐 아니라 계획과 집행 단계의 적절성을 독자적으로 평가하고 있다. 이로부터 제기되는 의문은 만약 기관이 어떤 사업을 진행할 목적으로 수립한 계획과 그 계획의 이행과정이 적절했다고 평가할 수 있다면 과연 그 기준은 무엇인가라는 점이다. 이는 다시 말하면 우리가 계획과 집행방법의 올바름을 평가함에 있어 이것이 좋은 결과를 가져 오는데 얼마나 기여했는가라는 점이 아닌 다른 독자적인 평가기준이라는 것이 과연 존재할 수 있는냐하는 의문이다.

만약 우리가 이미 어떤 기관의 사업 결과에 대해 평가하고 그 정보를 갖고 있다고 하자. 그리고 그 결과가 기관의 노력의 수준(혹은 효율성이나 공공성의 성취정도)에 대한 정확한 정보를 제공하고 있다고 하자. 이러한 경우 평가자 입장에서 기관의 성과를 평가하기 위해 추가적으로 계획과 집행에 대한 정보를 사용할 이유는 없다. 왜냐하면 이것이 평가자가 원래 추론하려고 했던 기관의 노력 수준에 대해 새로이 제공해주는 정보의 양이 아주 작기 때문이다(즉 이들 정보의 한계가치가 매우 미미하다는 뜻이다). 경우에 따라서는 이들 정보를 사용하는 것이 오히려 잘못된 결론을 도출하는 결과를 낳을 수도 있다. 만약 최종 결과가 좋지 않았는데 계획이나 집행 과정이 적절했다는 평가 결과가 나왔다고 하자. 이는 계획이나 집행 과정을 평가한 지표가 잘못되어 발생한 결과일 수도 있다.

특히 사업의 계획수립이나 집행과정의 적절성을 평가하는 것은 최종 결과에 대한 평가에 비해 훨씬 더 어려운 작업일 수밖에 없다. 무엇보다도 적절한 객관적 지표를 찾기 쉽지 않기 때문이다. 결국 계획이나 집행과정에 대한 평가는 정성적 판단에 의존하는 경우가 많은데 이로 인해 평가가 자의적

으로 이루어질 수 있다. 설사 평가의 자의성 때문이 아니더라도 측정오차(measurement error)로 인한 문제는 항상 존재한다. 따라서 계획이나 집행 과정에 대한 평가 결과는 최종 성과에 대한 평가에 비해 목표 달성을 위한 기관의 노력에 대한 정확한 정보를 담고 있다고 보기 어렵다.

문제는 이처럼 기관의 노력 수준을 제대로 반영하지 못하는 부정확한 지표에 근거하여 구성원의 보상체계를 설계할 경우 왜곡과 비효율성이 발생한다는 것이다. 바로 이러한 이유로 민간기업의 경우 임원들의 보수는 일반적으로 주가에 연동된다. 공공기관처럼 보수가 사업 계획이나 집행 과정의 평가결과에 의존하는 경우는 거의 찾아볼 수 없다.

마지막으로 민간기업의 성과 평가제도와 비교하여 공공기관 경평제도가 갖고 있는 독특한 요소는 모든 평가 지표들은 공공기관이 과제를 수행하기 전에 사전적으로 공표되고 일단 공표되고 나면 좀처럼 수정되지 않는다는 점이다. 효율성의 관점에서 이러한 평가방식은 다음과 같은 문제점을 갖고 있다.

먼저 사전적으로 제안된 지표에 따라서만 기관의 성과가 평가되기 때문에 기관으로 하여금 우발적 상황에 대처하도록 하는 유인을 크게 떨어뜨린다. 각 기관의 성과는 기본적으로 그 기관이 특정 과제를 위해 투입하는 노력의 정도에 의해 결정되지만 동시에 사전에 예측할 수 없는 불확실성과 우발적 요인(예컨대 외환위기나 국제유가의 하락과 같은 외부적 요인, 혹은 CEO의 갑작스러운 사망이나 생산설비의 화재와 같은 내부적 요인)에 의해서도 영향을 받는다. 경우에 따라서는 기관 사이의 성과의 차이가 이러한 우발적 상황에 대해 얼마나 유연하고 담대하게 대응했는가에 의해 결정되기도 한다.

문제는 현행 경평제도는 그 평가지표를 사전에 정해 놓음으로써 기관의 성과에 결정적 영향을 미칠 수 있는 우발 요인에 대한 적극적 대응 유인을 현저히 약화시킨다는 점이다. 설사 기관의 입장에서 우발 요인에 대한 대응이 기관의 설립목적에 부합하는 행동이라는 것을 안다고 하자. 그러나 그것이 평가지표에 빠져 있는 한 이에 대해 적극적으로 대처할 유인은 없다. 서론에서 거론한 인천공항공사의 딜레마가 그 대표적인 사례이다. 물론 이러한 주장에 대해서 지표 설정 후 발생한 우발적 요인에 대한 고려는 평가과

정에서 지표를 수정하는 방식으로 이루어지기 때문에 큰 문제가 없다고 항변할 수 있다.

그러나 기관이 위험기피적(risk averse)이라면 평가지표의 개선 요구가 받아들여지기 전에는 이를 위한 행동에 나서지 않을 것이다. 앞서의 인천공항 공사의 사례를 통해 설명해보자. 만약 보안에 대한 대규모 투자로 인해 공사의 부채비율이 상승하고 이로 인해 수익성이 떨어져 단기적으로 재무성치가 악화된다고 가정해보자. 따라서 기관장 입장에서 보면 보안에 대한 투자의 불가피성을 경영평가단이 인정하여 예컨대 보안과 안전에 대한 투자비용은 비용으로 간주하지 않겠다는 식으로 지표를 수정하기 전까지는 보안에 대한 투자를 진행하지 않을 것이다.

그렇다면 공공기관들의 평가지표 개선 요청은 얼마나 받아들여지고 있는 것일까? 실제 데이터에 따르면 2015년의 경우 기관들의 평가지표 수정 요청이 받아들여져 실제 지표의 개정으로 이루어지는 비중은 평균적으로 30% 수준에 불과했다.

〈표 III-1〉 2015년 평가지표 개선 요청 및 반영 결과

(단위: 건, %)

구분	공기업1	공기업2	기금관리형	위탁집행형	강소형	합계
요청 건수	47	66	81	63	145	402
1차 검토	26	22	17	21	43	129
2차 검토	15	28	20	15	44	122
최종	15	32	14	22	36	119
최종 비율	31.9	48.5	17.3	34.9	24.8	29.6

출처: 2016년 공공기관 경영평가단 내부 자료.

결론적으로 현행 공공기관에 대한 경평제도는 대리인으로 하여금 사회적으로 요구되는 수준의 노력을 이끌어내는 유인계약의 본래 취지에 부합하도록 정교하게 설계되어 있다고 보기 어렵다. 직관적으로 말해서 민간의 어떤 기업도 경평과 같은 방식으로 인센티브 계약을 체결하지 않고 있다는 사실이야말로 이 제도가 안고 있는 설계상의 문제를 단적으로 드러내주는 가장

강력한 증거인 셈이다. 만약 누군가가 현행 경평제도와 동일한 내용의 유인 계약을 설계하고 이를 민간기업에 적용하려 한다고 하자.

이 경우 우리는 시장에서 다음과 같은 내용의 연봉계약을 찾을 수 있어야 한다. 3월 주주총회에서 주주들이 기업의 ‘계획-집행-성과’라는 세부 항목에 대해 각각의 평가지표와 가중치를 정하고 오로지 이 기준에 따라 경영진을 평가한 뒤 성과급을 지급하겠다고 공표하는 것이다. 모두가 잘 알고 있 다시피 이러한 방식으로 계약을 체결하여 경영진의 연봉을 결정하고 있는 민간기업은 단 하나도 없다.

나. 약관과 경평제도 비교

앞에서 우리는 최고 경영진 시장에서 공공기관 경평제도와 유사한 방식으로 연봉계약을 체결하는 경우를 찾아보는 것은 거의 불가능하다고 주장했다. 그러나 경평과 같은 구조를 특징으로 하는 계약이 시장에 전혀 존재하지 않는 것은 아니다. 실제 계약체결 이전에 당사자가 해야 할 일을 세세하게 유형별로 규정하고 이의 준수 여부를 감시하는 형태의 계약은 보험시장에서 보험회사가 피보험자에게 제공하는 약관의 형태로 존재한다.

예컨대 생명보험회사들은 피보험자가 보험 구입 후 일정 기간 이내에 자살을 하면 보험금 지급을 거절하는 조항을 약관에 포함시킨다. 화재보험에도 유사한 조항이 있다. 건물에 화재 감지기 및 화재 진압 장비들을 설치할 경우 보험료를 할인해주는 것이 그 대표적인 사례이다(굴스비·레빗·시버슨, 2014, p. 670).

이러한 행태의 계약이 체결되는 이유는 도덕적 해이 문제 때문이다. 보험회사가 보험계약을 체결한 후 계약자의 행동을 관찰할 수 없는, 이른바 ‘정보 비대칭성’의 문제를 완화시키기 위해 보험회사는 피보험자의 전형적인 도덕적 해이 현상들을 유형화하여 이를 약관에 담는다. 이처럼 약관은 보험회사가 피보험자에게 보험의 적용을 받기 위한 전제 조건으로 도덕적 해이의 소지가 있는 행동들을 사전에 유형별로 규정한 뒤 그 이행 여부를 관찰하고 그 결과에 따라 피보험자의 의무이행 여부를 평가하여 이들에게 각기

상당한 보험금을 지급한다. 바로 이러한 점에서 약관은 현재의 경평 제도와 구조적으로 유사하다. 위의 보험계약의 사례로부터 경평제도의 도입이 사회 후생(social welfare)의 관점에서 보다 바람직한 결과를 가져올 수 있는 상황에 대한 시사점을 얻을 수 있다. 그것은 바로 경평제도가 공공기관의 도덕적 해이의 통제장치로 기능할 때 그 의미를 갖는다는 것이다. 이를 좀 더 자세히 설명해보자.

정부와 공공기관 사이에는 보험회사와 보험계약자처럼 정보의 비대칭성이 존재한다. 관료들은 해당 공공기관이 자신의 설립 목적과 법에서 정한 수입 과제를 이행하기 위해 필요한 노력을 제대로 기울이고 있는지를 관리 감독하는 것은 어려운 일이다. 그것은 관료들이 공공기관의 운영진에 비해 기관의 업무를 평가하는 데 필요한 고도의 전문적 지식이 부족해서 일 수도 있고 공공기관 스스로가 관료에게 정확한 정보를 제공하지 않기 때문일 수도 있다. 어떠한 이유이든 간에 이로부터 도덕적 해이의 문제가 발생한다.

또한 이러한 일반적인 정보비대칭성의 문제 외에 공공기관의 경우 이러한 도덕적 해이의 문제가 보다 악화될 수 있는 추가적 요인이 존재한다. 그것은 공공기관의 경우 민간기업과 달리 시장으로부터 퇴출될 수 있는 가능성, 즉 파산의 위험이 사실상 전무하다는 것이다. 이는 시장이 갖고 있는 공공기관의 부채에 대한 정부 보증에 대한 강한 기대 때문이다(최한수·이창민(2015)). 즉 공공기관의 여러 이해관계자들(채권자, 노동자, 그리고 공공기관의 기관장)은 공공기관은 정부의(암묵적) 보증 때문에 채무불이행과 같은 사건이 발생했을 때 그 충격으로부터 완벽하게 보호받을 것이라고 기대하고 있다. 한마디로 이들은 정부가 자신들에게 완전보험(full insurance)을 제공해주고 있다는 믿음을 갖고 있는 것이다.

위험에 대비할 수 있는 보험을 들고 있다는 생각은 피보험자로 하여금 위험선호적 경향을 부추긴다. 최한수·이창민(2015)은 공공기관의 과다 부채 문제가 바로 공공기관의 부채에 대한 정부의 암묵적 보증의 기대에서 비롯된 도덕적 해이의 결과라는 점을 보이고 있다. 과다 부채 외에 공공기관의 여러 비효율적 경영행태는 도덕적 해이에서 비롯된 것이 많다. 따라서 만약

경평의 주요 평가항목과 그 세부 지표가 공공기관에서 발생하는 전형적인 도덕적 해이의 징후들과 관련이 있다면 경평의 필요성은 보다 쉽게 인정될 수 있다.

다. 공공기관의 설립목적의 다차원성

일반적으로 공공기관은 ‘이윤극대화’를 목적으로 하는 사기업과 달리 ‘공공성’이라는 또 다른 목적을 갖고 있다. 이러한 이유로 공공기관의 성과는 사기업처럼 이윤의 크기나 주가와 같은 하나의 객관적 지표로 측정될 수 없다. 즉 특정 공공기관이 공공성의 추구라는 설립목적을 달성하기 위해 어느 정도 노력을 기울이고 있는지를 판단하려면 비계량지표의 사용은 불가피하다.

비단 공공성뿐만 아니라 효율성을 측정할 경우에도 사기업과 다른 공공기관의 특수성이 고려될 수밖에 없다.

공공기관의 경우 사기업과 달리 인사나 예산에 있어 자율성이나 재량권의 정도가 낮다. 또한 공공기관이 생산하는 재화나 서비스의 경우 그 가격 결정이 시장에서 공급과 수요의 원리에 의해 자유롭게 결정되기보다는 공익적 목적을 위해 통제받는 경우도 허다하다. 따라서 사기업처럼 그 결과지표만을 갖고 노력의 정도를 평가하는 것은 부당하며 오히려 미션 및 비전의 수립·전략 기획·사업 수행의 전 과정을 종합적으로 평가할 필요가 있다. 이러한 경영의 전 과정에 대한 평가는 실제 평가하고자 하는 항목의 내용과 다소 차이가 있다 하더라도 사용가능한 객관적(예컨대 리더십 및 책임경영 항목을 평가함에 있어 고객 만족도 조사를 사용하는 경우) 혹은 주관적 지표(리더십에 대한 평가단의 주관적 평가)로 측정된다.

이러한 주장은 일면 타당성을 갖고 있다. 그러나 이러한 주장이 보다 설득력을 가지려면 다음의 두 가지 이슈들을 추가적으로 검토해야 한다. 먼저 현행 경평지표들이 공공기관의 리더십이나 책임경영을 높이기 위한 기관의 노력들에 대한 정보를 얼마나 정확하게 전달하고 있는가라는 점이다. 또 다른 이슈로는 리더십과 책임경영 범주처럼 그 측정 결과의 신뢰성이 높지 않

은 항목이 실제 경영결과와 성과급의 결정에 어느 정도 영향을 끼치는가의 문제이다. 앞으로 본 장의 제2절과 제3절에서는 이러한 이슈들이 갖고 있는 현실적 함의에 대해 간단한 이론적 모형을 이용하여 좀 더 자세한 설명을 진행할 것이다.

2. 계약(contract)으로서의 경영평가제도

계약이론(contract theory)에 따르면 유인계약은 한 쪽 당사자가 다른 쪽의 행동을 관찰할 수 없는 상황에서 발생하는 도덕적 해이의 문제를 해결하기 위해 필요하다. 구체적으로 유인계약이 필요한 조건은 다음과 같이 설명될 수 있다.

(1) 주인(principal; 경우에 따라 주주 혹은 국가)과 대리인(agent; 경우에 따라 경영진 혹은 공공기관)이라는 두 경제주체가 존재하고 이들 사이에 이해관계가 완벽하게 일치하지 않으며 (2) 위험 성향(risk attitude)에 있어 대리인은 위험기피적(risk averse)이고 주인은 위험중립적(risk neutral)으로 두 주체의 위험 성향이 다르며 (3) 주인이 대리인의 행동(혹은 노력 수준)을 완벽하게 관찰할 수 없는 경우 주인이 대리인의 노력 수준을 높이기 위해서는 유인계약이 필요하다.

특히 대리인의 행동(혹은 노력 수준)과 행동의 결과 사이에 1대 1의 대응 관계가 없어야 한다. 즉 결과가 대리인의 노력 외의 요인에 의해서도 영향을 받기 때문에 주인은 발생한 결과를 관찰한 뒤 과연 대리인이 어느 정도 수준의 노력을 투입하였는지를 추론(inference)해야 한다는 것이다.

이 경우 주인은 대리인의 보상체계를 다음의 두 가지 방식으로 설계할 수 있다. 우선 관찰되는 결과와 상관없이 대리인에게 고정급을 지급하는 것이 가능하다. 이 경우 주인은 왜 결과가 좋지 않았는지 그 이유를 따지지 않는다. 만족스럽지 않은 결과는 대리인이 자신의 임무를 해태한 결과일 수도 있고 반대로 최선을 다했지만(예컨대 과거 IMF와 같은 외환위기의 사례처럼) 운이 좋지 않아 발생한 결과일 수도 있다. 고정급을 지급한다는 것의

의미는 이와 같은 계약과 관련된 모든 위험을 주인이 모두 떠안는다는 것을 의미한다. 예컨대 회사의 경우 영업이익을 내지 못한 경우에도 노동자에게 여전히 약속한 급여를 지급하는 경우가 이에 해당한다.

문제는 이처럼 결과의 실현과 관련된 모든 위험이 주인에게 이전될 경우 대리인은 나쁜 결과를 피하기 위한(혹은 바람직한 결과를 만들기 위한) 노력을 사회적으로 필요한 수준보다 적게 기울이게 된다. 이러한 문제를 해결하는 방법 중 하나가 대리인이 받는 보상을 성과에 연동시키는 것이다. 즉 주인은 대리인과 ‘성과 급여’와 같은 유인계약을 체결함으로써 대리인의 노력 수준을 끌어올릴 수 있는 것이다.

그러나 성과 급여를 제대로 설계하는 작업은 쉬운 일은 아니다. 보상체계를 설계함에 있어 고려해야 할 중요한 요소 중 하나가 대리인의 급여를 연동시키는 기준이 되는 성과를 제대로 측정하는 일이다. 만약 객관적이고 정교하지 못한 지표를 사용하여 대리인의 성과를 측정할 경우 어떤 문제가 발생할까?

Bolton and Dewatripont(2005, pp.137~139)는 비교적 일반적인 가정과 특별한 모형 구성²⁾하에서 대리인이 선택하는 노력 수준은 이를 측정하는 지표의 정확성이 떨어질수록 낮아진다는 점을 보이고 있다.

이를 간단한 수식을 통해 설명하면 다음과 같다. 대리인이 받는 보수(w)가 특정 성과지표(q)에 연동된 성과급의 특성을 갖고 있다. 이 경우 보수의 또 다른 구성변수 s 는 보수가 성과지표에 의해 어느 정도 영향을 받는지를 보여주는 파라미터(parameter) 값이다. 만약 s 가 0이라면 보수(w)는 성과와 무관하게 결정되어 t 라는 금액만을 받는다. 즉 이 경우 보수는 고정급이 되는 것이다. 반면에 s 가 0이 아닌 값을 갖는 경우 그 크기가 커질수록 보수가 성과에 굉장히 민감하게 반응하는 형태를 갖게 된다.

성과(q)는 대리인의 노력(a)과 오차항(ε)으로 구성되어 있다고 하자. 여기서 오차항은 두 가지로 해석될 수 있다. 첫 번째는 이를 계약의 두 주체인

2) 가정은 먼저 보상계약이 성과에 선형(linear)으로 설계되어 있어야 하고 노이즈가 정규분포(normally distributed)되어야 하고, 효용함수가 CARA(Constant Absolute Risk Aversion)이어야 한다.

주인과 대리인 모두가 통제할 수 없는 외부적 요인(예컨대 외환위기와 같은 외부 충격)이라고 해석하는 것이다. 또 다른 해석은 이를 노력의 수준을 측정하는 지표의 정확성(즉 측정오차)의 정도를 보여주는 것으로 이해하는 것이다. 즉 ε 가 크다는 것은 성과를 측정하는 지표가 대리인의 실제 투입 노력 수준에 대한 부정확한 정보를 제공하는 것을 의미한다.

대리인의 효용함수(U_A)는 다음과 같이 정의된다: $U_A(w, a) = -e^{n[w-c(a)]}$. 즉 대리인이 느끼는 효용은 그가 보수를 많이 받을수록 그리고 자신의 업무를 하는데 드는 노력 수준이 낮을수록 커진다. 주인의 보수함수(U_p)는 $U_p = q \cdot w$ 로 정의된다.

지금까지의 논의를 수학적으로 표현하면 다음과 같다.

$$\begin{aligned}
 U_p &= q \cdot w \\
 U_A(w, a) &= -e^{n[w-c(a)]} \\
 w &= sq + t \quad (t \geq 0) \\
 q &= a + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2) \\
 c(a) &= a^2
 \end{aligned}
 \tag{식 III-1}$$

이러한 조건에서 최적화 문제를 풀면 대리인의 최적 노력 수준(a^*)은 $a^* = a(\sigma^2)$ 처럼 σ^2 의 함수로 표현될 수 있다. 즉 대리인의 최적 노력 수준은 성과를 측정하는 지표의 정확성에 의존하게 된다. 다시 비교정학(comparative statics)을 통해 지표의 정확성과 최적 노력 수준과의 관계를 살펴보면 $\frac{da}{d\sigma^2} < 0$ 이라는 결과를 얻을 수 있다. 즉 σ^2 가 커질수록 대리인의 최적 노력 수준 a^* 의 값은 작아진다. 아울러 평가지표의 정확성과 보수의 성과민감도에 대한 관계도 확인할 수 있는데 $\frac{ds}{d\sigma^2} < 0$ 이란 결론을 도출할 수 있다. 즉 σ^2 가 커질수록 보수의 성과민감도 또한 감소해야 하는 것이다.

앞서 σ^2 는 외부에서 관찰하기 쉽지 않은 대리인의 노력의 수준을 관찰가능한 척도로 변환하는 과정에서 발생하는 측정오차라고 정의하였다. 만약 σ^2 가 무한대(infinity)라고 해보자. σ^2 가 무한대라는 것은 대리인의 노력 수

준을 측정하기 위해 우리가 관찰하는 지표가 실제로는 대리인이 그 과제를 달성하기 위해 얼마나 노력을 투입하는지에 대해 어떠한 가치 있는 정보도 제공해주지 못한다는 것을 의미한다.

한편 이는 다음과 같이 설명할 수도 있다. 만약 성과지표가 대리인의 노력으로 통제할 수 없는 외부적 요인에 의해 사실상 결정된다고 하자. 여기서 ‘통제할 수 없는 외부적 요인’은 앞서 언급한 IMF 외환 위기와 같은 거시경제적 요인인 경우가 많겠지만 반드시 여기에 한정될 이유는 없다. 예컨대 공공기관의 경우 정부의 개입도 통제할 수 없는 외부 요인의 하나이다.

이러한 조건에서 대리인의 노력 수준은 어떻게 결정되며 주인은 대리인의 보수의 성과민감도를 어느 수준에서 설정해야 할까? 이 두 경우 결론은 동일하다. 앞서 비교정확의 과정을 통해 우리가 확인할 수 있었던 사실은 만약 대리인의 성과를 측정하기 위해 사용하는 지표가 실제 대리인이 기울인 노력 수준에 대한 정확한 정보를 제공해주지 못한다면 대리인은 늘 사회적으로 바람직한 수준보다 적게 노력한다. 즉 대리인의 노력 수준은 지표의 정보로서의 가치(informativeness)가 하락하는 것에 따라 낮아진다. 따라서 σ^2 가 무한대인 경우처럼 극단적으로 성과지표가 대리인의 노력 수준을 보여주는 정보로서의 가치가 전무하다면 대리인은 아무런 노력을 기울이지 않게 된다. 노력이 성과에 전혀 영향을 못 미치는 상황에서는 대리인은 노력에 따른 비용을 0으로 만듦으로써 자신의 효용 수준을 높이려 하기 때문이다.

더 나아가 주인 또한 σ^2 가 무한대인 경우에서 대리인이 아무런 노력을 기울이지 않는다는 점을 알고 있기 때문에 자신이 관찰하는 지표(q)를 통해 대리인의 노력 수준을 추론하려 하지 않을 것이다.

이는 주인의 입장에서는 대리인의 노력 수준에 대한 정확한 정보를 제공하는 지표, 즉 다른 말로 측정오차가 적은 지표(q)를 찾는 것이 유인계약을 통해 대리인의 노력 수준을 끌어올리려 하는데 대단히 중요한 요인이라는 것을 의미한다. 아울러 이러한 노력의 수준을 가늠하게 해주는 객관적 지표를 찾는 것이야 말로 성과보수체계를 도입하고 운영하기 위한 전제조건이라는 점이다.

지금까지의 논의를 경평의 맥락에 맞게 재정리하면 다음과 같다. 먼저 현행 경평제도의 설계방식과 구성요소는 CEO의 성과보수계약보다 화재보험약관에 가깝다. 이는 정책당국이 경평제도를 공공기관의 도덕적 해이의 정도를 낮추어주는 정책 수단으로 사용하는 것을 정당해준다. 이러한 판단으로부터 우리는 본 연구에서 경평제도의 효과를 평가할 수 있는 하나의 명제를 도출할 수 있게 된다. 그것은 경평제도를 도입함으로써 해당 기관의 도덕적 해이의 징후들을 보여주는 지표의 개선이 이루어져야 한다는 것이다. 만약 그렇지 않았다면 이는 경평제도가 제대로 운영되고 있지 않다는 것을 의미한다.

이제 경평제도를 도덕적 해이의 통제장치로서가 아니라 CEO의 스톡옵션과 같은 일종의 성과보수시스템이라고 이해해보자. CEO의 스톡옵션의 취지처럼 이는 경평제도가 대리인의 노력 수준을 끌어올리게 하는 유인계약의 기능을 해야 한다는 것을 의미한다. 이의 가장 중요한 전제조건은 공공기관의 성과를 측정하는 지표들이 공공기관의 노력 수준에 대한 가능한 한 정확한 정보를 제공해주어야 한다. 만약 경평결과가 효율성이나 공공성을 높이기 위한 기관의 노력과 무관한 요인에 의해 결정된다면 경평의 유인계약으로서의 기능은 오히려 사회적으로 바람직하지 않은 결과를 가져올 수 있다. 또한 이러한 경우 보수의 성과민감도를 낮추어야 한다.

3. 복수의 과제를 대리인에게 맡기는 경우의 최적 계약이론

앞 절에서 살펴본 유인계약에 대한 이론적 고찰은 조직이 단일한 목표를 부여받아 이를 성취하려고 하는 경우의 문제를 다룬 것이다. 그러나 공공기관의 경우 기관의 목적 자체가 복잡하고 다차원적이다. 이들 기관의 경우 정부, 국민, 고객 등 다양한 이해관계자로부터 효율성과 공익성이란 목적을 이루기 위해 “균형 있게” 노력해줄 것을 요구받고 있다(라영재·윤태범, 2013).

본 절에서 소개하려는 Holmström and Milgrom(1991) 모델은 이처럼 복수의 과제를 한 명의 대리인에게 맡길 경우 발생하는 도덕적 해이(Moral hazard with multi-tasks)와 최적 계약의 문제를 다루고 있다. 이 모델의 기

본적 문제의식은 다음과 같다. 한 명의 대리인이 복수의 과제(혹은 복합적 성격을 갖는 하나의 과제)를 수행하도록 요구받았을 때 성과에 따른 보상 체계를 채택하는 것이 바람직한가의 문제이다.

주인이 대리인에게 복수의 과제를 동시에 진행할 것을 요구하였을 때 대리인이 직면하는 어려움은 더 이상 일반적 계약에서의 위험부담(risk sharing)과 유인 사이의 상충(trade-off)이라는 문제로 단순화되지 않는다.

예컨대 과제 A와 B를 동시에 수행해야 하는 대리인이 있다고 하자. 대리인의 보수는 과제 A의 결과에 달려 있다고 하자. 즉 과제 B의 결과는 대리인의 보수 크기에 영향을 주지 않는다고 하자. 이 경우 대리인이 직면하는 문제는 더 이상 앞서 사례에서 살펴본 것 처럼 두 주체 모두가 통제할 수 없는 위험에 따른 손실을 주인과 대리인 사이에 누가 부담할 것인가의 문제로 귀결되지 않는다. 복수의 과제에 직면한 대리인의 문제는 대리인 스스로가 자신의 보수의 수준에 영향을 미치는 과제 A와 이에 영향을 주지 못하는 과제 B 사이에 각각 어느 정도 노력과 시간을 할애해야 하는가라는 과제 간 노력 배분(allocation of effort among several tasks)의 문제로 바뀌게 된다 (Bolton and Dewatripont, 2005, pp. 217-219).

좀 더 현실적 사례를 갖고 이 문제를 설명해보자. 정부는 학생들이 교육을 통해 수리적 능력과 창의성이란 두 가지 덕목을 얻게 되기를 원한다. 그리고 교사들이 이러한 과제를 보다 효과적으로 수행할 수 있도록 성과급을 지급하기로 결정하였다. 학생들의 수리적 능력의 정도는 전국 단위 모의고사의 수학점수로 평가하기로 하였다. 반면에 창의성의 경우 이에 대한 객관적인 측정이 어려워 성과급 지급 기준에서 제외하였다. 이처럼 교사의 성과급을 오로지 자신이 가르치는 학생의 수학성적의 결과에 연동되도록 설계할 경우 교사들이 학생들의 창의성을 길러주기 위해 어느 정도 자신의 시간과 노력을 할애할 것인가?

Holmström and Milgrom(1991)은 이 질문을 다음과 같은 이론적 모델로 정식화하고 있다. 편의상 이 모델도 앞 절의 주인-대리인 모형과 동일한 가정과 모델 구성(model specifications)에 근거하여 설명해보자. 다른 것은 과

제(i)가 복수($i \neq 1, 2$)라는 것과 대리인의 보수 구조(w)가 과제 1과 과제 2의 결과(q_i)에 의존한다는 것이다. 다시 개별 과제의 성과는 각각의 과제에 대리인이 얼마만큼의 노력($a_i, i=1, 2$)을 기울이는지 여부와 측정오차(ε_i)에 의존한다.

이 경우 보수구조의 설계에 있어 중요한 역할을 하는 변수는 s_i 이다. s_i 는 대리인의 보수(w)가 어느 정도 각 과제의 평가지표인(q_i)에 의존해야 하는가를 나타내는 것으로 보수의 성과 민감성(sensitivity)를 포착하는 변수라고 보면 된다. 예를 들어 s_i 가 0이 경우 대리인의 보수는 과제 1의 성과에 전혀 영향을 받지 않는 것이다. 만약 s_i 가 1이라면 대리인의 보수는 과제 1의 성과에 전적으로 의존한다고 보면 된다. 만약 s_1 과 s_2 모두가 0이라면 대리인의 보수는 전혀 성과에 의존하지 않는 고정급 구조를 갖고 있는 것으로 이해하면 된다. 지금까지의 설명을 수식을 통해 설명하면 다음과 같다.

$$\begin{aligned}
 w &= t + s_1 q_1 + s_2 q_2 \\
 q_i &= a_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma_i^2), \quad i=1, 2 \quad (\text{식 III-2}) \\
 c(a_1, a_2) &= (c_1 a_1^2 + c_2 a_2^2) + 2\delta a_1 a_2
 \end{aligned}$$

여기서 중요한 것은 과제 1과 과제 2의 관계이다. 먼저 과제 1과 과제 2 사이에는 어떠한 외부성(externality)도 없다는 것에 주목할 필요가 있다. 즉 대리인이 과제 1에 쏟은 노력 a_1 은 q_1 에만 영향을 줄 뿐 q_2 에는 아무런 영향을 미치지 못한다(이는 반대의 경우도 마찬가지이다). 또한 과제 1을 하는 노력과 과제 2를 수행하는 노력이 서로 대체관계(substitute)에 있다는 점에 주의할 필요가 있다. 이는 과제 1에 쏟는 노력 수준이 다른 과제에 소요되는 한계비용(marginal cost)의 크기에 영향을 미친다는 것을 의미한다. 이를 잡아내는 모수가 δ 이다. 만약 $\delta > 0$ 이라면 두 과제 사이에는 서로 대체성이 있어 한 과제에 대한 노력 수준의 증가는 다른 작업에 투입되는 노력 수준이 감소된다는 것을 의미한다(김빛마로, 2016).

주인은 대리인이 과제 1과 과제 2 모두에 어느 정도 시간을 할애하여 성과가 나오기를 기대한다. 즉 주인의 효용함수는 과제 1과 과제 2의 성과 모

두에 달려 있다(앞서 교사의 사례에서 교육당국은 학생이 수학도 잘 풀지만 창의적으로 사고하기를 원한다). 그런데 문제는 과제 2(앞서의 교사의 사례에서 '창의성 발달'이란 목표)의 성취 여부를 판단하는 척도가 매우 불완전하다는 것이다. 즉 σ_2^2 가 매우 큰 값을 갖는다. 반면에 과제 1의 경우 이의 성취도를 비교할 객관적 기준이 존재한다. 즉 σ_1^2 은 별로 크지 않다고 가정하자. 이러한 조건에서 Holmström and Milgrom(1991)의 연구는 주인이 대리인으로 하여금 두 과제 모두에 '균형 있게' 노력해주기 위해서는 s_1 과 s_2 를 어떻게 결정해야 하는가에 대한 해법을 제시하고 있다.

이들의 결론은 다음과 같다. 과제 2의 성과를 측정하는 지표(q_2)가 대리인이 과제 2를 수행하기 위해 필요한 노력(a_2)에 대한 척도로써 불완전할수록(즉 σ_2^2 가 커질수록) s_2 는 작아져야 한다. 즉 이 경우 대리인의 보수(w)는 과제 2의 성과지표에 덜 의존하도록 설계되어야 한다. 극단적으로 과제 2의 성과를 측정하는 지표(q_2)가 a_2 에 대해 어떠한 유의미한 정보도 제공하지 못할 경우(즉 σ_2^2 가 무한대로 갈 경우) s_2 는 0이 되어야 한다. 즉 대리인의 보수는 과제 2의 성과지표에 의해 영향받지 않도록 설계되어야 한다.

Holmström and Milgrom(1991)의 통찰은 여기서 한 걸음 더 나아간다. 그렇다면 이 경우 보수는 과제 1의 성과에 연동되도록 해야 하는가? 만약 과제 1을 하는 노력(a_1)과 과제 2를 수행하는 노력(a_2)이 서로 대체관계에 있고 주인이 대리인이 과제 1과 과제 2 모두에 노력을 투입해주기를 원한다면 이 경우 s_1 도 0으로 놓아야 한다는 것이다. 이 경우 대리인의 보수 구조는 어떠한 성과에 대한 보상도 없는 고정급이 되어야 옳다. 그 이유를 앞에서 언급한 교사의 성과급의 사례를 들어 설명해보자.

학생들의 창의성 수준을 측정하는 지표가 대단히 부정확하게 설계되어 이 지표를 보고 교사가 학생들의 창의성을 높이기 위해 노력을 기울였는지를 판단하기가 사실상 불가능하다고 하자. 즉 이 경우 σ_2^2 가 무한대로 갈 경우 s_2 는 0이 되어야 한다. 그런데 이 경우 만약 s_1 이 0이 아니라고 해보자. 그렇다면 교사는 자신의 모든 시간을 이용하여 학생들의 수학점수를 높이는 데 열중한 뿐 학생들의 창의성을 길러주기 위한 노력은 시도조차도 하지 않을

것이다. 따라서 국가가 교사들이 자신에게 주어진 시간을 학생들의 수학 능력과 창의성 제고라는 두 가지 목표에 적절하게 배분하여 주기를 원한다면 교사의 급여를 고정급의 형식으로 지급하는 것이 옳다.

이러한 이론적 통찰은 경쟁제도의 설계에 있어 중요한 시사점을 제공해준다. 현행 경쟁제도는 공공기관으로 하여금 효율성과 공익성이라는 두 가지 목표를 ‘균형감 있게’ 추구할 것을 요구하고 있다. 효율성(이를 ‘과제 1’이라 하자)은 부채비율이나(노동·자본) 생산성과 같은 재무 지표를 사용하여 비교적 객관적으로 측정된다. 공익성(이를 ‘과제 2’라 부르자)은 주로 리더십이나 책임경영이란 항목을 통해 측정된다.³⁾ 현행 경쟁제도는 과제 1과 과제 2라는 각각의 과제에 대한 평가결과에 따라 각기 다른 성과급을 지불하고 있다.

문제는 과제 2의 평가지표(q_2)가 정말로 해당 기관의 공공성에 대한 정확한 척도인가라는 점이다. 만약 과제 2와 관련된 평가지표들이 실제 기관의 공익성 제고를 위한 노력 수준에 대한 정확한 정보를 제공해주지 못한다면 s_2 의 비중이 작아져야 한다. 이는 다시 말하면 과제 2의 결과가 공공기관 사이의 성과급의 차이를 결정짓는 주요한 요인이 되어서는 안된다는 것을 의미한다(이에 대한 논의는 제Ⅳ장 제3절에서 보다 자세히 설명할 것이다).

3) 아니면 효율성은 주로 계량지표에 의해 그 정도를 정성화하기 어려운 공익성은 주로 비계량지표에 의해 측정되고 있다고 놓자.

Ⅳ. 경영평가제도 설계의 적절성과 관련된 명제

앞 장의 서술된 계약이론을 근거로 본 장에서는 현행 경평제도가 적정하게 설계되고 운용되고 있는지를 판단할 것이다. 이를 검증할 수 있는 몇 가지 명제를 제시하고자 한다.

1. 도덕적 해이의 통제장치로서의 경영평가제도

본 연구에서 가장 먼저 검증하려는 이슈는 현행 경평제도가 공공기관의 도덕적 해이에 대한 통제장치로서의 역할과 기능을 다하고 있는가이다. 이러한 작업을 진행하기 위해서는 먼저 공공기관의 도덕적 해이의 정도를 잘 보여주는 지표들을 찾아야 한다.

다음의 이유로 부채비율은 공공기관의 도덕적 해이를 나타내는 척도로 사용되기에 적합하다. 최한수·이창민(2015)에 따르면 공공기관의 도덕적 해이 현상은 근본적으로 정부의(암묵적) 보증에 따라 민간기업과 달리 파산이라는 경영규율 기제가 작동하지 않는다는 데서 기인한다. 이로 인해 공공기관의 위험추구 성향이 높아질 수 있는데 이를 보여주는 가장 적절한 지표가 해당 공공기관의 부채규모이다.

물론 공공기관의 부채 유발 요인은 다양하다. 공공기관의 높은 부채비율은 비효율적 기업경영, 국책사업의 무리한 추진, 혹은 정부의 과도한 요금통제의 산물이다(이에 대한 문헌 정리는 최한수·이창민, 2015, pp. 23~24 참고). 그러나 투자에 필요한 자금을 조달하는 방법이라는 측면에서 보았을 때 공공기관과 민간기업의 가장 큰 차이는 공공기관은 많은 경우 자본이 아닌 부채를 통해 자본을 조달한다는 것이다. 이는 한편으로 제도적 제약 때문이기도 하지만 다른 한편으로는 민간기업에 비해 현저히 낮은 비용으로

자금을 조달할 수 있다는 상황에 기인한 것이다. 그리고 이러한 낮은 자본 조달비용은 공공기관에 대한 정부의(묵시적) 보증의 결과이다. 만약 이것이 없었다면 공공기관이 국책사업으로 추진했던 여러 프로젝트들(4대강 사업이나 자원외교)은 시장규율의 작동에 의해 좌초되거나 그 규모가 대폭 축소되었을 것이다. 따라서 부채비율은 정부의(묵시적) 보증에서 발생하는 도덕적 해이 정도를 보여주는 평가지표가 될 수 있다.

본 연구에서는 부채비율에 대한 정보를 이용하여 경평제도가 공공기관의 도덕적 해이를 억제하는 기제로 제대로 작동하고 있는지 여부를 살펴볼 것이다. 이를 위해 다음의 두 가지 명제를 검증할 것이다.

검증 명제

1. 부채비율이 높은 기관이 경영평가에서 불이익을 받을 것이다.
2. 이번 기에 경영평가 점수가 좋지 않은 공기업의 경우 다음 기 부채비율이 개선될 것이다.

검증 명제 1은 경평결과가 해당 공공기관의 부채비율에 대해 얼마나 유의미한 정보를 제공해주는가를 살펴보기 위한 것이다. 검증 명제 2는 경평제도의 규율적 효과(disciplinary effect)와 관련되어 있다. 만약 경평결과가 도덕적 해이와 관련된 경영관행을 바꾸는 기제로 제대로 작동하고 있다면 이번 기의 경평결과가 다음 기에 부채 비율에 영향을 미쳐야 한다. 특히 경평결과가 좋지 않았던 공공기관의 경우 다음 기에 부채비율의 정도가 개선되는 효과가 있어야 한다. 만약 그렇지 않다면 이는 경평제도가 도덕적 해이의 통제장치로서의 그 역할을 제대로 수행하고 있지 못하고 있다는 것을 시사한다.

부채비율 외에 복리후생비 또한 공공기관의 도덕적 해이의 정도를 보여주는 좋은 지표이다. 일반적으로 파산의 위험이 사라진 공공기관의 경우 그 종사자의 사적 편익과 관련된 비용 지출을 최적 수준보다 높게 잡을 유인이 크다. 물론 이러한 지출 항목 중 도덕적 해이와 관련된 가장 대표적 항목이 경영진(혹은 노동자)의 보수와 같은 인건비 항목일 것이다.

그러나 공공기관 인건비의 경우 기획재정부와 같은 예산당국의 지침에 의해 그 증가율이 비교적 철저하게 관리되고 있다. 이로 인해 기관 간 인건비 항목의 증가율의 차이를 찾는 것은 쉽지 않다. 따라서 본 연구에서는 인건비와 유사하면서도 예산 당국의 감독의 눈길이 상대적으로 덜 미쳐 있어 기관 간 차등 운영의 가능성이 있는 지표로 복리후생비를 선택하고 이것이 경쟁결과에 어떤 영향을 미치는지를 살펴보았다. 이와 관련된 것이 검증 명제 2-1이다.

검증 명제 2-1

복리후생비를 과다하게 지출하는 기관일수록 경영평가에서 불이익을 받을 것이다.

2. 평가지표설계의 적절성

앞서 설명했듯이 대리인의 최적 노력 수준은 이를 측정하는 지표가 부정확할수록 낮아질 수밖에 없다. 경쟁의 목적이 오로지 해당 공공기관의 경영효율성을 높이는 데 있다고 가정하자(즉 기관 고유사업을 통해 달성하고자 하는 공익성이란 목표는 잠시 접어두자). 경영효율성 제고라는 목적을 달성하기 위해서는 경쟁결과가 공공기관의 경영효율을 가장 잘 대변하는 정보라고 할 수 있는 재무정보에 의해 결정되어야 한다. 즉 재무성과가 좋으면 그 기관의 경쟁순위가 올라가야 한다. 그런데 어떠한 이유이든 간에 경쟁결과가 단지 심사를 대비할 수 있는 기관의 인적·물적 역량에 영향을 받는다고 하자. 만약 경쟁제도가 이러한 식으로 설계되어 있다면 공공기관들은 재무성과를 개선하여 경영효율성을 높이기 위한 근본적 노력보다는 그저 단기적으로 좋은 결과를 얻을 수 있는 편법을 찾아 여기에 노력과 시간을 투자할 것이다.

검증 명제 3

경평결과가 해당 기업의 공공성이나 효율성을 높이기 위한 노력과 무관한 요인에 의해 영향받아서 안된다.

공공성이나 효율성을 제고하기 위한 노력과 무관한 요인들은 무엇이 있을까? 선행 연구들은 경영평가를 받는 주체들이 보다 좋은 결과를 얻기 위해 평가준비에 적지 않은 인적·물적 자원을 동원하고 있음을 지적하고 있다. 예컨대 김완희(2010)의 경우 공공기관의 계량평가만 하더라도 그 산식 및 평가방법이 복잡해서 이를 이해하는 데 적지 않은 시간이 걸린다는 점을 지적하고 있다. 라영재·윤택범(2013)의 경우 평가를 준비하는 공공기관 입장에서 1년이란 평가 주기는 너무 짧아 평가담당자의 부담이 가중되고 있다는 점을 지적하고 있다. 특히 이러한 불만은 소규모 기관을 중심으로 꾸준히 제기되고 있다.

경평과정에 수반되는 과도한 평가 부담의 문제는 경평결과가 경평을 준비하는 기관의 자원동원능력에 좌우되는 부작용을 낳는다. 결국 기관의 자원동원능력은 기관의 크기에 의존한다. 특히 인력이 큰 조직일수록 경평 준비만을 전담하는 인력을 둘 수 있고 이것이 경평결과에 영향을 미칠 수 있다. 따라서 연구에서는 공공기관의 인력과 자산에 대한 정보를 사용하여 이러한 요인들이 경평결과에 영향을 주는지 살펴볼 것이다.

기관의 규모 외에 해당 공공기관의 업력(業歷) 또한 경평결과에 영향을 미칠 수 있다. 갓 설립된 기관보다 오래 전에 설립된 기관일수록 평가에 대한 노하우를 많이 축적하고 있을 수 있다. 이것이 조직 내부의 무형 지식으로 남아 담당자에게 전수된다면 경영평가에 좋은 점수를 받을 수 있다.

명제 3을 검증하는 또 다른 방법은 평가결과의 연도별 변동 정도를 추적하는 것이다. 평가결과의 연도별 변이를 살펴보는 것이 명제 3을 검증하는 방법이 되는 이유는 다음과 같다. 만약 경평결과가 공공기관의 효율성을 보여주는 지표에 의해 결정된다고 하자. 공공기관의 효율성은 많은 경우 단기

적으로 쉽게 변하지 않는 기관의 운영 능력에 의해 결정된다. 사실 이러한 능력은 장기적인 계획하의 지속적인 투자와 노력을 통해서만 향상될 수 있다. 따라서 이러한 요인이 경평결과를 결정한다고 하면 평가 상위그룹에 속해있던 기관이 1년 혹은 2년이라는 단기간에 하위그룹으로 떨어지는 일은 매우 드물게 발생할 것이다(반대의 경우도 마찬가지일 것이다).

반면에 경평결과가 평가대응 능력과 같은 비본질적 요인이나 운과 같은 일시적·외생적 요인에 의해 결정된다고 가정하자. 이 경우에는 특정 기관의 경평 순위의 연도별 변동 폭이 크게 나타날 가능성이 있다. 그리고 장기적으로는 ‘평균으로의 회귀(mean reversion)’와 같은 현상이 목격될 가능성도 있다.

3. 경영평가 성과급 설계의 적정성

앞서 언급했듯이 경평제도는 객관적 평가가 어려운 여러 항목들을 평가지표에 포함시키고 있다. 이 중에서도 리더십이나 책임경영이 대표적인 항목이다. 『2012년도 공공기관 경영평가 편람』에 따르면 리더십·책임경영 범주는 ▲ 리더십 ▲ 책임경영 ▲ 국민평가 ▲ 사회적 기여 ▲ 정부권장정책 등 5가지의 세부 지표로 구성되어 있다(기획재정부, 2011).

공기업의 경우 이 범주에 총 20점이 배점되어 있다. 리더십(5점), 책임경영(3점), 사회 공헌(2점)와 같은 비계량지표에 10점이, 그리고 국민평가(5점)와 정부권장정책(5점)과 같은 계량지표에 나머지 10점이 주어져 있다.⁴⁾ 여기서 언급하는 ‘리더십’의 정의는 다음과 같다: “기관의 비전과 목표를 구현하기 위한 경영전략의 수립과 시행, 주요 현안 과제 해결 등을 위한 기관장의 노력과 성과”. 이 밖에 책임경영은 주로 공사와 지배구조 및 윤리경영시스템 항목들을, 사회공헌은 일자리 창출, 중소기업 등과의 동반 성장, 사회봉사 활동의 항목들을 평가한다.

이러한 항목에 대한 기관의 노력을 평가할 때 발생하는 문제점은 크게 두

4) 준정부기관의 경우는 국민평가에 8점이 배정되어 계량점수가 13점, 비계량 점수가 10점이다.

가지로 요약될 수 있다. 첫 번째는 이미 앞에서 거론한 것으로 공공기관의 재무나 사업성과 외에 별도로 리더십이나 책임경영 항목을 평가하는 근거에 대한 의문이다. 연혁적으로 보면 이는 Malcom Baldrige 모델을 경영평가에 적용한 산물이다. 즉 평가영역을 그 과정의 관점에서 관리방법-실행-학습의 단계로 구분하는 것을 공공기관 평가에 적용한 것이다(김완희, 2010). 문제는 앞서도 잠시 언급했듯이 최종 성과에 대한 평가와 별개로 계획과 집행과정을 평가하는 것의 의미는 다음의 두 가지 경우에만 의미가 있다. 계획수립과 그 집행과정이 불량함에도 불구하고 결과가 양호하게 측정하는 것(혹은 그 반대). 평가의 일관성이란 관점에서 우리는 어떤 평가결과가 다른 평가결과보다 더 신뢰할 만한가를 판단해야 한다. 이는 결국 어떤 지표가 더 신뢰할 만한가의 문제로 귀결된다. 이 경우 일반적으로 성과에 대한 평가결과가 계획의 수립이나 수행 과정에 대한 평가보다 더 신뢰할 만하다.

사실 공공기관의 계획이나 과정에 대한 평가의 일환으로 진행되는 리더십이나 책임경영에 대한 평가가 유의미한 경우는 계획이나 과정이 좋았음에도 불구하고 그 기관의 재무 건전성이나 주요 사업 결과가 부정적으로 나왔을 때이다.⁵⁾ 그러나 이 경우조차도 여전히 리더십에 대한 평가가 성과를 높이 고자 하는 기관의 노력 수준에 대해 정확한 정보를 제공해주는가에 대한 의문은 여전히 남는다. 왜냐하면 본질적으로 리더십이나 책임경영 항목의 지표들이 기관의 노력과 성과를 제대로 반영한 지표인가의 문제가 남기 때문이다.

사실 이러한 문제는 비계량(혹은 정성적) 항목의 평가에 있어 보편적으로 나타나는 현상이다. 만약 어느 공공기관(예컨대 가스공사)이 특정 해에 중소기업과의 협력이나 사회봉사활동을 소홀히 했고 이러한 이유로 사회공헌항목의 점수가 낮았다고 가정해보자. 이러한 낮은 사회공헌항목 점수가 가스공사의 공공성이나 효율성 제고를 위한 노력에 대한 정확한 정보를 제공해

5) 마지막 경우 외부 충격(예컨대 금융위기)으로 인해 결과가 안 좋은 경우를 생각해 볼 수 있다. 즉 리더십은 좋았는데 외부 충격으로 인해 그것이 경영효율이나 주요 사업의 성과에 반영되지 않았다는 것이다. 그러나 이 경우에도 경영평가는 공공기관의 유형별 평가라는 상대평가(relative performance) 방식을 채택하고 있기 때문에 이를 리더십과 같은 항목을 도입하는 것으로 해결할 이유는 없다.

주는가?

실제 공공기관 경영평가 위원들을 상대로 한 조사에 따르면 리더십이나 책임경영 평가항목의 국민평가나 정부권장정책 지표의 경우 기관의 노력과 성과를 반영하는 지표인가라는 질문에 긍정적으로 답변한 비율은 각각 약 32%와 34% 수준에 불과했다. 오히려 이는 측정오차가 굉장히 큰 평가지표라고 봐야 한다(라영재·윤태범, 2013). 정부 역시 리더십·책임경영평가 범주를 아예 폐지하고 이를 경영성과 협약제의 평가 요소로 이관한다는 안을 발표한 바 있다(라영재·윤태범, 2013).⁶⁾

현행 제도에서 리더십·책임경영 범주에 대한 기관의 노력을 평가하는 것을 문제 삼는 가장 큰 이유는 이러한 범주에 대한 평가결과가 기관의 경영순위에 큰 영향을 미쳐 사실상 경영성과급의 크기를 결정하는 상황이 발생할 수 있기 때문이다. 만약 경영의 결과가 이처럼 부정확한 지표에 의해 좌우되었을 때 나타나는 문제점에 대해서는 앞에서 이미 설명한 바 있다. 여기서는 이를 간단한 수식을 통해 설명해보자.

즉 현재의 공공기관 종사자의 보수(w)는 기본적으로 경영효율, 리더십 및 책임경영, 주요 사업에 대한 평가지표의 점수(이를 각각 q_1 , q_2 , q_3 라고 하자)의 합에 의존한다. 여기서 보수는 경영결과에 영향을 받지 않는 부분(t)과 이에 연동된 성과급($sf(q_1+q_2+q_3)$)으로 구성되어 있다. s 는 전체 보수에서 성과급이 차지하는 비중이다.

$$w = t + sf(q_1 + q_2 + q_3) \quad (\text{식 IV-1})$$

$f(\)$ 는 경영 등급 함수

그런데 만약 실제 공공기관들이 경영효율이나 주요 사업분야의 점수의 기관별 편차가 거의 없다면 경영등급은 사실상 리더십 및 책임경영 항목의 지표(q_2)의 결과가 좌우하게 된다. 이 경우 보수함수(w)는 사실상 아래의 식처럼 변한다.

6) 그러나 이는 실제 실현되지 않았다.

$$w = t + s_2 f(q_2) \quad (\text{식 IV-2})$$

$$q_2 = a_2 + \varepsilon_2, \varepsilon \sim N(0, \sigma^2)$$

리더십 지표는 그 지표를 측정하고자 했던 본질적인 요소(예컨대 기관의 공공성과 효율성을 제고하는 데 기여하는 기관장의 노력)와 측정오차(ε_2)에 의해 영향을 받는다. 현재의 상황은 리더십 지표(q_2)가 적절하게 설정되지 못하여 지표가 기관장의 노력에 대한 정확한 정보를 제공해주지 못하는 상황(ε_2 가 큰 경우)이다. 이로 인해 리더십을 높이기 위한 노력(a_2)은 제대로 평가받지 못한다. 그렇다면 이 경우 공공기관 직원들의 보수는 사실상 자신의 노력과 무관한 외적 요인에 의해 결정되는 것과 마찬가지이다. 지금까지의 분석으로부터 본 연구의 마지막 검증 명제가 도출된다.

검증 명제 4

리더십이나 책임경영항목 지표처럼 해당 기관의 공공성에 대한 정확한 정보를 담고 있다고 보기 어려운 항목에 의해 경평결과가 크게 영향을 받아서는 안 된다.

V. 실증분석 결과

1. 공공기관 경영평가 결과에 대한 기초통계 분석

가. 공공기관 경영평가 순위에 대한 기초통계 분석

본 연구는 2008년부터 2013년까지 경평의 대상이 된 공공기관 전체(총 138개 공공기관)를 분석대상으로 설정하고 공공기관 경영제도가 일종의 유인 계약의 기제로 작용하는가를 실증적으로 분석한다.

제II장에서도 소개한 바처럼 경영제도는 크게 세 가지 지표, 1) 리더십·책임경영, 2) 경영효율, 3) 주요사업을 평가하며, 개별 지표별 세부 평가 항목을 바탕으로 경영실적을 평가하여 점수를 환산하고, 개별 지표 점수에 대한 적절한 가중치를 적용하여 백점 만점의 총점수를 계산한다. 이를 토대로 공공기관 유형별 순위를 매기고, 최종적으로 등급을 확정한다. 평가 등급은 S에서 E로 6단계이나, S등급은 아주 드물게 관측되었다. 그후 이러한 최종 등급을 토대로 매년 공공기관 경영결과를 공표한다.

이처럼 공공기관 경영 제도는 평가 점수를 바탕으로 공공기관의 성과를 유형별로 상대적인 평가를 하는 제도라고 볼 수 있으며, 그렇기 때문에 공공기관 유형 내 경영평가 순위가 공공기관 경영의 최종결과물이라고 볼 수 있다. 이에 본 연구에서는 경영의 연도별 순위를 공공기관 경영실적을 대표하는 최종지표로 간주하고, 경영제도에 대한 실증 분석을 시도할 때 주요 종속변수로 활용할 것이다.

여기서는 먼저 이러한 경영 순위가 어떠한 행태를 나타내고 있으며, 그 함의가 무엇인지에 대해 분석해 보고자 한다.

순위(rank)는 상대적인 개념이기 때문에, 이를 바로 계량적인 지표로 활용

하는 것은 올바른 활용법이라 할 수 없다. 이는 100개의 기관 중 한 계단 상승과, 1000개의 기관 중 한 계단 상승이 같은 의미를 지니지 않기 때문이다. 따라서 본 연구에서는 경쟁 대상에 포함되는(유형별) 공공기관의 수가 매년 변화한다는 것을 고려하여 연도별·공공기관 유형별 표준화 순위(standardized rank)를 다음과 같이 구성하였다.

$$\text{표준화 순위: } zRank_{ijt} = \frac{Rank_{ijt}}{(\# \text{ of firms})_{jt}} \times 100 \quad (\text{식 V-1})$$

구체적으로 t 기의 j 유형 공공기관 i 의 표준화 순위($zRank_{ijt}$)는 t 기의 공공기관 i 의 j 유형 내 순위($Rank_{ijt}$)를 t 기 j 유형의 공공기관의 수($(\# \text{ of firms})_{jt}$)로 나눈 값에 100을 곱한 값으로 정의하였다.

이와 같은 표준화 순위는 일종의 가중치 조정 순위라고 볼 수 있으며, 거리개념을 차용하면 이해하기 쉽다. 즉 표준화 순위는 n 개의 공공기관이 0에서 100 사이에 동일한 간격($100/n$)으로 평가점수의 순서대로 정렬한 상태라고 볼 수 있으며, 이에 최저 순위(혹은 최대값)는 항상 100이 된다. 따라서 표준화 순위에서 1계단 상승은 100개의 기관 중에서는 1계단 상승을, 1000개의 기관 중에서는 10계단 상승을 의미한다. 이와 같이 표준화 순위는 상대적인 개념을 차용하고 있기 때문에, 공공기관 유형별·연도별 비교를 가능케 한다는 장점을 지닌다.

이러한 표준화 순위를 활용하여 우리는 공공기관 경쟁제도의 실효성과 관련한 다음 두 가지 질문에 대한 답을 찾는 것이 가능하다. 첫 번째 질문은 경쟁결과에 연도별 변동성이 어떠한 양상을 나타내며 그 크기는 어느 정도인가이다. 이는 공공기관의 경영 개선의 노력에 따라 순위가 변동될 수 있는 가능성 유무와 그 크기에 대한 질문으로 볼 수 있다.

두 번째는 경쟁순위에 지속성이 나타나는가이다. 이 질문은 실제로 공공기관 경쟁결과에 빈익빈 부익부 현상이 나타나는가, 즉 고평가를 받은 공공기관이 지속적으로 고평가를 받고 저평가를 받는 기관이 계속 저평가를 받

는 현상이 나타나는가를 묻고 있다. 이는 반대로 고평가를 받은 기관이 저평가를 받고, 저평가를 받은 기관이 고평가를 받는 순위 전이 확률(transition probability)의 크기에 대한 질문이며, 동시에 개별 기관의 경영 개선의 노력이 실제 결과로 반영될 확률의 크기에 대한 질문이기도 하다.

기본적으로 시장에서 효율성 제고의 압력과 동시에 인센티브를 지속적으로 받는 사기업과 달리, 공공기관은 경영 효율성 제고에 대한 인센티브가 미미하다. 공공기관 경평제도는 이러한 인센티브를 제공하여 경영활동의 효율성을 제고하는 데 목적이 있으므로, 경영평가의 최종 결과물인 순위에 유의미한 변동이 나타나는 것은 경평 제도의 실효성을 판단할 때 매우 중요한 판단 근거가 될 수 있다. 상기한 두 가지 질문은 경평순위에 유의미한 변동이 나타나고 있는가를 묻고 있는 것으로 볼 수 있다.

공공기관 경평순위 변동성의 크기를 검정하기 위해, 앞서 정의한 표준화 순위의 1계 차분값(1st difference)을 통해 순위 변동폭에 대한 기초통계량을 살펴보자. 이를 위해 표준화 순위 변동은 $\Delta nRank_{ijt} = nRank_{ijt} - nRank_{ijt-1}$ 로 정의한다. 여기서 순위가 높을수록 그 값이 작기 때문에 차분한 값이 음의 값을 가질수록 순위가 상승한 것을 의미한다는 점에 유의할 필요가 있다.

먼저 <표 V-1>은 표준화 순위 변동의 기초통계량을 보여주는데, 여기서 우리는 몇 가지의 사실을 확인할 수 있다. 첫째는 표준화 순위 변동의 평균값이 0에 근접한다는 것이다. 표준화 순위 변동의 평균값은 큰 폭의 순위 상승이 많으면 음의 값을, 큰 폭의 순위 하락이 많으면 양의 값을 가지게 되나, 평균값이 대체로 0에 근접한다는 것은 순위 변동이 어느 한쪽에 치우침이 없이 나타났다는 것을 의미한다.

둘째, 표준화 순위 변동의 최댓값이 100이라는 점을 감안할 때, 표준편차가 30.2와 32.7이라는 사실은 순위의 연도별 변동성이 상당히 큰 편이라는 것을 의미한다. 또한 표준화 순위 변동 절댓값의 평균값은 25.5로 나타났는데, 이는 개별 공공기관이 자신들의 순위가 매년 평균적으로 25.5% 감소하거나 증가할 것을 기대할 수 있음을 뜻하여, 순위에 상당히 큰 폭의 변동성이 존재한다는 것을 보여준다.

마지막으로 표준화 순위 변동의 최솟값과 최댓값이 각각 -100과 100에 근접한다는 사실을 알 수 있는데, 이는 실제로 공공기관의 경평 순위가 최상위 순위에서 최하위 순위로, 최하위 순위에서 최상위 순위로의 이동이 존재한다는 것을 의미한다.

〈표 V-1〉 표준화 순위 변동의 기초통계(2007 ~ 2013년)

변수	관측치	평균	표준편차	최솟값	최댓값
표준화 순위 변동	596	-0.6023	32.6696	-92.1569	93.3333
표준화 순위 변동의 절댓값	596	25.5017	20.4019	0.0000	93.3333

자료: 한국조세재정연구원 내부자료 이용 저자 작성

〈표 V-2〉는 〈표 V-1〉의 기초통계량을 연도별로 계산하여 제시하였는데, 주요 현상들은 앞의 결과와 크게 다르지 않다. 즉 표준화 순위 변동의 평균값이 0에 근접하고 있다는 것과 순위의 변동성이 상당히 크다는 점에서 동일한 패턴을 보여준다.

〈표 V-2〉 연도별 표준화 순위 변동의 기초통계량

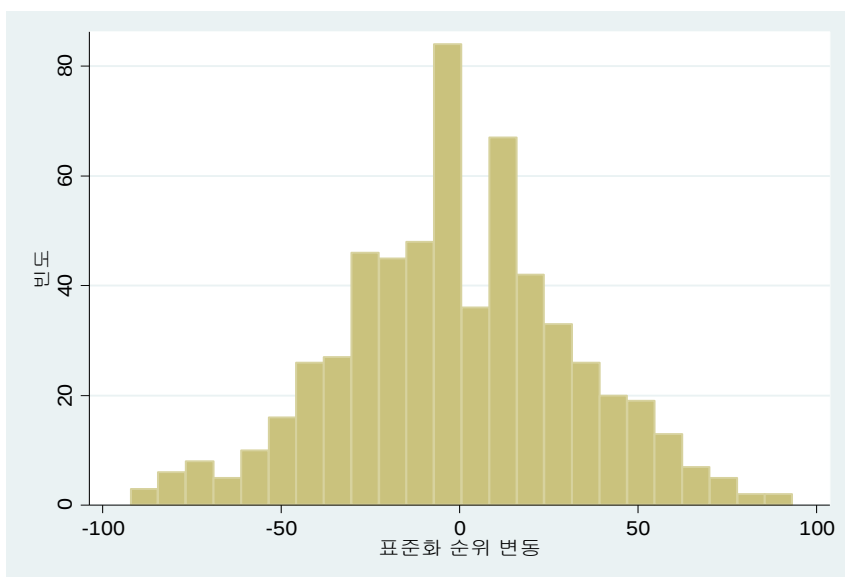
년도	표준화 순위 변동		
	관측치	평균	표준편차
2008	100	1.0000	36.6093
2009	84	-0.9133	34.0343
2010	94	-0.0163	31.9876
2011	99	-1.2206	30.2410
2012	109	-1.0164	32.8802
2013	110	-1.3552	30.9282
Total	596	-0.6023	32.6696
년도	표준화 순위 변동의 절댓값		
	관측치	평균	표준편차
2008	100	29.5329	21.4533
2009	84	26.7548	20.8504
2010	94	25.4783	19.1595
2011	99	22.8428	19.7207
2012	109	24.3321	22.0140
2013	110	24.4521	18.8419
Total	596	25.5017	20.4019

자료: 한국조세재정연구원 내부자료 이용 저자 작성

[그림 V-1]과 [그림 V-2]는 표준화 순위의 변동성을 더욱 자세히 알아보기 위한 것으로서, [그림 V-1]은 표준화 순위 변동의 분포를, [그림 V-2]는 연도별 표준화 순위 변동의 분포를 보여준다.

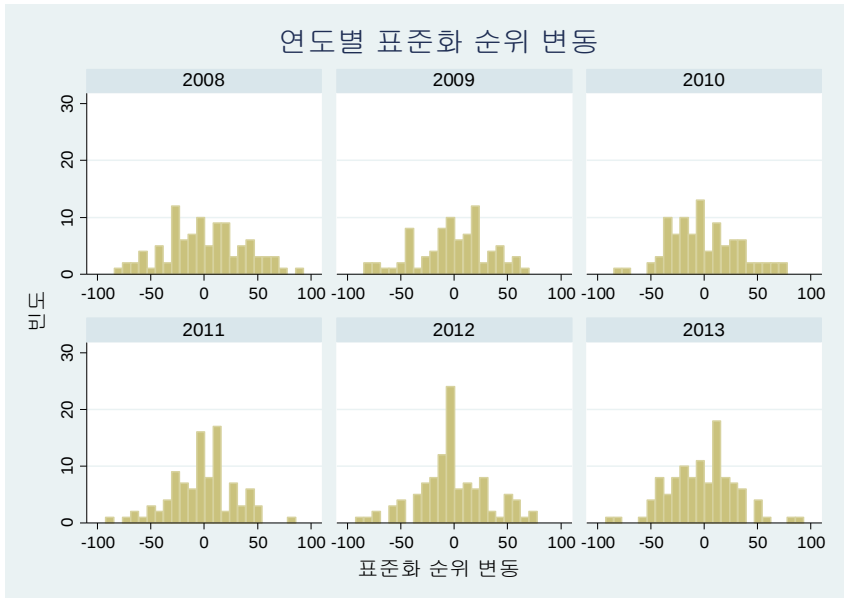
[그림 V-1]에서 한 가지 관찰되는 현상은 순위 변동의 값들이 최댓값과 최솟값에 걸쳐서 매우 넓고 고르게 퍼져 있으며, 정규분포의 형태를 닮아 있다는 것이다. 연도별 표준화 순위 변동의 분포를 보여주는 [그림 V-2]도 대체로 비슷한 형태가 나타나고 있으나, 연도별로 약간의 차이점이 존재한다. 2012년 이전의 분포는 이후 연도에 비해서 약간의 끌림 현상이 나타남과 동시에 분포의 형태 또한 상대적으로 고르지 않다. 반면 최근(2012년과 2013년)의 경향을 보면, 순위 변동의 분포가 더욱 고르고 넓게 퍼진 정규분포의 모습을 닮아가고 있다는 것을 확인할 수 있다.

[그림 V-1] 표준화 순위 변동의 분포



자료: 한국조세재정연구원 내부자료 이용 저자 작성.

[그림 V-2] 연도별 표준화 순위 변동의 분포



자료: 한국조세재정연구원 내부자료 이용 저자 작성.

이상의 결과에서 순위의 변동이 평균적으로 0으로 수렴하는 현상은 자연스럽다고 볼 수 있지만, 순위의 분포가 상당히 고르고 넓게 퍼져 있으며 그 변동 폭이 평균적으로 20%를 상회하고 있다는 사실은 현재 관찰되는 순위 변동이 유의미한 변화인가에 대한 의문점을 남긴다. 물론 순위 변동이 활발하게 일어나는 것은 경영 개선을 위한 공공기관 간의 치열한 경쟁의 실제 발현이라고 해석할 수도 있다. 그럼에도 불구하고 상대적으로 변동 폭이 큰 순위 변경은 경쟁제도가 실제로 공공기관의 경영활동의 혁신성, 경영능력, 혹은 경영성과를 제대로 평가하고 있는가에 대해서는 여전히 의문을 제기하게 만든다.

공공기관의 경영능력 및 성과는 단기적인 노력에 의해서도 개선될 수 있지만, 기본적으로 기관이 가지고 있는 고유의 내재적 능력에 의해 좌우될 것으로 예상할 수 있다. 그리고 이러한 기관의 내재적 능력은 장기적인 투자와 개선을 통해서 증진될 것으로 생각된다. 그렇기 때문에 순위 변동이

연도별로 빈번하게 발생하고 그 변동 폭 또한 크다는 사실은 현재의 경평제도가 경영활동의 근원적인 부분보다는 단기적인 경영 개선 노력에 더욱 집중하여 평가에 반영하도록 설계되어 있다는 근거를 제시하고 있는 것으로 볼 수 있다.

이와 같은 사실은 **검증 명제 3**에서 제시한 것처럼 경영평가의 결과가 효율성이나 공공성 제고를 위한 기관의 노력 수준과 관련이 없는 비본질적 요인에 의해서 쉽게 영향을 받을 가능성이 높다는 것을 의미한다. 그렇기 때문에, 현재의 경평제도에 적절한 평가지표에 대한 개선이 필요하다는 것을 암시하고 있다.

공공기관 경평 과정에서 매년 제기되는 불만 중 하나는 현재의 제도가 특정한 공공기관이 매년 좋은 평가를 받도록 디자인되어 있다는 주장이다. 이는 고평가를 받은 공공기관이 지속적으로 고평가를 받고 저평가를 받는 기관이 계속 저평가를 받는 빈익빈 부익부 현상이 경영평가에서도 나타나고 있으며, 그 강도 또한 높다는 것이다. 물론 경평순위에 어느 정도의 지속성이 나타나는 것은 자연스러운 현상이지만, 그 지속성의 정도가 매우 높으면 공공기관의 경영 개선노력의 의지를 저하시킬 수 있기 때문에 문제가 될 수 있다. 따라서 여기서는 표준화 순위의 전이확률행렬을 계산하여 순위의 지속성의 정도에 대해 살펴본다.

전이확률행렬은 전기의 상태 공간(state space)에서 현재의 상태 공간으로 이행하는 확률로 정의할 수 있는데, n 개의 상태(state)가 있다고 할 때 전이확률행렬은 $n \times n$ 행렬이 되며 (i, j) 요소(element)는 전기의 i 의 상태에서 현재 j 의 상태로 전이할 확률을 의미한다.

표준화 순위의 전이확률행렬을 구하기 위해서 먼저 경영평가 최종 점수를 기준으로 30% 이하는 하위그룹(LOW), 30% 초과 70% 이하는 중간그룹(MIDDLE), 70% 초과는 상위그룹(HIGH)으로 정의하여 상태공간을 구성하였다. 그후 t-1기에서 t기로 개별 순위 그룹이 이동하는 경로를 추적하여 표준화 순위의 전이확률행렬을 계산하였다.

〈표 V-3〉은 전체 자료를 활용하여 표준화 순위에 대한 전이이행확률을

구하였는데, 각각의 값들은 전기의 순위 그룹에서 다음 기의 순위 그룹으로 전이될 확률을 의미한다. 예컨대, 〈표 V-3〉의 (LOW, MIDDLE) 요소값(element)은 t-1기에 LOW일 때, t기에 MIDDLE로 전환될 확률이 9%라는 것을 알려 준다. 따라서 전이확률행렬의 대각요소(diagonal element)들 (즉, LOW → LOW, MIDDLE → MIDDLE, HIGH → HIGH)은 과거의 상태가 현재에도 지속적으로 나타날 확률을 의미하기 때문에 순위의 지속성을 보는 척도가 될 수 있다. 반면 비대각요소(off-diagonal element)들은 현재의 상태가 과거의 상태를 이탈할 확률을 의미하게 된다.

이를 근거로 〈표 V-3〉을 해석하면, 먼저 개별 기관들이 과거의 순위 그룹에 지속적으로 머무를 확률이 약 45%라고 볼 수 있다. 이는 반대로 개별 기관들이 자신들의 과거 순위 그룹에서 이탈할 확률은 약 55%라는 것을 의미한다. 따라서 이러한 결과는 순위의 빈익빈 부익부 현상은 일부 존재하지만 이탈율이 매우 크기 때문에 지속성의 강도는 높지 않다는 근거를 제시한다.

〈표 V-4〉는 연도별 표준화 순위 전이확률행렬을 보여주는데, 결과는 〈표 V-3〉이 시사하는 바와 동일하다는 것을 확인할 수 있다. 따라서 객관적인 자료를 통해 보았을 때, 경영평가 제도가 특정한 공공기관이 특혜를 받는 구조로 디자인 되었다는 주장은 설득력이 낮은 것으로 판단된다. 오히려 순위 그룹의 지속성이 높고 그룹 이탈률이 약 55%로 매우 높다는 사실은 앞서 표준화 순위의 변동성에서 논의하였던 것과 같이 **검증 명제 3**에 대한 동일한 함의를 내포하고 있다고 생각된다.

기관의 경영성가를 결정하는 기관의 노력(혹은 내적인 능력)은 기관 고유의 자원으로 축적되기 때문에 연속성·지속성의 특성을 가지고 있다. 반면, 단기적인 경영활동 개선 노력으로 인한 경영성가는 상당 부분 불연속적·일시적일 것으로 예상되기 때문에 단기적인 순위 변동에 영향을 줄 것으로 기대된다. 따라서 순위그룹의 이탈률이 매우 높다는 것은 현재 경영 지표가 단기적인 경영 활동에 상대적으로 큰 가중치를 두어 평가하도록 고안되었다는 주장의 근거가 될 수 있으며, 이는 **검증 명제 3**에서 제시한 바람직한 평가지표의 기준과는 거리가 멀다고 볼 수 있다.

〈표 V-3〉 표준화 순위 전이확률행렬(transition probability matrix)

(t-1) 기 \ t 기	LOW	MIDDLE	HIGH
LOW	0.16	0.11	0.06
MIDDLE	0.09	0.13	0.12
HIGH	0.07	0.11	0.16

주: 전체 관측치는 596개.
자료: 한국조세재정연구원 내부자료 이용 저자 작성.

〈표 V-4〉 표준화 순위 변동의 연도별 전이확률행렬

2007년 \ 2008년	LOW	MIDDLE	HIGH
LOW	0.12	0.12	0.09
MIDDLE	0.12	0.11	0.10
HIGH	0.09	0.11	0.14
2008년 \ 2009년	LOW	MIDDLE	HIGH
LOW	0.14	0.12	0.06
MIDDLE	0.11	0.13	0.11
HIGH	0.06	0.10	0.18
2009년 \ 2010년	LOW	MIDDLE	HIGH
LOW	0.18	0.10	0.04
MIDDLE	0.07	0.12	0.16
HIGH	0.07	0.12	0.14
2010년 \ 2011년	LOW	MIDDLE	HIGH
LOW	0.20	0.05	0.07
MIDDLE	0.07	0.15	0.12
HIGH	0.05	0.12	0.16
2011년 \ 2012년	LOW	MIDDLE	HIGH
LOW	0.17	0.12	0.06
MIDDLE	0.08	0.12	0.12
HIGH	0.07	0.11	0.16
2012년 \ 2013년	LOW	MIDDLE	HIGH
LOW	0.15	0.14	0.04
MIDDLE	0.09	0.16	0.09
HIGH	0.05	0.08	0.19

주: 전체 관측치는 위에서부터 각각 100개, 84개, 94개, 99개, 109개, 110개.
자료: 한국조세재정연구원 내부자료 이용 저자 작성.

나. 공공기관 경영평가 주요 지표에 대한 기초통계 분석

공공기관 경영의 최종등급을 결정하는 총점수나 순위는 상대적인 개념이기 때문에, 연도별 변화가 있어도 표준화를 통해 충분히 비교가 가능하다. 반면 경영평가의 주요 지표 혹은 평가 항목의 경우에는 평가 기준 혹은 세부 항목이 바뀌면 전혀 다른 지표가 되기 때문에 의미 있는 비교 분석이 어렵다.

실제로 현재 본 연구의 분석기간(2008~2013년)에도 경영평가의 기준이 두 차례 전면적으로 교체가 이루어졌으며, 바로 이러한 점이 실증분석을 통해서 경영평가의 실효성을 검토하는 데 주요한 장애요인으로 작용한다. 따라서 여기서는 먼저 이러한 경영평가 지표의 일관성을 점검해 보고자 한다.

2011~2013년 기준으로 공공기관 경영 제도의 주요 지표는 크게 계량과 비계량지표 나뉘며, 대분류 항목은 리더십·책임경영(이하 리더십), 경영효율, 주요사업 세 가지 부문으로 구분되어 총 6개의 지표로 구성된다. 이러한 대분류 항목에 대한 주요 평가항목은 〈표 V-5〉에 정리하였다. 이 밖에도 각각의 주요 평가지표에 대한 세부평가항목들이 존재하여, 경영평가의 지표는 매우 다양하고 복잡하게 구성되어 있다.⁷⁾

〈표 V-5〉 경영평가 범주 및 평가지표 체계

평가범주	주요 평가지표
리더십·책임경영	• 리더십 • 책임경영 • 국민평가 • 사회적 기여(사회공헌, 정부권장정책)
경영효율	• 업무효율 • 조직 및 인적자원관리 • 재무예산관리 및 성과 (재무예산관리, 재무예산성과, 계량관리업무비) * 연금금 유형은 기금운용관리 및 성과 • 보수 및 성과관리(보수 및 성과관리, 총인건비 인상률) • 노사관리
주요사업	공공기관의 주요사업별 계획·활동·성과를 종합적으로 평가

자료: 라영재·윤태범(2013), p. 61.

7) 세부평가항목에 대한 자세한 사항은 기획재정부, 「2012년도 공공기관 경영평가편람」, 2011을 참조.

또한 이러한 세부평가항목들도 경쟁구조가 바뀔 때마다 생성과 소멸을 반복하며, 그 가중치나 점수 편성 등도 자주 변경되기 때문에 일관성이 결여되어 있다. 따라서 본 연구에서는 최대한 일관성 있게 활용할 수 있는 6개의 대분류 항목, 즉 리더십 계량, 경영효율 계량, 주요사업 계량, 리더십 비계량, 경영효율 비계량, 주요사업 비계량의 지표들을 활용하여 분석에 사용하고자 한다.

비록 본 연구에서 6개의 대분류에 초점을 맞추어 분석을 실시한다고 하여도, 개별 지표의 일관성을 유지하는 데에는 여전히 많은 장애요소들이 존재한다. 특히 경쟁 제도 전반에 걸친 구조적인 변화가 나타났을 때(2009년과 2011년)에는 대분류 항목에서도 역시 지표 간 연계가 어려운 것이 사실이다. 또한 개별 지표 간 연계가 가능하더라도, 각각의 지표에 대한 가중치와 점수 편성 또한 달라지기 때문에 경쟁지표를 연속성 있게 분석을 하는 데 어려움이 따른다. 그렇기 때문에 경쟁지표의 일관성을 최대한 유지하기 위해서는 분석기간을 나누어 설정할 수밖에 없다. 이에 대한 근거를 살펴보기 위해서 먼저 각각의 주요 평가지표들의 분포를 살펴보겠다.

[그림 V-3]과 [그림 V-4]는 각각 계량과 비계량 평가지표들의 연도별 상자 그림(box plot)을 보여 주는데, 이를 통해 우리는 주요 평가지표들의 연도별 분포 형태를 확인하고 일관성을 점검할 수 있다.⁸⁾ 이 그림들을 통해 특징적으로 알 수 있는 사실은 주요 평가지표들이 2011년을 기점으로 상당히 다른 형태를 보여준다는 것이다.

먼저 2011년 이전에는 리더십 계량, 경영효율 계량, 주요사업 비계량이 6개의 대분류 항목으로 구분되지 못하고 다른 평가지표에 편입되었다는 것을 확인할 수 있다. 다음으로 2011년을 기점으로 두 기간(즉 2008~2010년과 2011~2013년)의 개별 지표의 분포가 상당히 유사하다는 것을 알 수 있다. 특히 2011~2013년 동안의 개별 지표의 분포는 그 범위와 수준이 매우 유사

8) 상자 그림(box plot)은 한 개의 그림에 최댓값, 최솟값, 중앙값, 사분편차를 사용하여 자료의 측정값들이 어떤 모양으로 분포하는가는 보여주기 때문에 개별 평가지표의 연도별 일관성을 점검하는 데 유용하게 사용할 수 있다.

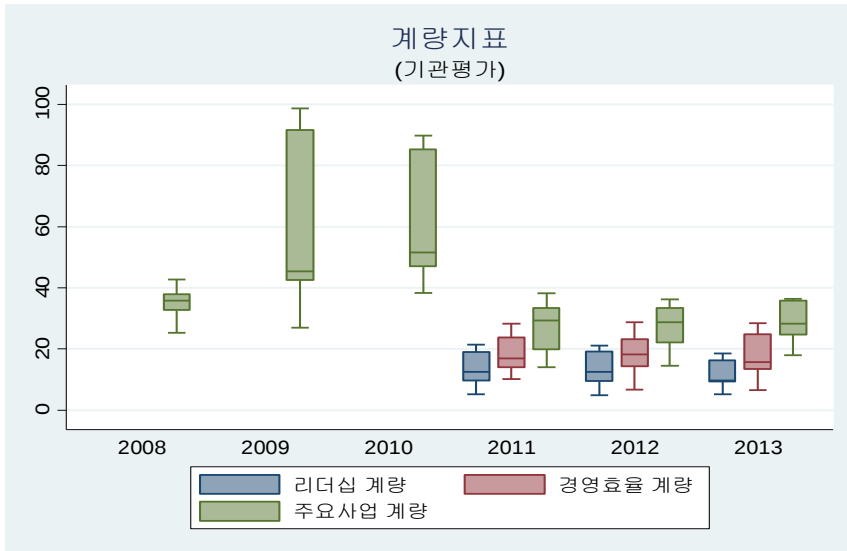
하고, 연도별로 상당히 일관성이 있는 것으로 보인다. 반면 2008~2010년 기간에서는 2008년 주요사업 계량 지표를 제외하고는 다른 모든 평가 지표들이 대체로 유사한 분포를 따르고 있는 것으로 보인다.

이렇듯 개별 평가지표들은 연도별로 상이한 분포를 가지고 있기 때문에, 개별 지표들에 무리하게 연속성을 부여하여 실증분석에 활용하는 것은 두 가지 면에서 타당하지 않다고 생각된다.

하나는, 상기한 바와 같이 경영평가 지표는 2011년을 기점으로 가중치나 점수편성의 측면에서 분절되어 있으며 세부항목 또한 완벽하게 일치하지 않기 때문에, 이를 무리하게 연결하게 되면 개별 평가지표들이 담고 있는 정보의 질과 구성이 동질적이지 않게 된다.

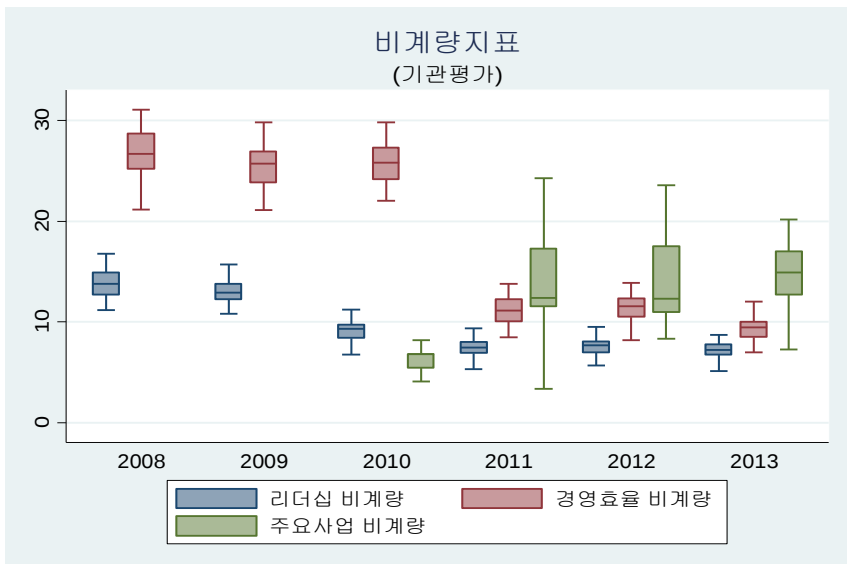
다른 하나는, 비록 개별 평가지표들이 지니고 있는 정보의 질을 유지한 채 연결시킬 수 있다고 하여도, 앞에서 확인하였듯이 두 기간 간에 평가지표의 수가 같지 않기 때문에 이들 지표에 연속성을 부여하게 되면 현재 평가지표들이 가지고 있는 정보에 어느 정도의 손실을 감수해야 할 수밖에 없다. 따라서 최대한 정보의 손실을 막고 일관성을 확보한 평가지표를 구성하기 위해서, 본 연구에서는 두 개의 분석기간을 기간 1(2008~2010년)과 기간 2(2011~2013년)로 나누어 정의하고 경평제도의 실효성에 대한 실증분석을 시도하고자 한다.

[그림 V-3] 연도별 경영평가 계량지표 분포



자료: 한국조세재정연구원 내부자료 이용 저자 작성

[그림 V-4] 연도별 경영평가 비계량지표 분포



자료: 한국조세재정연구원 내부자료 이용 저자 작성

〈표 V-6〉과 〈표 V-7〉은 두 개의 분석기간에 따른 개별 평가지표들의 기초통계량을 보여준다. 두 표를 통해서, 개별 평가지표의 기초 통계량이 두 분석기간 간에 상당히 다른 모습을 보이고 있다는 것을 알 수 있다. 이는 같은 지표라고 하여도 두 분석기간 간 정보의 양과 질이 다르다는 것을 보여주는 근거가 되며, 일관성을 유지하기 위해서 분석기간을 나누어야 한다는 본 연구의 방향을 지지해주고 있다고 생각된다.

〈표 V-6〉 경영평가 주요지표의 기초통계(2008 ~ 2010년)

변수	관측치	평균	표준편차	최솟값	최댓값
주요사업(계량)	296	53.8471	22.5692	23.1790	98.6600
리더십(비계량)	186	12.2153	2.3095	6.3000	16.8000
경영효율(비계량)	186	26.0219	2.3245	21.1000	31.1000

자료: 한국조세재정연구원 내부자료 이용 저자 작성.

〈표 V-7〉 경영평가 주요지표의 기초통계(2011 ~ 2013년)

변수	관측치	평균	표준편차	최솟값	최댓값
리더십(계량)	336	13.5122	4.6599	4.9253	21.4233
경영효율(계량)	337	18.5736	5.8077	6.6094	28.8102
주요사업(계량)	337	27.9246	6.2308	13.9940	38.1714
리더십(비계량)	184	7.4562	0.8523	5.1000	9.7895
경영효율(비계량)	184	10.5567	1.5973	5.8462	13.9000
주요사업(비계량)	337	14.1747	3.8159	3.3333	24.3000

자료: 한국조세재정연구원 내부자료 이용 저자 작성.

2. 검증 명제 실증분석

가. 경영평가의 결정요인 분석(검증 명제 4에 대한 검증)

먼저 공공기관 경평결과에 영향을 미치는 중요한 결정요인에 대해 알아보기 위해 개별 평가지표가 표준화 순위에 미치는 영향에 대한 회귀분석을 시행하였다. 물론 경영평가의 순위는 평가지표에 가중치를 적용하여 합산한 총점수로 결정되기 때문에 평가지표가 순위에 미치는 영향은 전반적으로 상당히 높을 것으로 쉽게 예상할 수 있다. 그럼에도 평가 지표가 경평순위에 미치는 영향의 정도는 개별 지표별로 다를 수 있기 때문에, 이를 점검해 볼 필요가 있다.

이와 같은 과정에서 우리는 경영평가 결과에 영향을 미치는 개별 지표의 상이한 중요도를 점검함으로써 **검증 명제 4**에 대한 가설을 검증할 수 있다. 즉 양적으로 측정이 어려운 지표들, 즉 리더십·책임경영의 항목 혹은 비계량 지표들이 그렇지 않는 지표들에 비해 경영평가에 미치는 중요도가 어떻게 다른지를 비교함으로써, 우리는 경평지표의 적절성에 대해 판단해 볼 수 있다.

〈표 V-8〉은 개별 평가지표가 표준화 순위에 미치는 영향에 대한 회귀분석 결과이다. 각 열은 기간 1과 기간 2에 대한 회귀분석의 결과이고, 공공기관 유형 내 표준화 역순위를 종속변수로 사용하였다. 또한 공공기관 유형별 이질적 효과와 연도별 경영평가의 특수성을 통제하기 위해 공공기관 유형과 연도별 시간 고정효과를 회귀분석에 포함하였다. 여기서 해석의 편의를 위해 앞에서 사용한 표준화 순위 대신에 표준화 역순위를 사용한 점에 주의할 필요가 있다.⁹⁾ 따라서 각 평가지표에 대한 양의 계수는 개별 지표들이 순위를 상승시키는 데 기여한다는 것을 의미한다.

9) 표준화 역순위는 표준화 순위와 마찬가지로 0과 100 사이에 값을 가지지만, 순위가 높을수록 높은 값을 가진다(예컨대, 1위는 100). 따라서 표준화 역순위에 대한 회귀분석 결과는 계수의 부호가 바뀌는 효과만이 나타난다.

〈표 V-8〉 경영평가 순위의 결정요인 분석

변수	종속변수: (유형 내) 표준화 역순위	
	기간 1: 2008~2010년	기간 2: 2011~2013년
리더십(계량)		4.0567 ^{***} (0.9937)
경영효율(계량)		4.6201 ^{***} (0.5384)
주요사업(계량)	4.9797 ^{***} (0.2552)	3.1667 ^{***} (0.5627)
리더십(비계량)	7.3128 ^{***} (1.0788)	4.7511 ^{**} (2.1322)
경영효율(비계량)	5.3011 ^{***} (0.6889)	7.3045 ^{***} (1.5143)
주요사업(비계량)		5.5456 ^{***} (0.9085)
공공기관 유형 더미	YES	YES
연도별 시간 고정효과	YES	YES
R ²	0.71	0.85
관측치	184	186

주: 1. 괄호 안의 값은 표준오차를 의미.

2. *, **, ***는 각각 10%, 5%, 1%의 유의수준을 의미.

자료: 한국조세재정연구원 내부자료 이용 저자 작성.

〈표 V-8〉을 통해 우리는 두 가지 결과를 도출해 낼 수 있다.

하나는, 기간 1에서는 리더십 지표가 순위에 미치는 영향은 가장 큰 것으로 나타났으며, 기간 2에서도 리더십 지표의 중요도가 상대적으로 매우 높다는 것을 알 수 있다. 이는 현재의 공공기관 경쟁제도가 기관들이 수행하고 있는 주요사업의 성과나 경영효율성 개선보다는 리더십·책임경영에 그 결과가 좌우되는 경향이 존재한다는 것을 의미한다.

다른 하나는 우리는 비계량지표의 계수가 계량지표의 계수들보다 상대적

으로 크다는 것을 알 수 있는데, 이는 비계량지표가 계량지표들보다 경영평가의 결과에 미치는 영향이 상대적으로 크다는 것으로 볼 수 있다. 따라서 현재의 공공기관 경평제도하에서 경평결과는 정량적 자료(quantitativ data)보다는 정성적 자료(quality data)에 더욱 의존하고 있는 것으로 판단된다.

이상의 결과는 양적으로 측정하기 어려워 측정오차가 큰 지표들, 즉 리더십·책임경영 혹은 비계량지표들이 경평결과에 더욱 주요한 요인으로 작용하고 있다는 것을 매우 일관되게 보여주고 있어, **검증 명제 4**와 상반되는 결론을 지지하고 있다.

이처럼 오차가 큰 지표들이 경영평가에 미치는 영향이 크다는 것은 경영평가의 결과가 다분히 운(luck)적인 요소에 좌우될 수 있는 가능성이 존재함을 시사하며, 그렇기 때문에 현재 경영평가제도의 평가지표가 과연 적절하게 설계되어 있는가에 대한 재검토가 필요할 것으로 생각된다.

현재까지 살펴본 바와 같이 평가지표의 적절성을 개별 지표의 특성(측정가능성)을 통해 평가할 수도 있지만, 경영평가 자체의 일관성을 통해서도 살펴볼 수 있다. 특히, 경영평가제도는 상기한 바와 같이 두 차례 대대적인 개정이 있었으며, 세부적으로는 매년 개정이 있어 왔다. 이와 같은 잦은 개정으로 인해 개별 지표가 평가결과에 미치는 중요도가 매년 변화한다면, 이는 공공기관이 경평결과를 개선하기 위한 목표를 설정하는 데 혼선을 야기할 가능성이 존재한다.

〈표 V-9〉 연도별 경영평가 순위의 결정요인 분석

변수	종속변수: (유형 내) 표준화 역순위					
	기간 1: 2008~2010년			기간 2: 2011~2013년		
	2008년	2009년	2010년	2011년	2012년	2013년
리더십(계량)	—	—	—	4.5485 ^{**} (1.7612)	2.9464 (1.9738)	5.7552 ^{***} (1.6518)
경영효율(계량)	—	—	—	3.9812 ^{***} (1.2396)	3.2702 ^{***} (1.0441)	6.0098 ^{***} (0.8834)
주요사업(계량)	5.3070 ^{***} (0.3784)	3.9450 ^{***} (0.4219)	6.9745 ^{***} (0.5368)	1.6888 (1.1961)	1.9694 (1.1947)	6.0892 ^{***} (0.8037)
리더십(비계량)	6.8426 ^{***} (1.4723)	9.3429 ^{***} (2.6518)	7.2693 ^{***} (1.6831)	2.5510 (4.2064)	11.5049 [*] (5.8138)	6.4592 ^{**} (2.6284)
경영효율(비계량)	4.0082 ^{***} (1.0198)	5.9066 ^{***} (1.6274)	6.3863 ^{***} (0.9646)	9.3750 ^{***} (2.8362)	3.0693 (3.0619)	5.9627 ^{**} (2.2482)
주요사업(비계량)	—	—	—	7.4014 ^{***} (2.1865)	4.0939 (3.2652)	5.6428 ^{***} (1.3035)
공공기관 유형 더미	YES	YES	YES	YES	YES	YES
R ²	0.87	0.86	0.91	0.69	0.72	0.86
관측치	75	56	55	62	60	62

주: 1. 괄호 안의 값은 표준오차를 의미.

2. *, **, ***는 각각 10%, 5%, 1%의 유의수준을 의미.

자료: 한국조세재정연구원 내부자료 이용 저자 작성.

이러한 가능성을 검정하기 위해, 본 연구에서는 연도별로 개별 지표가 경영평가의 결과에 미치는 중요도가 어떻게 변화하는가를 추정해 보았다. 구체적으로 이는 〈표 V-8〉과 동일한 분석을 연도별로 시행하는 것을 통해 가능하며, 그 결과를 〈표 V-9〉에 제시하였다. 추정 결과, 앞선 결과와 유사하게 리더십(비계량) 지표는 상당히 일관적으로 중요도가 가장 큰 것으로 나타났다. 반면 그 외의 지표들에서는 평가지표의 상대적인 중요도가 연도별로 부침이 존재하며, 이에 연도별 일관성을 유지하고 있다고 보기에는 힘든 것으로 판단된다.

나. 재무정보를 활용한 검증 명제 실증분석

주지하다시피 본 연구는 공공기관 경영제도를 일종의 유인계약으로 간주하고, 경제학적으로 중요한 경영평가제도의 목적을 공공기관의 도덕적 해이를 통제하는 역할에서 그 정당성을 찾고 있다. 즉 민간기업과 달리 공공기관은 정부의(암묵적) 보증으로 인해 파산이라는 경영규율 기제가 작용하지 않기 때문에 도덕적 해이가 나타날 가능성이 높다. 이러한 경제 환경에서 경영평가제도는 공공기관의 경영효율성을 증진시켜 도덕적 해이의 문제를 규율하는 역할을 할 수 있다는 것이다.

우리는 이미 제IV장에서 도덕적 해이의 통제장치로서의 경영평가의 역할에 대한 다양한 검증 명제를 제시한 바 있으며, 여기서는 공공기관의 재무정보와 경영평가 결과와 연계한 자료를 바탕으로 각각의 검증 명제를 실증적으로 검정한다.

경영평가 결과와 연계하기 위한 공공기관의 주요 재무정보는 알리오(www.alio.go.kr)를 통해 수집하였다. 재무정보와 연계한 결과 <표 V-10>에서 제시한 바와 같이 총 112개의 공공기관이 분석 대상에 포함되었다. 또한 알리오는 최근 5년간의 재무정보만을 공개하고 있기 때문에, 가용한 범위 내에서 분석기간을 기간 2(2011~2013년)로 설정하였다. 이처럼 구성된 자료를 활용하여, 본 연구의 실증분석에 사용된 주요 재무 변수의 기초통계량을 <표 V-11>에 제시하였다.

〈표 V-10〉 기관유형별 분석대상 공공기관

기관유형	기관명
시장형	부산항만공사, 인천국제공항공사, 한국가스공사, 한국공항공사, 한국광물자원공사, 한국남동발전, 한국남부발전, 한국동서발전, 한국서부발전, 한국석유공사, 한국수력원자력, 한국전력공사, 한국중부발전, 한국지역난방공사
준시장형	대한석탄공사, 대한주택보증주식회사, 여수광양항만공사, 울산항만공사, 인천항만공사, 제주국제자유도시개발센터, 한국감정원, 한국관광공사, 한국도로공사, 한국마사회, 한국방송광고진흥공사, 한국수자원공사, 한국조폐공사, 한국철도공사, 한국토지주택공사, 해양환경관리공단
위탁집행형	건강보험심사평가원, 교통안전공단, 국립공원관리공단, 국민건강보험공단, 국제방송교류재단, 국토교통과학기술진흥원, 농림수산물교육문화정보원, 농림수산물기술기획평가원, 농업기술실용화재단, 대한무역투자진흥공사, 대한지적공사, 도로교통공단, 독립기념관, 선박안전기술공단, 소상공인시장진흥공단, 에너지관리공단, 우체국금융개발원, 우체국물류지원단, 정보통신산업진흥원, 중소기업기술정보진흥원, 축산물안전관리인증원, 축산물품질평가원, 한국가스안전공사, 한국고용정보원, 한국과학창의재단, 한국광해관리공단, 한국교육학술정보원, 한국기상산업진흥원, 한국노인인력개발원, 한국농수산식품유통공사, 한국농어촌공사, 한국디자인진흥원, 한국방송통신전파진흥원, 한국보건복지인력개발원, 한국보건복지정보개발원, 한국보건산업진흥원, 한국보훈복지의료공단, 한국산업기술진흥원, 한국산업기술평가관리원, 한국산업단지공단, 한국산업안전보건공단, 한국산업인력공단, 한국석유관리원, 한국세라믹기술원, 한국소방산업기술원, 한국소비자원, 한국수산물관리공단, 한국시설안전공단, 한국에너지기술평가원, 한국연구재단, 한국예탁결제원, 한국우편사업진흥원, 한국원자력안전기술원, 한국인터넷진흥원, 한국임업진흥원, 한국장애인고용공단, 한국장학재단, 한국전기안전공사, 한국전력거래소, 한국정보화진흥원, 한국철도시설공단, 한국청소년상담복지개발원, 한국청소년활동진흥원, 한국콘텐츠진흥원, 한국해양수산연수원, 한국환경공단, 한국환경산업기술원
기금관리형	공무원연금공단, 국민연금공단, 근로복지공단, 기술신용보증기금, 사립학교교직원연금공단, 신용보증기금, 영화진흥위원회, 예금보험공사, 중소기업진흥공단, 한국무역보험공사, 한국문화예술위원회, 한국언론진흥재단, 한국원자력환경공단, 한국자산관리공사, 한국주택금융공사
총계	112개 공공기관

자료: 한국조세재정연구원 내부자료 이용 저자 작성.

〈표 V-11〉 주요 재무정보의 기초통계량

변수명	관측치	평균	표준편차	최솟값	최댓값
부채/자본 비율	297	14.8676	162,3401	-530,559	2421,092
ROA(당기순이익/자산)	289	0,021806	0,109946	-0,90125	0,761841
자산(억원)	297	242,8653	4059,902	0,001204	69973,19
정규직 1인당 평균보수액(백만원)	318	64,91797	12,37033	35,894	101,021
정규직 1인당 복리후생비(백만원)	318	1,666072	1,939957	0,015	15,442
정규직 상시근로자 수(천명)	318	1,563577	3,720349	0,0498	29,732
정규직 평균근속년수(년)	318	12,97264	4,553116	0,81	22,79
공공기관 업력(년)	320	23,38125	20,05349	0	105
기관장출신 더미					
정치인	320	0,04375	0,204859	0	1
관료	320	0,528125	0,49999	0	1
공공기관 유형 더미					
기금관리형	320	0,134375	0,341589	0	1
시장형	320	0,13125	0,338202	0	1
위탁집행형	320	0,596875	0,4912937	0	1
준시장형	320	0,1375	0,344914	0	1

자료: 재무자료는 Alio 자료 이용 저자 작성. 기관장 출신에 대한 정보는 한국조세재정연구원 내부자료 이용 저자 작성.

본 연구에서 주목하고 있는 재무정보를 활용한 실증분석은 1) 공공기관의 주요 재무변수가 경영결과(표준화 역순위 및 평가등급)에 미치는 효과와 2) 전기의 경영평가의 결과가 공공기관의 주요 재무변수에 미치는 효과, 크게 두 가지 테마를 중심으로 이루어져 있다.

먼저 〈표 V-12〉는 주요 재무변수가 경영평가 결과에 미치는 효과에 대해 추정하였다. (1)결과 (3)열에서는 표준화 역순위를, (2)결과 (4)열에서는 공공기관 경영평가등급을 종속변수로 사용하였다. 경영평가 등급은 A(S 포함)에서 E까지 총 5단계¹⁰⁾로 구분하였으며, 종속변수의 잠재변수(latent

10) 본 연구의 분석기간 동안 경영평가에서 S등급을 받는 공공기관은 단 1개의 관측치만이 존재하기 때문에 A등급에 포함시켜 분석하여도 무방하다고 생각된다.

variable)적 성격을 고려하여 순서로짓모형(ordered logit model)을 적용하였다. 또한 실증분석 결과의 강건성(robustness)을 보여주기 위해 마지막 두 열에서는 기금관리형 공공기관을 제외한 샘플만을 활용하여 앞선 두 열과 동일한 분석을 시행하였다.

기금관리형 공공기관의 재무 정보는 이들이 운용하는 기금으로 인해 그 액수가 과대 평가될 가능성이 있으며, 또한 기금형 공공기관의 실질적인 생산활동은 전무하다고 볼 수 있기 때문에 재무정보가 왜곡될 수 있는 우려가 있다.

이러한 이유 때문에 선행연구(예컨대, 하세정·오영민·라영재, 2014)에서는 기금관리형 공공기관을 실증분석에서 제외하기도 하였다. 본 연구에서는 기금관리형 공공기관에 대해서 기금계정을 제외한 고유계정에 대한 재무정보만을 고려하였기 때문에 재무정보의 왜곡이 최소화되었을 것이라 예상된다. 그럼에도 기금관리형 공공기관을 분석에 포함하는 것으로 인해 문제가 발생할 수 있는 가능성을 고려하여, 기금관리형 공공기관을 제외한 샘플을 활용한 실증분석을 추가적으로 제시하였다.

〈표 V-13〉과 〈표 V-14〉는 전기의 경영결과(전기의 경영평가 표준화 순위 혹은 평가등급 더미)가 공공기관의 주요 재무변수(부채/자본 비율, ROA, 정규직 복리후생비, 정규직 평균보수액)에 미치는 효과를 전체 표본과 기금관리형 공공기관을 제외한 표본을 대상으로 추정한 결과를 보여준다. 여기서 주의해야 할 점은 각각의 재무정보가 서로 관계를 가지며 동시에 생성되기 때문에 내생성으로 인한 동시 편의(simultaneity bias)의 문제가 발생할 수 있다는 것이다. 본 분석에서는 이와 같은 내생성을 고려하여 충분히 외생적이라고 판단되는 변수들만을 독립변수에 포함하였다.

이제 이와 같은 실증분석의 결과를 제IV장에 제시된 검증 명제별로 살펴 보도록 하자. 우선 원활한 논의를 위해서 재무정보를 활용한 실증분석 결과를 통해 경영평가 설계의 적절성을 평가할 수 있는 검증 명제를 다시 나열하면 다음과 같다(**검증 명제 4**는 제외).

검증 명제

1. 부채비율이 높은 기관이 경영평가에서 불이익을 받을 것이다.
2. 이번 기에 경영평가 점수가 좋지 않은 공기업의 경우 다음 기 부채 비율이 개선될 것이다.
- 2-1. 복리후생비를 과다하게 지출하는 기관일수록 경영평가에서 불이익을 받을 것이다.
3. 경영결과가 해당 기업의 공공성이나 효율성을 높이기 위한 노력과 무관한 요인에 의해 영향을 받아서는 안 된다.

〈표 V-12〉 공공기관의 재무정보가 경영평가 결과에 미치는 효과

	전체 표본		기금관리형 제외 표본	
	표준화 순위 (1)	평가등급 (2)	표준화 순위 (3)	평가등급 (4)
부채/자본 비율	0.0019 (0.0102)	-0.0001 (0.0006)	0.0084 (0.0107)	0.0002 (0.0006)
ROA(당기순이익/자산)	31.6870 (15.7616)**	1.6931 (1.2237)	46.7246 (18.4782)**	3.2808 (1.3632)**
자산(조원)	0.6858 (0.4075)*	0.0099 (0.0237)	0.6996 (0.4079)*	0.0105 (0.0238)
기관 업력	0.1246 (0.0932)	0.0100 (0.0059)*	0.1328 (0.0961)	0.0107 (0.0061)*
정규직 상시근로자 수(천명)	-0.4721 (0.4946)	0.0079 (0.0309)	-0.4569 (0.4978)	0.0087 (0.0309)
정규직 1인당 복리후생비(백만원)	0.0426 (0.8603)	-0.0315 (0.0564)	-0.0963 (0.8689)	-0.0501 (0.0570)
정규직 1인당 평균보수(백만원)	0.4823 (0.1608)***	0.0423 (0.0116)***	0.4311 (0.1691)**	0.0377 (0.0121)***
정규직 평균근속년수	1.9999 (1.5088)	0.1891 (0.1086)*	2.3176 (1.5147)	0.2071 (0.1088)*
(정규직 평균근속년수) ²	-0.1201 (0.0650)*	-0.0112 (0.0046)**	-0.1318 (0.0654)**	-0.0119 (0.0046)**
기관장 출신: 정치인	-3.7461 (9.1066)	-0.3269 (0.5678)	-3.9216 (9.1282)	-0.3396 (0.5685)
기관장 출신: 관료	9.2688 (3.6205)**	0.4619 (0.2371)*	8.1155 (3.7197)**	0.3927 (0.2435)

〈표 V-12〉 계속

	전체 표본		기금관리형 제외 표본	
	표준화 순위 (1)	평가등급 (2)	표준화 순위 (3)	평가등급 (4)
공공기관 유형 더미	YES	YES	YES	YES
연도별 시간 고정효과	YES	YES	YES	YES
R^2 / LR χ^2	0.11	54.25	0.11	49.28
N	289	289	269	269

주: 1. 괄호 안의 값은 표준오차를 의미.

2. *, **, ***는 각각 10%, 5%, 1%의 유의수준을 의미.

자료: 저자 작성.

〈표 V-13〉 경영평가 결과가 주요 재무정보에 미치는 효과(전체 표본)

	부채/자본 비율		ROA		정규직 1인당 복리후생비		정규직 1인당 평균보수	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
표준화 순위 _{t-1}	0.0131 (0.2115)		0.0005** (0.0003)		0.0014 (0.0044)		0.0973*** (0.0259)	
경영평가 A등급 더미 _{t-1}		-7.2873 (33.5689)		0.0446 (0.0398)		-0.0875 (0.7206)		15.4277*** (4.2058)
경영평가 B등급 더미 _{t-1}		2.2850 (31.6113)		0.0194 (0.0373)		-0.3233 (0.6790)		9.7024** (3.9630)
경영평가 C등급 더미 _{t-1}		-5.1297 (32.0488)		0.0060 (0.0379)		-0.0440 (0.6876)		9.2905** (4.0132)
경영평가 D등급 더미 _{t-1}		-9.6596 (36.5716)		-0.0028 (0.0434)		-0.5352 (0.7810)		3.5441 (4.5582)
기관 업력	-0.3464 (0.3059)	-0.3601 (0.3094)	0.0002 (0.0004)	0.0002 (0.0004)	0.0107* (0.0065)	0.0112* (0.0065)	-0.0151 (0.0385)	-0.0072 (0.0382)
기관장 출신: 정치인	-6.6384 (35.8361)	-3.9366 (36.2797)	-0.0323 (0.0422)	-0.0348 (0.0428)	-0.0291 (0.6686)	-0.0473 (0.6727)	-1.9269 (3.9717)	-2.5420 (3.9262)
기관장 출신: 관료	11.1788 (12.6913)	9.6762 (13.0376)	-0.0035 (0.0151)	-0.0007 (0.0156)	0.0962 (0.2636)	0.1121 (0.2697)	-3.9099** (1.5656)	-3.3445** (1.5742)
공공기관 유형 더미	YES	YES	YES	YES	YES	YES	YES	YES
연도별 시간 고정효과	YES	YES	YES	YES	YES	YES	YES	YES
R^2	0.01	0.02	0.06	0.06	0.03	0.04	0.29	0.32
N	191	191	186	186	207	207	207	207

주: 1. 괄호 안의 값은 표준오차를 의미.

2. *, **, ***는 각각 10%, 5%, 1%의 유의수준을 의미.

3. 경영평가 등급더미는 E등급이 기준등급(base).

자료: 저자 작성.

〈표 V-14〉 경영평가 결과가 주요 재무정보에 미치는 효과
(기금관리형 제외 표본)

	부채/자본 비율		ROA		정규직 1인당 복리후생비		정규직 1인당 평균보수	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
표준화 순위 _{t-1}	0.0015 (0.2226)		0.0005 (0.0003)*		0.0014 (0.0048)		0.0881 (0.0281)***	
경영평가 A등급 더미 _{t-1}		-9.1824 (34.9607)		0.0500 (0.0407)		-0.1746 (0.7614)		14.3889 (4.3217)***
경영평가 B등급 더미 _{t-1}		2.3762 (32.7117)		0.0150 (0.0379)		-0.2437 (0.7124)		9.9379 (4.0437)**
경영평가 C등급 더미 _{t-1}		-4.9988 (33.1267)		0.0084 (0.0384)		-0.0856 (0.7215)		9.6619 (4.0950)**
경영평가 D등급 더미 _{t-1}		-10.0149 (37.7553)		-0.0024 (0.0440)		-0.4948 (0.8223)		3.9387 (4.6672)
기관 압력	-0.3475 (0.3202)	-0.3567 (0.3239)	0.0002 (0.0004)	0.0002 (0.0004)	0.0113 (0.0070)	0.0114 (0.0071)	-0.0213 (0.0404)	-0.0153 (0.0400)
기관장 출신: 정치인	-6.6321 (36.9936)	-3.9262 (37.4575)	-0.0325 (0.0428)	-0.0355 (0.0434)	-0.3174 (0.8059)	-0.3367 (0.8158)	-4.7126 (4.6622)	-5.4401 (4.6304)
기관장 출신: 관료	12.0777 (13.3794)	10.0516 (13.8267)	-0.0011 (0.0157)	0.0036 (0.0163)	0.2092 (0.2915)	0.2070 (0.3011)	-3.7497 (1.6862)**	-3.2248 (1.7092)*
공공기관 유형 더미	YES	YES	YES	YES	YES	YES	YES	YES
연도별 시간 고정효과	YES	YES	YES	YES	YES	YES	YES	YES
R^2	0.01	0.02	0.04	0.05	0.04	0.04	0.29	0.32
N	179	179	174	174	179	179	179	179

주: 1. 괄호 안의 값은 표준오차를 의미.

2. *, **, ***는 각각 10%, 5%, 1%의 유의수준을 의미.

3. 경영평가 등급더미는 E등급이 기준등급(base).

자료: 저자 작성.

처음 세 가지 검증 명제는 경영평가제도가 도덕적 해이의 통제장치로서 기능할 수 있는지를 평가하며, 네 번째 검증 명제는 평가지표의 적절성 여부를 평가한다.

먼저 우리는 경평결과와 부채비율과의 관계를 살펴보는 것을 통해 **검증 명제 1**과 **검증 명제 2**를 검정할 수 있다. 〈표 V-12〉를 보면, 부채/자본 비율의 계수는 모두가 통계적으로 유의하지 않은 것으로 나타났으며, 이는 경

평 순위나 평가등급이 부채비율에 영향을 받지 않는다는 것을 의미한다. 또한 <표 V-13>과 <표 V-14>의 (1)열과 (2)열은 직전 연도의 표준화 순위 혹은 경영등급이 현재의 부채/자본 비율의 변화에 영향을 주지 않는다는 것을 보여준다.

이와 같은 결과는 경영평가제도의 설계가 **검증 명제 1**과 **검증 명제 2**에서 제시한 가설과는 무관하다는 것을 보여주고 있다. 즉 공공기관의 위험추구(risk taking) 행위로 인해 나타날 수 있는 도덕적 해이의 문제가 경영평가에 적절히 반영되지 않았으며, 경영평가가 차후의 도덕적 해이의 문제를 규율할 수 있는 기제로서의 역할도 수행하지 못하고 있음을 시사한다.

부채비율이 도덕적 해이의 규율 기제로 경영평가에 반영되지 않은 것에는 몇 가지 이유가 있는 것으로 판단된다. 그중 가장 큰 이유는 공공기관 부채의 변화가 외부적인 요인들에 의해 좌우되는 경향이 존재한다는 것이다.

지난 2011년 기획재정부에서 LH 공사의 국민주택기금 차입금, 한국수자원 공사의 4대강, 아라뱃길 단지 사업과 관련된 부채의 경우 이를 ‘정부정책 이행 등을 위해 해당 기관에서 통제 불가능한 사유로 발생한 부채로 정의’하면서 이를 부채비율 평가 시 제외하였다(갑사원, 2013). 이러한 평가기준의 조정이 갖는 의미는 부채비율이 경영평가상의 재무 건전성을 대표하는 의미 있는 정보를 제공해주지 못하고 있다는 것을 의미한다.

즉 앞서 설명한 모형의 맥락에서 다시 설명하면, 부채비율은 경영성과에 대한 충분한 정보를 담고 있다기보다는 측정오차로 인한 잡음을 더욱 많이 반영하고 있다는 것을 뜻한다. 이처럼 부채비율이 외부적인 요인에 의해서 결정될 때에는 도덕적 해이의 문제를 부채비율을 통해 식별하는 것이 매우 어렵다. 이러한 상황에서 부채비율과 경영결과가 무관한 것으로 나타난 결과는 어쩌면 매우 자연스러운 현상이 아닐 수 없다.

반면 수익성을 대표하는 변수인 ROA(당기순이익/자산)의 경우에는 경영평가의 순위나 평가등급에 통계적으로 유의미한 양의 효과(표 V-12 참고)가 있으며, 직전 연도의 경영평가의 결과가 ROA에 미치는 평균적인 영향(표 V-13과 표 V-14의 (3)열) 또한 어느 정도 통계적으로 유의한 것으로

나타났다. 이러한 결과는 경평제도가 공공기관의 수익성에 대한 정보를 어느 정도 평가에 반영하고 있다는 것을 의미하며, 이는 공공기관이 수익성 향상을 통해 경평결과를 개선하기 위해 노력하고 있는 것으로 해석할 수 있다.

다음으로 도덕적 해이의 또 다른 지표로 본 연구에서 고려하고 있는 복리후생비가 경영평가 결과와 어떠한 관계를 형성하고 있는가를 검토하는 것을 통해 **검증 명제 2-1**을 다루어 보자. 앞서 제IV장에서 언급하였듯이, 정부의 (암묵적) 보증으로 인해 파산의 위험이 사라진 경우에 공공기관들은 모든 소비 지출을 최적 수준보다 높일 유인이 존재한다. 따라서 공공기관의 도덕적 해이가 팽배하다면 인건비, 그중 복리후생비의 지출이 증가할 것으로 기대된다. 그러나 경평제도가 이를 규율할 수 있는 방식으로 설계되었다면, 복리후생비 지출이 경평결과와 부(-)의 관계에 있을 것으로 예상할 수 있다.

〈표 V-12〉 ~ 〈표 V-14〉의 결과는 복리후생비 지출과 경평의 관계가 앞선 부채비율의 결과와 동일하게 아무런 관련이 없다는 것을 보여준다. 즉 복리후생비가 경평결과에 영향을 미치지 못하며, 또한 직전 연도의 경평결과가 현재의 복리후생비 지출에 영향을 미치지 못한 것으로 나타났다. 결론적으로 현재의 경평제도에서 복리후생비 또한 공공기관의 도덕적 해이에 대한 유의미한 정보를 담고 있지 못한 것으로 평가할 수 있다.

본 연구에서는 정규직 1인당 평균보수액(혹은 연봉)이 근로자에게 주어지는 혜택이라는 점에서 복리후생비와 유사한 성격을 띠고 있으나, 기관의 도덕적 해이를 가늠할 수 있는 지표로 간주하지는 않았다. 이는 평균보수액이 근로성과나 근로자의 생산성을 반영한다는 점에서 복리후생비와 성격을 달리 하기 때문이다. 따라서 1인당 평균보수액이 경평결과에 미치는 효과는 앞선 논의와 달리 근로자의 생산성의 측면에서 논의하는 것이 타당할 것으로 생각된다.

〈표 V-12〉의 결과는 1인당 평균보수액이 경평결과에 매우 유의한 양의 효과가 있는 것을 보여 준다. 이는 근로자의 근로 능력이 평균보수액에 반영되어 있다고 가정할 때 매우 자연스러운 결과로 보인다. 또한 〈표 V-13〉과 〈표 V-14〉의 마지막 두 열은 직전 연도의 경평순위나 평가등급이 현재

의 임금에 매우 유의한 양의 효과가 있다는 것을 보여 주고 있다. 이와 같은 결과 또한 경평결과가 근로자의 성과급에 반영되기 때문에 자연스러운 현상으로 볼 수 있다.

이제 마지막으로 **검증 명제 3**를 검정하기 위한 실증분석 결과에 대해 살펴보자. 앞선 검증 명제들은 도덕적 해이의 통제 장치라는 관점에서 경영평가를 바라보고 분석을 하였지만, **검증 명제 3**에서는 경영평가지표 자체의 적절성에 대한 부분을 논하고자 한다.

경평제도는 공공기관의 경영효율성을 평가하여 공공기관의 경영 성과를 증진시키는 데 목적이 있으므로, 경영평가 결과가 기관의 내적가치(intrinsic value)나 노력을 충분히 반영하고 있어야 할 것으로 생각된다. 따라서 경평 결과는 기관의 경영효율성을 대표하고 있는 재무정보에 의해서 결정되는 것이 바람직한 것으로 볼 수 있다. 반면 경평결과가 공공기관의 내적가치나 노력에 의해 결정되기보다는 이와 무관한 요인들, 특히 경영평가를 대비하기 위해 투입한 인적·물적 자원에 의해 결정된다면, 기관들은 경영효율성 증진을 위한 노력을 줄일 것으로 예상할 수 있다.

기업의 내적가치와 무관하지만 경평결과에 영향을 미칠 수 있는 외적 요인은 매우 다양하게 존재한다. 예컨대, 선행연구에서 지적한 것처럼 공공기관이 경영평가를 대처할 때 상당한 양의 인적·물적 자원을 사용하기 때문에, 크기가 큰 기관일수록 이들의 조직 내에 경영평가에 대한 대처를 전담하는 인력을 동원하여 더 좋은 결과를 받을 수 있는 가능성이 있다. 또한 오래된 기관일수록 신생 기관들보다는 그 동안의 경영평가에 대한 노하우를 축적하고 있기 때문에 경영평가에 더욱 잘 대처할 가능성이 존재한다.

따라서 본 연구에서는 경영효율성과 직접적인 관련은 적으나 경영평가 결과에 영향을 미칠 수 있는 변수들로, 기관의 크기(자산 혹은 상시종사자수)와 업력 등을 고려하여 실증분석을 시도하였다.¹¹⁾

11) 기업의 내적가치와 무관하지만 경영평가 결과에 영향을 줄 수 있는 비본질적 요인은 기업의 능력과 무관하기 때문에 대체로 일시적·불연속적 경향이 있으며, 그렇기 때문에 등급의 변동성에 큰 영향을 미칠 것으로 예상된다. 반면 업력과 기관의 크기는 비본질적 요인임에도 불구하고 지속적이고 연속적인 성격을 갖기 때문에 등급의 변동성의 증

먼저 <표 V-12>에서 기업의 크기를 대변하는 변수인 자산과 상시종사자 수가 경영평가 결과에 미치는 영향을 살펴보면, 유의수준은 높지는 않았으나 자산크기가 경영평가 순위에 긍정적으로 영향을 미치는 것으로 나타났다((1)결과 (3)결과). 반면 정규직 상시종사자 수의 계수는 통계적으로 유의하지 않아, 정규직 근로자 수가 많을수록 경영평가가 유리하다는 가설은 기각되었다. 또 다른 외적 요인에 대한 변수로 기관의 업력은 유의수준은 높지 않으나 경영평가 등급과 양(+)의 관계에 있다는 것을 알 수 있다((2)결과 (4)결과). 이러한 결과는 기관의 설립연도가 오래될수록 경평결과가 높다는 것을 의미한다.

이와 같은 실증분석 결과는 뚜렷하지는 않지만 경평결과가 기관의 내적가치나 노력과 무관한 요인들에 어느 정도 영향을 받고 있다는 것을 지지하고 있다(**검증 명제 3**과는 상반되는 결과). 즉 오래 전에 설립된 공공기관일수록 또한 기관의 규모가 클수록, 변경이 잦은 경평 기준에 잘 대처하여 상대적으로 좋은 결과를 받은 것으로 해석된다.

이 밖에도 <표 V-12>에서는 경평결과에 미칠 수 있는 다양한 외적 변수 혹은 요인에 대해서도 고려하고 있는데, 이는 상당히 흥미로운 결과를 보여 준다.

먼저 근로자의 평균근속년수의 경우 제공항이 상당히 유의한 수준에서 음(-)의 계수를 띠고 있는 것을 알 수 있는데, 이는 경영평가의 결과가 근로자의 평균근속년수와 역 U자형의 관계가 있다는 것을 보여준다. 추정된 계수들을 토대로 계산해 본 결과, 평균근속년수와 경영평가 결과의 관계가 평균근속년수 약 8.3~8.7년을 기점으로 양(+)에서 음(-)으로 전환된 것으로 나타났다. 이와 같은 결과는 기관 내에 근로자의 구성원의 나이가 대체로 적을수록 혹은 많을수록 경영평가 결과가 좋지 않다는 것을 의미한다.

가에 영향을 적게 미칠 것으로 예상된다. 따라서 업력과 기관의 크기가 경영평가에 미치는 결과가 비본질적 요인이 등급의 변동성에 미치는 효과와 일견 상충된 것으로 보이나, 이는 별개의 효과로 보는 것이 타당한 것으로 판단된다. 즉 업력과 기관의 크기는 지속적으로 경영평가에 미치는 외적 요인으로 분류하고, 그 외에 대부분은 관찰되지는 않지만 단기적·불연속적으로 경영평가에 미치며 등급의 변동성에 영향을 미치는 외적 요인으로 분류함이 타당할 것으로 사료된다.

다음으로 기관장 출신과 경영평가 결과와의 관계를 살펴보면, 우리는 기관장 출신이 관료를 뜻하는 더미 변수가 경영결과에 상당히 유의한 수준의 양의 효과를 가진다는 것을 알 수 있다. 즉 기관장 출신이 관료인 경우 상대적으로 경영결과가 좋다는 것이다. 하지만 이러한 결과에 대한 해석은 매우 조심스럽다. 아주 긍정적으로 해석한다면 관료출신 기관장의 경우 기존의 조직생활의 경험으로 인해 경영관리능력이 상대적으로 뛰어나기 때문에 경영결과가 좋게 나타난 것으로 볼 수 있다. 또 다른 해석으로는 관료 출신의 경우 공공기관에 대한 정보가 상대적으로 풍부하기 때문에 경영성과가 더욱 좋을 것으로 예상되는 공공기관의 기관장을 역임할 가능성 또한 존재한다. 이유는 명확하지는 않지만 기관장 출신은 경영평가와 뚜렷한 관련이 있는 것으로 보인다.

Ⅵ. 결론 및 정책적 시사점

1. 결론 요약

본 연구에서는 다음의 명제를 통해 현행 경평제도가 도입 취지에 부합하는 적절한 역할을 수행하고 있는지를 살펴보았다.

1. 도덕적 해이 문제가 심각한 기관(예컨대 부채비율나 복리후생비 지출 수준이 높은 기관)이 경영평가에서 불이익을 받아야 한다.
2. 이번 기에 경영평가 점수가 좋지 않은 공공기관의 경우 다음 기에 도덕적 해이 문제를 해결하기 위해 더 많은 노력(예컨대 부채비율이나 복리후생비 개선)을 해야 한다.
3. 경평결과가 효율성이나 공공성 제고를 위한 해당 기관의 노력과 무관한 요인(기관의 크기나 업력 등)에 의해 결정되어서는 안 된다.
4. 경평결과가 효율성이나 공공성을 개선하기 위한 기관의 노력에 대한 정확한 정보를 담고 있지 않은 항목(리더십·책임경영 범주)에 의해 크게 영향을 받아서는 안 된다.

첫 번째 명제를 검증하기 위하여 도덕적 해이의 기준을 부채비율과 복리후생비로 설정하였다. 해당 공공기관의 부채비율이 높거나 복리후생비의 지출이 높다면 그 기관에 도덕적 해이의 문제가 있다고 보는 것이다. 분석 결과 부채비율이 높은 기관이 그렇지 않은 기관에 비해서 경영평가 결과가 좋지 않게 나오거나 복리후생비의 지출이 높다고 하여 경영평가에서 불이익을 받는 것은 아닌 것으로 확인되었다. 두 번째 명제 또한 충족하지 못하는 것으로 확인되었다. 이번 기에 경영평가 결과가 좋지 않았다고 해서 다음 기에 부채비율이 낮아지는 현상은 발견할 수 없었다.

세 번째 명제를 검증하기 위하여 기관의 크기는 기관의 자산의 크기로 측정하였다. 그 결과 자산의 크기가 큰 공공기관일수록 소규모 공공기관보다 평가를 더 잘 받는 것으로 나타났다. 또한 해당기관이 오래되었을수록, 즉 업력이 쌓일수록 평가에서 좋은 결과를 받았다. 경영평가의 결과가 기관의 설립 목적을 달성하기 위한 노력과 상관없는 요인들의 영향을 받는다는 것이다. 또한 경영평가 순위의 변동성을 살펴본 결과 경평결과 기준으로 상·중·하의 어느 특정 그룹에 계속해서 머물 확률은 45%에 불과했고, 과거 자신이 머물렀던 그룹에서 벗어나 다른 그룹으로 이탈할 확률은 55%나 되었다. 특히 전년도에 상위그룹에 있다가 다음 연도에 하위그룹으로 떨어지거나 반대로 하위그룹에 있다가 상위그룹으로 올라갈 확률이 13%가 되었는데 이는 100개의 기관이 있다면 13개의 기관은 1년이라는 단기간 안에 큰 폭의 순위변동을 경험했다는 것을 의미한다.

네 번째 명제의 분석결과, 경영평가 결과가 리더십 및 책임 경영 지표의 영향을 많이 받는 것으로 나타났다. 이는 공공기관의 노력에 대한 제대로 된 정보를 담고 있다고 보기 어려운 지표가 경영평가 결과의 순위에 가장 큰 영향을 주는 것을 의미한다. 지금까지 살펴본 결과 경평제도는 우리가 제도의 최적 설계와 관련된 4가지 명제를 모두 다 충족시키지 못하는 것으로 나타났다.

2. 정책적 시사점 및 제도 개선방향

본 절에서는 앞서 분석한 결과에 근거하여 경평제도의 개선방향을 제안해 본다.

가. 경영평가제도의 운영목적에 대한 제안

첫 번째 개선방향은 경평제도의 본원적 목적을 최대한 간결하게 하자는 것이다. 앞서 설명한 것처럼 현행 제도는 공공기관으로 하여금 변화하는 경영환경에 대한 신속한 대처를 어렵게 만드는 비효율적 방식으로 설계되어

있다. 이러한 비효율성으로 인해 경평제도의 유인계약적 성격은 크게 약화될 수밖에 없다. 현행 제도하에서 경평성과급은 노동시장에서 최고 경영진의 스톡옵션 계약과 같은 역할을 할 수 없다. 경평제도의 유인계약적 성격을 강화하기 위해서는 경평제도의 전면적 손질이 불가피하다. 문제는 지난 30년 이상 운영되어 온 경평제도의 틀을 바꾸는 것은 많은 시간과 노력을 필요로 한다는 것이다.

따라서 정부가 현행 제도의 평가틀을 그대로 유지한 채 운영해나가고 싶다면 이를 공공기관의 도덕적 해이에 대한 통제 장치로 목적을 분명히 한 뒤 운영하는 것이 바람직하다. 예컨대 경평제도의 평가지표를 부채비율이나 복리후생비와 같은 몇 가지 객관적 재무지표로 한정된 뒤 그 수준에 따라 기관에게 페널티를 부과하는 것이다. 기관의 부채비율이나 복리후생비 같은 운영비용의 상승을 방지할 경우 이후 국가재정에도 상당한 부담 요인이 될 수 있다는 점에서 이러한 규제의 필요성은 충분히 인정된다. 공공기관의 경영효율성 향상은 도덕적 해이에서 기인하는 비효율을 줄임으로써 이루어지는 것이다.

이러한 개선안의 가장 큰 장점은 현행 제도의 틀을 유지하는 가운데 그 평가지표들을 간결하게 만듦으로써 새로운 성과 평가제도 도입에 따른 사회적 비용을 줄일 수 있다는 점이다. 또한 평가대상자인 공공기관으로 하여금 이들이 성취해야 할 목표를 분명히 함으로써 제도에 대한 순응성을 높여준다는 데에 있다. 이러한 개선방향은 최근 경평의 운영방향을 부채관리와 방만경영 평가 강화에 놓는 정부의 정책기조와도 일치한다는 점에서 장점이 있다.

나. 경영평가지표 설계의 문제

두 번째 개선안은 평가지표의 설계와 관련된 이슈이다. 만약 정책결정자가 경평을 통해 기관의 성과향상 노력을 제고하기를 원한다면 현행 제도에 대한 대폭적 손질은 불가피하다. 계약이론에 따르면 유인계약이 효과적으로

작동하기 위해 필요한 가장 중요한 전제조건 중 하나는 성과 지표들이 대리인의 노력 수준의 변화를 제대로 반영해야 한다는 것이다. 만약 경평결과가 공공기관의 노력과 무관한 요인에 의해 영향을 받는다면 그만큼 경평의 유인계약으로서의 기능은 약화될 수밖에 없다.

따라서 경평이 사회적으로 바람직한 수준의 노력을 이끌어내기 위한 기제로 사용되기 위해서는 무엇보다도 그 성과지표를 제대로 정해야 한다. 이는 공공기관의 인력이나 자산 크기, 그리고 업력처럼 경평에 대한 단기적 대응 능력과 관련이 있는 요인들이 경평결과에 영향을 주어서는 안 된다는 것을 의미한다. 이러한 현상이 발생하는 것을 막기 위해서는 무엇보다도 기관들의 평가부담을 줄여줄 필요가 있다.

공공기관경영연구원(2015)의 경우 평가부담을 덜어주기 위해 단기적으로 ▲ 경영실적 보고서 분량 축소 및 간편 표준양식 제시 ▲ 집중실사 및 합동 간담회 진행 ▲ 지표체계 및 지표 수 간소화라는 대안을 내놓고 있으며 라영재·윤태범(2013)은 평가 주기를 2~3년으로 재조정할 것을 제안하고 있다.

본 연구 역시 평가부담을 줄이기 위해 우선 평가지표 수를 대폭 축소하고 지표의 세부 평가 내용도 최대한 단순화해 만들 것을 제안한다. 예를 들어 정부권장정책의 경우를 보자. 이는 해당 기관의 공공성의 정도를 보여주는 지표라기보다는 특정 부처의 과제의 이행을 담보하기 위해 사용되는 공공기관의 관리 통제수단(공공기관경영연구원, 2015)의 성격이 더 강하다. 해당 기관의 공공성 측정 지표는 기관의 고유 목적과 직접 관련된 주요사업의 성과로 한정하고 일반적인 공공기관 관리정책 관련 사안들은 경평지표에서 제외해야 한다.

평가지표의 축소 및 간소화와 더불어 지표 설계에 있어 유념해야 할 것은 평가지표가 기관의 노력 수준에 대해 ‘왜곡없는’ 정보를 제공해주어야 한다는 것이다. 예컨대 김지영(2011)은 많은 공기업에서 평가 등급이 실제 생산성의 움직임을 정확히 반영하지 못하고 경우에 따라 오히려 변동 방향이 반대로 가는 현상을 발견했다. 그는 이러한 현상이 발생하는 이유로 경평의 지표들이 실제 생산성 증가율이 아니라 기업별로 설정된 목표치를 바탕으로

산정되기 때문임을 지적했다. 이처럼 생산성은 하락하였는데 평가순위가 상승한 사례들이 많이 발생할수록 평가제도의 신뢰성은 허물어진다. 예컨대 기관의 경평성과급이 재무지표상의 결과보다 경평결과에 더 의존하게 됨에 따라 일반인들은 그 성과급이 ‘나눠먹기식’으로 배분되고 있다는 생각을 갖게 된다.

혹자는 공공기관의 성과가 그 기관이 통제할 수 없는 위험에 의해 결정되는 것은 불공평하기 때문에 ‘조정된 재무정보’나 실적치를 이용하여 이를 조정해줄 필요가 있다고 항변한다. 실제 경평 과정에서 공공기관의 부채비율을 산정할 때 ‘4대강 사업’이나 해외자원개발 사업과 같은 주요 정부정책이 행 과정에서 발생한 부채는 해당 기관에 의해 통제가능하지 않은 사유로 인해 발생한 부채로 보아 이를 제외한 적이 있다.

그러나 최적 계약이론 관점에서 이러한 주장은 문제가 있다. 만약 기관이 성과급(s)을 받고 그 성과급이 기본적으로 그 기관의 재무지표에 의존하고 있다($s(q_2)$)고 하자. 그리고 관찰되는 재무지표(q_2)에는 효율성 제고를 위한 기관의 노력(a_2)과 그 기관이 통제할 수 없는 위험(ε_2)에 대한 정보가 모두 담겨 있다고 하자.

$$\begin{aligned} s &= s(q_2) \\ q_2 &= a_2 + \varepsilon_2, \end{aligned} \quad (\text{식 VI-1})$$

위의 수식을 통해 ‘조정된 재무정보’를 사용하는 것은 다음과 같이 설명할 수 있다. 통제할 수 없는 위험이 매우 크다는 것은 ε_2 값이 아주 크기 때문에 q_2 를 그대로 사용하여 성과보수의 수준을 결정해서는 안 된다는 것이다. 이 경우 q_2 에 따라 성과급(s)을 지급할 경우 이끌어낼 수 있는 노력 수준(a_2)은 통제할 수 없는 위험 수준이 그리 높지 않은 경우에 비교했을 때 낮아지기 때문이다. 문제는 이것을 해결하는 방법이다. 현행 경평제도는 실제 재무지표(q_2) 대신 ‘조정된 재무정보’ $\tilde{q}_2$ 를 사용하고 있는 셈이다. 문제는 $\tilde{q}_2$ 역시 자의적 기준에 따라 작성되기 때문에 $\tilde{q}_2$ 역시 a_2 를 추론하는 데

필요한 정보를 제공해주지 못한다는 것이다. 결국 이러한 문제를 해결하는 방법은 무엇일까? 계약이론에 따르면 이는 조정된 재무정보를 사용하여 성과급($s(q^*)$)을 지급할 것이 아니라 공공기관의 보수에서 성과보수 비중을 낮추는 것이다.

다. 경영평가성과급 설계에 대해

앞에서 우리는 경영성과급이 리더십이나 책임경영 지표처럼 공공성이나 효율성을 개선하기 위한 해당 기관의 노력과 직접 관련도 없을 뿐만 아니라 그 측정 결과의 정확성도 담보하기 어려운 지표에 의해 가장 크게 영향을 받는다는 점을 보았다. 그리고 Holmström and Milgrom(1991) 이론의 결론에 따라 이러한 경우 전체 보수 중 경영성과급의 비중을 낮추는 것이 사회적으로 최적일 수 있다는 점 역시 앞에서 이미 설명했다.

이와 관련하여 몇 가지 사항들을 추가적으로 검토해보자. 먼저 경영성과급의 적정 비율과 격차의 문제이다. 현행 제도는 기관유형에 따라 월 기본급의 100%에서 300% 이내의 성과급을 받는다. 이에 대해 성과급의 크기가 클 뿐만 아니라 기관 간 격차가 과도하다는 견해가 많다. 라영재·윤택범(2013)이 준정부기관 종사자들을 상대로 실시한 여론조사에 따르면 기관 사이에 경영성과급의 규모 및 기관 간 격차를 개선해야 한다는 견해가 거의 50% 비중에 다다랐다. 경영성과급의 수준과 격차를 재조정할 필요가 있다.

이를 위해 Holmström and Milgrom(1991)이 말하는 것처럼 측정이 용이한 것과 측정이 불가능한 것을 나누어 각각의 주체에게 일을 맡기는 아이디어를 경영성과급의 설계에 적용할 필요가 있다. 즉 해당 기관의 활동이 평가하기 어려운 것이라면 아예 경영성과급을 폐지하고 기본급만을 주는 방식을 채택해야 한다. 또 반대로 경영보다 해당 기관의 활동을 평가할 수 있는 더 좋은 기제(예컨대 상장기업의 경우 주가)가 있다면 이들 기관 역시 경영 대상에서 제외하는 것이 맞다. 이는 자연스럽게 경영의 대상 기관을 재조정하는 결과를 낳을 것이다. 이들 기관에 대해서 경영을 도덕적 해이에 대한 통제장치로 전환하여 경영을 통해 이를 살펴보는 방식이 타당하다.

참고문헌

- 감사원, 『감사결과보고서: 공기업 재무 및 사업구조 관리실태』, 2013.
- 공공기관경영연구원, 『공공기관 경영평가제도 개선방향 연구』, 2015.
- 굴스비·레빗·시버슨, 『미시경제학』, 시그마프레스, 2014.
- 김빛마로, 「2016년 노벨 경제학상: 계약이론에의 공헌」, 『재정포럼』, 제245호, 한국조세재정연구원, 2016.
- 김완희, 『공공기관의 핵심성과 평가를 위한 계량지표 체계의 개선방안 연구』, 한국조세재정연구원, 2010.
- 김주찬·이민창 『공공기관 경영평가실적의 효과성에 대한 직무 담당자의 인식조사』, 2014.
- 김지영, 『공기업 생산성에 관한 연구』, 한국조세재정연구원, 2011.
- 기획재정부, 「공공기관 경영평가제도 대폭 개편」, 보도자료, 2010. 12. 29.
- _____, 『2012년도 공공기관 경영평가 편람』, 2011.
- _____, 「2015년도 공공기관 경영실적 평가결과 확정」, 2016a.
- _____, 「2016년도 공기업·준정부기관 예산집행지침」, 2016b.
- 라영재·윤태범, 『공공기관 경영평가제도 분석과 새로운 모형 개발』, 한국조세재정연구원, 2013.
- 박미정, 『공공기관 경영실적 평가제도의 현황과 과제』, 국회입법조사처, 2010.
- 이석원, 「조정성과지표를 사용한 공공기관의 성과측정: 정부투자기관경영평가에의 적용」, 『한국행정학보』, 제39권 제4호, 2005.
- 이오·유승현, 『공공기관 경영평가제도의 문제점 및 개선방안 연구 - '08년도 경영실적평가제도를 중심으로』, 감사연구원, 2010.
- 이원희·라영재, 『공공기관 경영평가 30년, 회고와 전망』, 한국조세재정연구

원, 2015.

이창길 · 최성락, 「공공기관 경영평가에서 상대평가의 오류가능성 분석」,
『행정논총』, 제48권 제1호, 2009.

장지인 · 곽채기 · 신완선 · 오철호, 『공공기관 경영평가제도의 변천과정 연구
Ⅱ』, 한국조세재정연구원, 2013.

하세정 · 오영민 · 라영재, 『지속가능한 공공기관 부채관리를 위한 정책적 대
응방안』, 한국조세재정연구원, 2014.

한국조세재정연구원, 『공공기관 경영평가제도 변천과정 연구』, 2010.

최한수 · 이창민, 『공기업 부채와 도덕적 해이: 정부의 암묵적 보증으로 발생
한 보조금 규모 추정』, 한국조세재정연구원, 2015.

Bolton, Patrick, and Mathias Dewatripont, *Contract theory*, MIT press,
2005.

Hölmstrom, Bengt. and Paul Milgrom, "Multitask principal-agent analyses:
Incentive contracts, asset ownership, and job design," *Journal of
Law, Economics & Organization*, 7(1991), pp. 24~52.

Wilson, James Q., *Bureaucracy: What government agencies do and why
they do it*, Basic Books, 1989.

〈웹사이트 자료〉

공공기관경영정보공개시스템, <http://www.alio.go.kr>(접속일: 2016. 10. 17).

『조선일보』, 「공기업 경영평가 점수 포기하더라도 인천공항 보안 비용 충분
히 쓸 것」, 2016. 2. 3, [http://news.chosun.com/site/data/html_dir/
2016/02/03/2016020300349.html](http://news.chosun.com/site/data/html_dir/2016/02/03/2016020300349.html)(접속일: 2016. 10. 17).

공공기관 경영실적 평가제도 실증 연구

최한수 · 홍우형

본 논문에서는 Holmström and Milgrom(1991)의 다중 임무하의 주인-대리인 간의 최적 계약이론 이용하여 현행 공공기관 경평제도 설계방식의 적정성을 검토하고 이의 개선방향을 제시할 것이다. 규범적 관점에서 경평제도 설계의 적정성을 판단할 기준은 다음과 같다.

- (1) 경평이 도덕적 해이의 통제장치로서 작동하고 있는가.
- (2) 경평결과가 효율성이나 공공성 제고를 위한 기관의 노력과 무관한 요인에 결정되지 않도록 설계되어 있는가.
- (3) 평가결과가 효율성이나 공공성 제고를 위한 기관의 노력에 대한 부정확한 정보를 담고 있는 항목에 의해 좌우되지 않도록 설계되어 있는가.

이러한 명제를 실증적으로 검증하기 위해 지난 2008~2013년도 사이에 약 100여 개의 공공기관을 상대로 진행된 경평결과를 이용하여 분석을 진행해 보았다. 그 주요 결과는 다음과 같다.

첫째, 현행 경평제도는 공공기관의 도덕적 해이를 규율하는 데 큰 역할을 하지 못하고 있다.

둘째, 평가결과가 해당 기관의 공공성이나 효율성 제고를 위한 노력과 상관없는 비본질적 요인에 의해 영향을 받는다.

셋째, 경평결과가 리더십이나 책임경영처럼 본질적으로 측정오차가 클 뿐

만 아니라 공공기관의 효율성이나 공공성을 위한 노력 수준에 대한 부정확한 정보를 담고 있는 요인에 의해 좌우되었다. 이는 경평성과급의 크기가 사실상 리더십이나 책임경영의 평가결과에 의해 결정된다는 점을 의미한다.

본 연구의 정책적 시사점은 다음과 같다.

첫째, 경평제도를 도덕적 해이의 통제장치로 사용하기 위해서는 그 평가지표를 부채비율이나 복리후생비로 한정하는 것이 타당하다.

둘째, 경평결과가 비본질적 요인에 의해 영향받는 상황을 개선하기 위해 평가 지표를 대폭 축소하고 평가내용도 최대한 단순해지도록 재설계해야 한다.

셋째, 현재처럼 경평성과급의 크기가 리더십이나 책임경영처럼 공공기관의 노력 수준에 대한 부정확한 정보를 제공해주는 지표에 의해 결정되는 상황이 지속된다면 경평성과급의 크기를 줄여나가거나 극단적인 경우 폐지하는 것이 사회적으로 최적일 수 있다.

A Study of the Management Evaluation System of Public Enterprises in Korea

Hansoo Choi · Woo-Hyung Hong

Using Holmström and Milgrom's optimal contract model under multiple tasks, we attempt to justify why we need the Management Evaluation System (MES) of public enterprises and provide an insight for how we make MES better. Three hypotheses about an optimal MES design are established and empirically tested. Employing 2008 to 2013's data about MES of public enterprises in Korea, we find that the current MES fails to achieve the first best outcomes: MES is unlikely to solve moral hazard of the public enterprises and induce them to undertake the optimal level of efforts. The main reason is that there are a large amount of measurement errors for MES results, which suggests the rankings of MES are uninformative of underlying values of public enterprises. To improve MES, we should make MES simpler and scale back the wage-MES performance sensitivity.

■ 저자약력

최한수

서울대학교 국제경제학과 졸업
미국 피츠버그대학교 경제학 박사
현, 한국조세재정연구원 부연구위원

홍우형

서강대학교 정치외교학과 졸업
서강대학교 경제학 석사
미국 워싱턴주립대학교 경제학 박사
현, 한국조세재정연구원 부연구위원

자료 수집 및 정리

권선정 한국조세재정연구원 연구원

연구보고서 16-17

공공기관 경영실적 평가제도 실증 연구

발행	행	2016년 12월 30일
저자	자	최한수·홍우형
발행인	인	박형수
발행처	처	한국조세재정연구원
주소	소	30147 세종특별자치시 시청대로 336
전화	화	(044)414-2114(代)
홈페이지	지	www.kipf.re.kr
등록	록	1993. 7. 15. 제2014-24호
정가	가	5,000원
조판 및 인쇄		(주)정인애드
I S B N		978-89-8191-854-5 93320