

신자유주의적 개혁의 이상과 현실

이 준 구

세계 여러 나라에서 공공부문의 효율성을 획기적으로 높이기 위한 개혁 프로그램의 실험이 활발하게 이루어지고 있다. 최근 들어 성과에 기초한 유인을 제공하는 신자유주의적 개혁방식이 많은 관심을 끌고 있다. 미국 주정부 차원에서 시작된 성과에 기초한 책임 정책, 그리고 이것의 연장선상에서 연방정부가 도입한 NCLB 정책은 이 신자유주의적 개혁 노력의 대표적 사례라고 할 수 있다. 이 교육개혁 프로그램이 미국 교육에 어떤 영향을 미쳤는지를 심도 있게 분석해 보면 과연 신자유주의적 개혁이 공공부문의 대대적 혁신을 가져올 수 있는지에 대한 분명한 답을 얻을 수 있다.

한 마디로 말해 강력한 유인을 제공해 교육개혁을 이루려는 시도는 약속된 결과를 가져다주는 데 실패했다. 성과 평가의 기초가 되는 시험의 성적은 향상된 것으로 나타났지만, 그것이 진정한 의미에서의 교육의 질 향상을 뜻하는 것은 아니었다. 성적의 향상은 성적 그 자체를 올리려는 학교와 교사의 노력이 가져온 결과라는 해석이 설득력을 얻고 있다. 그 과정에서 오히려 교육의 내용이 왜곡되는 부작용만 가져왔다고 볼 수 있다.

신자유주의적 방식에 의한 교육개혁 시도가 가져온 가장 비극적인 결과는 가장 깨끗해야 할 교육부문에 부정행위가 횡행하게 되었다는 점이다. 학생들의 부정행위를 적발해 엄격하게 다스려야 할 학교와 교사가 스스로 부정행위를 저지르는 어이없는 일이 벌어지게 된 것이다. 이는 강력한 유인의 제공이 가져다주는 긍정적 효과를 모두 합친 것을 상쇄하고도 남을 부정적 효과다.

미국의 교육개혁 실험에서 분명하게 드러난 것처럼, 강한 유인을 부여함으로써 공공부문을 개혁한다는 것은 생각하는 것만큼 좋은 방법이 아니다. 특히 유인의 제공이 마치 공공부문의 비효율성을 치료하는 만병통치약이나 되는 듯 과도한 기대를 거는 것은 결코 바람직하지 않다. 공공부문의 개혁에 오히려 더 효과를 발휘할 수 있는 방법은 공공의 이익을 위해 일한다는 자부심을 느끼게 만드는 것일지도 모른다.

주제어: 유인계약, 성과에 기초한 책임원칙, 시험맞춤교육

1. 머리말

공공부문의 개혁은 언제나 논의만 무성할 뿐 실제로 이루어진 바는 거의 없다고 해도 과언이 아니다. 기회 있을 때마다 공공부문을 획기적으로 개혁해야 한다는 말이 나오고 수많은 개혁방안이 제시되기도 했지만, 공공부문은 언제나 비효율과 비능률의 대명사처럼 여겨

지는 것이 보통이다. 이와 같은 사정은 비단 우리나라뿐 아니라 세계 어느 나라에서든 공통적으로 관찰할 수 있다. 어느 나라 국민이든 간에 자신의 정부가 효율적으로 운영되고 있다고 믿는 경우는 무척 드문 것이 사실이다.

사실 어떤 조직을 개혁해 그 효율성을 획기적으로 높이는 것은 민간부문의 경우에도 무척 어려운 과제다. 잘 살펴보면 민간부문의 기업들도 수없이 많은 비효율성의 요인들을 안고 있는 것을 발견하게 된다. 기업은 마치 효율성의 대명사라도 되는 양 인식되고 있지만, 실제로는 더욱 효율적으로 만들 수 있는 여지가 생각 밖으로 클 수 있다. 아무리 효율적으로 경영되고 있는 기업이라 할지라도 기본적으로 본인-대리인의 문제(principal-agent problem)에서 자유로울 수 없기 때문이다.

기업의 경우에는 이윤이라는 효율성 평가의 명확한 잣대가 존재하기 때문에 그래도 사정이 조금 나은 편이다. 공공부문의 경우에는 그 본질상 다양한 목표를 추구하고 있기 때문에 이윤 같은 하나의 잣대를 통해 그것의 효율성을 평가할 수 없다. 이와 같이 평가기준 그 자체가 모호하기 때문에 공공부문의 효율성을 높이는 일은 출발 지점부터 난관에 직면하게 된다. 현재 효율적으로 운영되고 있는지의 여부조차 분명하게 평가할 수 없는 상황에서 좀 더 효율적인 조직으로 탈바꿈하도록 만들기 위해 무엇을 해야 할지에 대한 답을 얻기는 더욱 어려울 수밖에 없다.

더군다나 공공부문의 경우에는 서로 다른 목표를 추구하고 있는 수많은 이해 당사자들이 관련되어 있어 개혁이 더욱 어려운 측면이 있다. 국민, 정치가, 관료의 이해관계가 서로 엇갈리고 있기 때문에 공공부문의 개혁에 대한 입장에서도 차이가 날 수밖에 없다. 겉으로 모든 집단이 공공부문의 효율성을 높여야 한다는 원론적인 입장에서 동의하고 있다 할지라도 구체적인 내용에서는 서로 다른 생각을 갖고 있을 가능성이 크다. 더군다나 어떤 집단이 공공연히 자신의 사사로운 이익을 추구하는 행태까지 보이게 되면 문제는 한층 더 복잡해진다.

기업은 경영의 효율성을 높이는 데 경쟁이라든가 유인계약(incentive contract) 같은 것을 활용할 수 있다. 그러나 공공부문의 경우에는 유효한 경쟁이 존재하지 않는 경우가 대부분일 뿐 아니라, 유인계약을 활용할 여지도 상대적으로 적다. 유인계약을 활용한다는 것은 성과에 기초한 유인(performance-based incentives)을 제공함으로써 대리인이 자발적으로 열심히 일하게 만든다는 것을 뜻한다. 그런데 이 정책이 좋은 성과를 거두기 위해서는 무엇보다 우선 성과라는 것이 무엇을 뜻하는지 분명하게 정의해 놓아야 하며, 그 성과를 객관적으로 정확하게 평가할 수 있는 기준도 마련되어 있어야 한다. 그러나 공공부문처럼 다양한 목표를 추구하는 조직의 경우에는 성과의 정의도 어렵고 그것을 정확하게 평가하는

것도 어렵다. 그렇기 때문에 유인계약을 활용할 여지가 기업에 비해 상대적으로 적을 수밖에 없는 것이다.

그럼에도 불구하고 성과에 기초한 유인을 통해 공공부문의 개혁을 추진하려는 움직임은 끊임없이 계속되어 왔다. 그것 말고는 공공부문을 획기적으로 개혁할 수 있는 다른 더 좋은 방법을 찾기가 힘들었던 사정도 이에 대한 관심을 높이는 한 이유가 되었을 것이라고 짐작한다. 성과에 기초한 유인은 1980년대 뉴질랜드, 오스트레일리아, 그리고 영국 등지에서 시작되어 다른 나라들로 확산되어 간 성과중심경영(performance-based management) 정책에서 핵심적 역할을 수행하고 있다.⁽¹⁾ 이 정책을 지지하는 사람들은 공공부문 비효율성의 원인을 명백한 기준이 없어 성과에 대한 책임을 물을 수 없다는 데서 찾았다. 그들은 정부 운영의 성과가 객관적으로 측정될 수 있으며, 그 측정결과를 활용함으로써 정부 운영의 효율성을 획기적으로 높일 수 있다고 믿었다.

1990년대에 들어오면서 미국에서도 이와 비슷한 방법으로 공공부문의 개혁을 추진하려는 움직임이 활기를 띠기 시작했다. 이와 같은 움직임은 1993년 제정된 Government Performance and Results Act로 구체화되기에 이르렀는데, 이 법률은 연방정부의 여러 기관들로 하여금 행정의 효율성을 증대시키기 위해 성과측정(performance measurement) 결과를 적극 활용하도록 촉구하고 있다. 이 법률의 제정과 동시에 성과에 대한 목표를 연방예산의 주요 지출항목과 연계시키는 성과주의예산제도(performance budgeting)가 본격적으로 실험되기 시작했다. 과거의 프로그램 예산제도와 영기준 예산제도를 대신해 성과중심예산제도가 효율성을 추구하는 새로운 예산제도로서 각광 받기 시작한 것이다.

그런데 이와 같은 성과중심경영이 실제로 공공부문의 효율성을 높이는 데 얼마나 기여했는지에 대해서는 명확한 결론을 내리기 어렵다. 이론적으로 보면 성과에 기초한 유인을 제공함으로써 관료들의 행동을 바람직한 방향으로 유도할 수 있다는 논리에 아무 문제가 없다. 그러나 현실에서 관료들이 주어진 유인에 대해 기대한 것과 똑같은 반응을 보이리라는 보장이 없다. 자신의 이익을 먼저 챙기려 하는 관료는 그 유인에 대해 전략적으로 반응할 가능성이 높는데, 실제로 그런 일이 발생한다면 성과에 기초한 유인은 공공부문의 효율성을 높이는 데 큰 도움이 되지 못한다. 그 동안 몇 나라에서 성과중심경영 정책을 실시한 경험에 비추어 보면 그와 같은 결과가 나타날 가능성이 상당히 크다는 느낌을 받는다.

사실 민간부문에서의 유인계약도 대리인의 전략적 반응 때문에 기대했던 효과를 거두기 어려운 경우가 많다. 유인계약이 효과를 발휘할 수 있는 여건의 측면에서 볼 때, 민간부문

(1) 이 정책은 결과를 중시한다는 뜻에서 결과중심경영(managing for results)정책이라고도 불린다. 성과중심경영정책의 기본적 성격과 구체적 적용사례에 관해서는 Wholey(1999)를 참조하라.

이 공공부문보다 훨씬 더 유리한 상황에 있음은 의문의 여지가 없다. 그러나 민간부문에서의 유인계약도 그 효과를 제약하는 여러 가지 종류의 어려움에 직면하게 되는 것이다. 따라서 성과중심경영이 공공부문의 효율성을 획기적으로 높여주는 특효약이 되리라고 기대하는 것은 무리가 아닐 수 없다. 현실적으로 보아 공공부문에서 성과에 기초한 유인을 제공함으로써 관료들의 행동을 바람직한 방향으로 유도하는 것은 아주 어려운 과제가 될 가능성이 크다.

이 논문에서는 성과에 기초한 유인의 원칙을 적용한 정책 중 하나의 대표적 사례를 집중적으로 분석해 봄으로써 그것이 공공부문의 효율성을 높이는 데 과연 어느 정도의 기여를 했는지를 검증해 보려고 한다. 그것은 교육개혁의 한 방안으로 학생들에게 표준화된 시험을 치게 하고 그 성적에 대해 학교와 교사들이 책임을 지게 하는 정책, 즉 '성과에 기초한 책임(performance-based accountability)' 정책이다.⁽²⁾ 1980년대부터 미국 각 주에서 채택해 왔고, 2000년대 초반에는 연방정부 차원에서도 채택하게 된 이 정책은 교육 전문가들의 많은 관심을 끌어들였다.

이 정책이 몇 개의 주에서 좋은 성과를 거둔 것으로 알려지면서 교육개혁을 이룰 수 있는 특효약처럼 선전되기 시작했고, 급기야 연방정부도 이를 채택하기에 이르렀다. 그러나 이 정책의 효과를 분석한 수많은 논문들은 대체로 부정적인 평가를 내리고 있는 것을 볼 수 있다. 이 정책이 교육의 질을 현저히 개선하는 데 별 기여를 하지 못했을 뿐 아니라, 학교와 교사들의 전략적 행위로 인해 부작용을 낳는 사례도 많이 발견되었다는 것이다. 그 정책이 좋은 성과를 거두었다는 쪽으로 결론을 내고 있는 연구결과가 전혀 없는 것은 아니다. 그러나 부정적인 쪽으로 결론을 내고 있는 연구결과에 비하면 그 숫자가 엄청나게 적은 것이 사실이다.

이와 같은 미국의 경험은 우리에게도 시사하는 바가 크다. 우리나라에서도 2008년 새 정부가 들어서면서 이와 비슷한 방향으로 교육개혁을 시도하려는 움직임이 감지되기 시작했다. 사회 일각의 거센 반대에도 불구하고 정부는 전국의 초중학교를 대상으로 일제고사를 실시하기로 결정한 바 있다. 정부가 일제고사 실시를 통해 궁극적으로 목표하고 있는 것은 미국처럼 성과에 기초한 책임 정책을 채택하는 데 있는 것처럼 보인다. 그렇기 때문에 미국에서 지난 2, 30년 동안의 기간에 걸쳐 이 정책을 시험한 결과가 과연 어떤 것인지에 대한 철저한 검증이 더욱 절실하게 요구된다고 말할 수 있다.

본 논문은 다음과 같이 구성되어 있다. 다음 절에서는 일반적으로 유인계약이 갖는 문제

(2) 학생들의 시험성적에 대해 책임을 지게 만든다는 뜻에서 '시험에 기초한 책임(test-based accountability)' 정책이라고도 불린다.

점을 이론적 차원에서 논의하고 있다. 민간부문에서의 유인계약이 갖는 문제점이 주요 논의대상이지만, 공공부문에도 그대로 적용될 수 있을 뿐 아니라 오히려 문제점이 더욱 증폭되어 나타날 수 있다. 제3절에서는 사례연구로서 최근 미국정부의 교육개혁 노력에서 중심적 위치를 차지하고 있는 ‘성과에 기초한 책임’ 정책이 미국 교육에 어떤 변화를 가져왔는지를 중점적으로 논의하려고 한다. 마지막 절에서는 논의를 요약하고 이와 관련한 정책적 시사점을 찾아보려고 한다.

2. 유인계약의 문제점

유인계약이란 대리인이 자발적으로 열심히 일하도록 만들기 위해 그가 거둔 성과에 따라 보수를 달리 지급하기로 약속한 계약을 뜻한다. 사람은 유인에 반응하는 존재이기 때문에 열심히 일한 데 대한 보상이 지급된다는 사실이 자발적인 노력을 촉진하는 결과를 가져올 가능성이 크다. 그러나 현실에서 이런 논리적 예상이 그대로 들어맞는다는 보장은 없고, 따라서 유인계약이 효율성 제고의 특효약이 된다는 기대가 충족되지 못할 수 있다. 수많은 기업들이 유인계약의 장점을 모를 리 없는데도 이를 적극적으로 활용하지 않는 현실이 무엇을 말해주고 있을까?

일반적으로 유인계약에서 중심적 역할을 하는 것은 어떤 특정한 측면에서 평가한 성과에 기초해 보수를 지급하는 성과급(pay-for-performance)제도다. 이 성과급제도가 바람직한 결과를 가져오기 위해서는 무엇보다 우선 성과 평가의 대상이 되는 측면을 적절하게 선택해야 한다. 예를 들어, 기업 임원의 성과를 평가한다고 할 때 매출액, 이윤, 주식가격 등 다양한 측면이 고려될 수 있는데, 그 중 어떤 것을 평가의 기준으로 삼아야 하는지가 중요한 문제가 되는 것이다. 관련 있는 모든 측면을 고려한다 하더라도 이들 사이에서 어떤 가중치를 부여하는 것이 적절한지의 문제가 있다. 임원이 아닌 일반 직원의 성과를 평가하는 기준을 찾는 것은 더욱 어려운 일이 될 수 있다.

어떤 기준을 선택해 사용한다 하더라도 성과를 정확하게 평가한다는 것은 아주 어려운 일이다. 성과측정의 기준으로 사용되는 지표가 여러 가지 요인에 영향을 받기 때문에 그것의 값이 성과를 정확하게 반영하지 못하는 경우가 많다. 그렇기 때문에 기업은 객관적인 성과측정 지표에 의한 평가를 포기하고 아예 주관적인 성과 평가의 결과에 의해 유인을 제공하는 방법을 선택하기도 한다. 이와 같은 주관적인 평가는 근로자의 성과를 좀 더 총체적인 관점에서 평가할 수 있다는 장점을 갖지만, 이것 또한 나름대로 여러 가지 문제점을 갖고 있다.

무엇보다 우선 주관적 평가제도하에서는 자신의 성과를 평가하는 상급자의 환심을 사기 위해 아까운 자원을 낭비하는 경향이 나타난다. 객관적인 기준에 따른 평가의 경우와 달리, 주관적 평가의 경우에는 평가하는 사람 개인의 의견이 반영될 여지가 크다. 그렇기 때문에 상급자에게 좋은 인상을 심어주기 위한 목적의 행위에 자원을 쏟아 붓는 행위가 나타날 수 있는 것이다. 이런 행위가 본질적인 성과에는 아무런 영향을 주지 못하면서 자원의 낭비만을 가져오게 된다는 것은 두말할 나위도 없다.

프렌더개스트[Prendergast(1999)]는 주관적 평가제도가 이 문제 말고도 관대편향(寬大偏向, leniency bias)과 중심편향(中心偏向, centrality bias)의 문제를 일으킨다고 말한다. 관대편향이란 것은 상급자가 인간적 관계 때문에 냉정한 평가를 내리려 하지 않는 경향을 뜻한다. 열심히 일하지 않는다는 사실을 안다 하더라도, 같은 직장에서 함께 일하는 사람에게 냉정한 평가를 내리기는 어렵다. 그러다 보니 두루뭉술한 평가로 일관하게 되고 이에 따라 대부분의 평가가 비슷하게 보이는 결과가 나타난다. 그렇게 되면 성과가 아주 좋은 사람과 아주 나쁜 사람을 구별하기 힘들게 되는데, 이런 현상을 중심편향이라고 부른다. 평가 결과가 중간수준 부근에서 맴도는 경향이 나타난다는 뜻에서 이런 이름이 붙었다.

결론적으로 말해 유인계약에서 핵심적 사항이라고 할 수 있는 성과측정의 문제에 대해서는 어떤 정답이 있을 수 없다. 즉, 객관적 기준에 의한 평가나 주관적 기준에 의한 평가나 모두 문제점을 갖고 있기 때문에 만족스런 해결책이 될 수 없다는 말이다. 유인계약에 거는 기대에 비해 현실에서 그 효과가 그리 크지 않은 중요한 이유 중 하나를 바로 여기에서 찾을 수 있다. 성과측정의 문제가 제대로 해결되지 못하는 한 어떤 유인계약이라 할지라도 기대하는 효과를 거둘 수 없다.

성과측정의 문제는 대리인의 전략적 행위에 의해 더욱 복잡하게 꼬일 수 있다. 어떤 유인계약이라 할지라도 대리인의 모든 행위를 일정한 방향으로 완벽하게 유도할 수 없다. 대리인은 대리계약의 허점을 찾아내 자신의 이익에 부합되는 방향으로 이를 활용하는 왜곡된 행동을 보일 소지가 있다. 대리인이 적극적으로 유인계약의 허점을 찾는 행위에 나서게 되면 성과측정 지표의 왜곡은 한층 더 심각한 차원으로 치달게 된다. 문제는 이런 현상이 일어날 가능성이 결코 작지 않다는 데 있다. 이런 현상은 단지 이론적 가능성의 차원에 머무는 것이 아니라 현실 경제에서 자주 관찰될 수 있다는 데 문제가 있다.

베이커[Baker(1992)]는 기업이 연구개발에 중사하는 과학자와 맺은 유인계약의 예를 들어 이 문제를 설명하고 있다. 기업은 어떤 과학자가 실제로 얼마나 많은 노력을 기울였는지를 직접 관찰할 수도 없고, 그 과학자가 만들어낸 결과의 진정한 가치가 얼마인지도 모른다. 따라서 그 둘 중 어느 것도 유인계약의 대상이 될 수 없기 때문에 다른 성과 지표에

기초해 유인계약을 체결하는 길을 선택하는 것이 보통이다. 이 경우 유인계약의 대상으로 가장 그럴 듯한 성과지표는 과학자가 얻어낸 특허의 건수다. 그렇다면 특허 건수에 따라 보수를 지급하는 유인계약이 과학자로 하여금 최선의 노력을 기울이게 만드는 유인을 줄 수 있을까?

어떤 특허이냐에 따라 과학자가 기울여야 하는 노력의 크기가 다르다고 가정하자. 즉, 어떤 특허를 얻는 데는 많은 노력을 기울여야 하는 반면, 어떤 특허의 경우에는 별 노력을 기울이지 않아도 쉽게 얻을 수 있다고 가정하는 것이다. 과학자 자신은 특허로 이르는 어떤 특정한 연구프로젝트가 얼마만큼의 노력을 요구하는지 알고 있지만 기업은 그것을 모르는 정보의 비대칭성이 문제를 일으키게 된다. 기업은 평균적으로 하나의 특허를 얻는데 $1/P_e$ 단위의 노력이 들어가야 하며, 특허의 평균 가치는 X 달러라는 사실만 알고 있다(여기에서 P_e 는 과학자가 기울인 노력의 한계생산을 뜻하고 있다).

만약 기업이 과학자와 특허 한 건당 X 달러를 지급하는 유인계약을 체결한다면 과학자는 자신이 갖고 있는 정보의 우월성을 이용해 전략적 행동을 하게 된다. 즉, 그는 노력이 적게 드는 특허 연구프로젝트에 많은 노력을 기울이는 한편, 노력이 많이 드는 특허 연구프로젝트에는 적은 노력밖에 기울이지 않는 전략적 행동을 하게 되는 것이다. 베이커는 이런 상황에서의 최적성과급은 P_e 의 분산을 고려해 X 달러를 적절한 수준으로 낮춘 것이라는 결론을 도출하고 있다. 그에 따르면 이처럼 특허 건수에 따른 성과급의 수준을 하향조정함으로써 과학자의 전략적 행동을 줄일 수 있다는 것이다.

베이커는 대리인의 전략적 행동이 문제가 되는 경우 유인계약을 적절히 조정함으로써 문제를 해결할 수 있다는 해법을 제시하고 있다. 그러나 대리인의 전략적 행동이 유인계약 그 자체를 무의미하게 만들 수 있다. 즉, 유인계약이 대리인의 자발적 노력을 부추기는 유인을 제공할 수 없기 때문에 아예 성과와 무관한 고정급을 지급하는 쪽이 더 낫다는 결론에까지 이르게 만들 수 있는 것이다. 홈스트롬-밀그롬[Holmstrom and Milgrom(1991)]의 ‘다중과업본인-대리인 모형(multitask principal-agent model)’이 바로 이와 같은 가능성에 대해 관심의 초점을 맞추고 있다.

다중과업본인-대리인 모형은 대리인이 여러 가지 과업을 동시에 수행하는 경우나 대리인이 수행하는 하나의 과업에 여러 가지 측면이 존재하는 경우를 분석 대상으로 삼는다. 이 모형에서 유인계약은 대리인으로 하여금 더 열심히 일하도록 만드는 유인을 주는 데 그치는 것이 아니라, 여러 가지 과업 사이에서 노력을 재조정하도록 만드는 역할까지 수행한다. 대리인에게 지급되는 보수를 어떤 특정한 과업과 연동시키면 다른 과업의 수행에 기울일 주의와 노력을 바로 그 과업으로 재배분하는 것이 그 좋은 예다.

일반적으로 기업은 유인계약하에서 근로자가 수행하는 다양한 과업을 하나의 숫자로 대표시킨 지표에 의해 그의 보수를 결정한다. 그런데 그 지표는 근로자가 수행하는 모든 과업을 포괄할 수 없기 때문에 기업의 목표 달성에 도움이 되는 과업인데도 그 지표에 반영되지 않는 경우가 생길 수 있다. 지표에 반영되지 않는 과업에 사용될 시간과 노력을 반영되는 과업으로 재배분하는 일이 실제로 일어난다면 측정지표상에 나타난 성과의 향상은 별 의미를 갖지 못할 수 있다. 그와 같은 표면적인 성과의 향상이 기업의 목표 달성에 도움이 되는데도 불구하고 지표에 반영되지 않아 소홀히 한 과업 탓에 발생한 성과 하락에 의해 상쇄되기 때문이다.

홈스트롬-밀그롬은 유인계약에 대해 대리인이 보일 수 있는 반응이 매우 복잡한 데다가 성과측정의 수단이 아주 불완전하게 갖춰져 있는 상황에서 대리인의 행동을 바람직한 방향으로 이끌어가는 것은 아주 어려운 일이라고 지적한다. 다중과업본인-대리인 모형으로 대표되는 현실의 상황에서 대리인이 오직 하나의 과업만을 수행한다고 가정한 종전의 1차원적 본인-대리인 모형에서 도출한 결론은 별 도움이 되지 않는다는 것이 그들의 주장이다. 어떤 한 과업의 수행 실적에 기초해 유인을 제공하는 방법은 기업의 전반적 성과를 향상시키는 데 크게 기여할 수 없기 때문이다. 그들은 이런 상황에서의 최적유인계약은 측정된 성과와 관련 없이 고정급을 지급하는 것이 될 수 있다는 다소 파격적인 결론을 도출하고 있다. 현실에서 기업들이 성과급제도를 활용하는 사례가 생각 밖으로 적다는 사실을 이를 뒷받침해 주고 있다.

그들은 1차원적 본인-대리인 모형에서 도출한 결론이 별 도움이 되지 않음을 보여주는 현실의 사례로 다음 절에서 본격적으로 논의할 미국의 교육개혁 시도를 들었다. 표준화된 시험 성적에 기초해 학교와 교사의 책임을 묻는 정책, 즉 성과에 기초한 책임 정책의 지지자들은 적절한 유인 제공을 통해 학교와 교사들이 자발적으로 교육에 많은 노력을 쏟아 붓게 만들 수 있다고 주장한다. 이에 비해 그 정책에 반대하는 사람들은 그와 같은 유인이 학교와 교사들의 전략적 반응에 의해 왜곡되기 때문에 교육의 질이 향상되는 효과가 나오기 어렵다고 반박한다. 곧 보게 되겠지만, 이 정책을 실험한 결과를 보면 반대파의 주장이 더 큰 설득력을 갖는 것처럼 보인다. 다중과업본인-대리인 모형은 그와 같은 전략적 반응이 나타날 가능성을 정확하게 예측하고 있다는 점에서 매우 큰 설득력을 갖는다.

그런데 지금까지의 논의는 유인계약이 갖는 문제점이 상대적으로 적은 민간부문의 기업을 대상으로 하는 것이었다. 기업의 경우에도 유인계약이 효율성의 제고를 가져오는 수단으로 많은 문제점을 갖는다는 사실을 고려하면, 이와 비슷한 방식으로 공공부문의 개혁을 추구할 때 얼마나 많은 문제점이 발생할 것인지 능히 짐작할 수 있다. 디싯[Dixit(2002)]은

성과에 기초한 유인이 공공부문 개혁에 효과를 발휘하기 어렵게 만드는 공공부문의 특성 몇 가지를 지적하고 있다.

그 중 첫 번째 것이 다중본인(multiple principals)의 특성인데, 공공부문의 경우에는 기업과 달리 본인이 여럿 존재한다는 뜻이다. 국민 이외에도 다양한 집단이 본인의 역할을 하고 있는데, 이들이 관심을 갖는 사항 혹은 추구하는 목표는 아주 다양하기 때문에 공공부문이 무엇을 추구해야 하는지 목표 그 자체에 대해서도 엇갈린 견해를 가질 수 있다. 공공부문이 추구해야 할 목표에 대해서까지 이견이 존재하는 상황에서 공공부문의 개혁 작업은 어려울 수밖에 없다.

딕шит은 윌슨[Wilson(1989)]의 다음과 같은 비유를 인용해 다중본인의 존재가 일으키는 문제에 대해 설명하고 있다. 이 비유는 본인의 역할을 하는 수많은 개인 혹은 집단들이 정부의 정책에 영향을 미치기 위해 일종의 비협조게임을 벌이고 있는 양상을 묘사하고 있다.

“(미국에서) 정책을 수립하는 일은 마치 선술집에서 벌어지는 주먹싸움과도 같다. 아무나 그 싸움에 가담할 수 있고 어느 누구와도 상대해 싸울 수 있을 뿐 아니라 편을 바꿔 싸우기도 한다. 심판을 맡고 있는 사람도 없고 정해진 라운드만 싸우는 것이 아니라 무한정 싸움이 계속되거나 모두가 피곤에 지쳐 나가떨어질 때까지 싸움이 계속된다[Wilson(1989, pp. 299-300)].”

딕шит이 지적하는 공공부문의 두 번째 특성은 다중과업(multiple tasks)이다. 앞에서 논의한 것처럼 다중과업이라는 특성은 기업에서의 유인계약의 경우에도 문제를 일으킨다. 그런데 공공부문의 경우에는 한층 더 심각한 양상으로 확대되어 나타난다. 공공부문의 경우에는 과업 그 자체뿐 아니라 이를 통해 성취하려 하는 목표(goals)도 여러 가지로 복잡하게 얽혀 있을 수 있기 때문이다. 이와 같은 상황은 어떤 것을 기준으로 개혁을 추구하며 그 성과를 어떻게 평가해야 하는지에 대해 명확한 답을 얻기 어렵게 만든다. 예컨대 기업의 경우에는 이윤이나 주시가격 같은 명확한 목표와 평가의 기준이 존재하지만, 공공부문은 그런 것이 존재하지 않기 때문에 개혁을 추진하는 일이 한층 더 어려울 수밖에 없다.

이와 더불어 딕шит은 민간부문에서 볼 수 있는 경쟁이 결여되어 있다는 점이 개혁 작업을 한층 더 어렵게 만들 수 있다고 말한다. 다른 기업들과 치열한 경쟁을 벌여야 하는 상황은 효율성을 극대화해야 살아남을 수 있는 분위기를 만든다. 그러나 공공부문은 실질적으로 독점상태에 있기 때문에 이런 외부적 압력에서 자유롭고, 따라서 자발적으로 효율성을 높이라는 의욕에 문제가 있을 수 있다. 바꿔 말하면 라이벤스타인[Leibenstein(1966)]이 말한

X-비효율성(X-inefficiency)과 비슷한 성격의 문제점을 가질 수 있다는 뜻이다. 이런 상황에서는 성과에 기초한 유인을 제공한다 해도 공공부문의 효율성을 높이는 데 큰 효과를 내지 못할 수 있다.

결론적으로 말해 공공부문은 민간부문에 비해 유인계약이 효과를 발휘하기 어려운 점을 여럿 갖고 있다. 민간부문에서도 유인계약이 여러 가지 문제점을 갖는 터에, 공공부문의 경우에는 이런 어려움까지 추가되기 때문에 공공부문을 개혁하는 특효약이 될 수 없다. 그런데도 성과에 기초한 유인을 제공함으로써 공공부문을 효과적으로 개혁할 수 있다는 믿음을 갖고 있는 사람이 생각 밖으로 많다. 미국 교육개혁의 기수로 등장한 성과에 기초한 책임 정책은 적절한 유인의 제공이 공공부문 개혁의 핵심이라고 본다는 점에서 그와 같은 믿음을 구체화시켜 표현하고 있다. 문제는 이와 같은 방향으로의 개혁 시도가 과연 현실에서 어떤 결과를 가져왔는지에 있으며, 다음 절에서의 논의는 바로 이 점에 관심의 초점을 맞추고 있다.

3. 성과에 기초한 책임 정책이 교육의 개혁을 가져 왔는가?

미국의 부시(G. Bush) 대통령은 후보 시절 자신이 텍사스 주지사로 재임했던 시절 교육 개혁에 성공해 소위 '텍사스의 기적(Texas Miracle)'을 만들어 냈다는 사실을 부각시켜 자신이야말로 미국 교육을 개혁할 책임자임을 자처했다. 그리고 대통령으로 취임하자마자 텍사스주와 비슷한 방식으로 개혁을 추구하는 No Child Left Behind(NCLB) 법안을 통과시키기 위해 전력을 기울였다. 부시대통령은 이 법안의 서문에서 미국의 수많은 아동들이 낮은 기대수준, 지식의 결여, 그리고 자신감의 상실로 인해 경쟁대열에서 뒤처지고 있는 현실을 지적하고, 그와 같은 교육개혁을 통해 어느 누구도 경쟁대열에서 뒤처지지 않게 만들 것임을 다짐하고 있다. 바로 그와 같은 다짐을 반영하는 NCLB 정책은 부시행정부가 추진한 교육 개혁의 상징처럼 여겨지고 있다.

이 NCLB 정책의 핵심을 구성하는 것은 '성과에 기초한 책임'이라는 원칙이다. 이 정책 하에서 각 학교는 책임을 지고 모든 학생들이 높은 수준의 교육을 받을 수 있도록 만들어야 하며, 주정부는 학생들의 학업성취 수준을 높이는 데 성공한 학교와 실패한 학교에게 적절한 포상과 제재를 가할 수 있는 체제를 준비해 놓도록 요구되고 있다. 학업성취의 수준은 학생들이 해마다 치르는 표준화된 시험(standardized test)의 평균 성적에 의해 평가되며, 연방정부의 자금을 받는 학교는 매년 이 성적을 어느 정도로 향상시키겠다는 약속을 해야 한다. 만약 어떤 학교의 실제 성과가 Adequate Yearly Progress(AYP)라고 부르는 이

성과개선 목표에 미달하면 그 학교는 열악한 성과를 거둔 학교로 간주되어 이에 상응하는 조치의 대상이 된다.

예를 들어 4년 연속 AYP 달성에 실패하면 그 학교는 ‘시정조치(corrective action)’가 필요한 학교라는 불명예스러운 명칭을 받게 되는데, 그 시정조치라는 것은 교직원의 대폭 교체, 새로운 교과목의 도입, 수업시간 연장 등을 포함한다.⁽³⁾ 5년 연속으로 실패하게 되면 학교 전체를 재편하는 계획이 만들어지며, 6년째까지도 실패하면 그 계획이 실행에 옮겨지게 된다. 이 경우 학교가 폐쇄될 수도 있고, 민간회사가 학교를 운영하게 될 수도 있고, 주정부의 교육청이 학교를 직접 운영하게 될 수도 있다.

이와 같이 성과가 나쁜 학교는 제재를 받는 반면에 좋은 성과를 거둔 학교는 예산상의 보너스를 받는 혜택을 누리게 된다. 바로 이 성과에 기초한 유인을 제공함으로써 학교와 교사로 하여금 자발적으로 열심히 학생들을 가르치려 노력하게 만들 수 있다는 데 NCLB 정책의 주안점이 있다. 신자유주의적 색채를 강하게 띠고 있던 부시행정부로서는 이와 같은 방향으로 교육부문의 개혁을 추구하는 것이 아주 자연스러운 선택이었다. 성과에 기초한 유인만이 계엄을 피우는 학교와 교사를 분발하게 만들 수 있는 유일한 방법이라는 생각은 신자유주의적 개혁 구상과 아주 잘 어울리는 것이기 때문이다.

그런데 연방정부가 마치 백지 위에 그림을 그리듯 아무것도 없는 상황에서 머리를 짜내 NCLB 정책을 만들어낸 것은 아니었다. 실제로는 주정부 차원에서 이와 비슷한 정책이 오랜 기간 동안 실시되어 왔고, 부시행정부에 의해 연방정부 차원에서도 이를 도입하기에 이른 것이다. 1980년을 전후해 몇 개의 주정부가 표준화된 시험을 실시하기 시작했고, 이와 같은 움직임이 다른 주들로 확산되어 가면서 2000년 즈음에는 전국적인 추세가 되었다. 케인-스테거[Kane and Staiger(2002)]에 따르면, 2000년 봄 현재 40개 주가 표준화된 시험 성적으로 학교의 성과를 평가하는 제도를 실시하고 있으며, 20개 주는 이에서 한 걸음 더 나아가 이 성과평가의 결과에 기초해 금전적인 상과 벌을 준다고 한다.

여러 주에서 실시된 교육개혁 프로그램 중 NCLB 정책과 여러 가지 점에서 가장 비슷한 성격을 갖는 것은 플로리다주의 ‘A플러스 프로그램(A-Plus Program)’이다. 이 프로그램은 읽기, 쓰기, 수학 세 부문에서 학업성취도를 측정하는 Florida Comprehensive Assessment Test(FCAT)의 성적에 기초해 마치 대학에서 학점을 부여하듯 각 학교에 A로부터 F에 이르는 성적을 매긴다. 만약 어떤 학교가 4년 동안 두 번에 걸쳐 F를 받으면 그 학교 학생은 다

(3) 2년 연속 AYP 달성에 실패한 학교는 공개적으로 ‘개선이 필요한(in need of improvement)’ 학교라는 명칭을 받게 되며, 문제점을 보완하는 2개년 개선계획을 세워 제출해야 한다. 그리고 그 학교 학생들은 동일 학군 안에 있는 더 좋은 학교로의 전학을 선택할 수 있다.

른 공립학교나 사립학교로 전학할 수 있는 기회를 갖게 된다.⁽⁴⁾ F라는 평가는 그 자체가 불명예스러운 뿐더러, 이를 두 번 받아 학생이 모두 빠져나가면 결국 폐교의 위협에까지 직면하게 될 수 있다.

이와 같은 강력한 유인이 적용되는 상황에서 학교와 교사가 F를 피하기 위해 필사적인 노력을 하리라는 것은 당연한 일이다. 이에 따라 교육의 질도 현저하게 개선되기를 기대할 수 있다는 것이 플로리다 주정부의 논리다. 라우즈 등[Rouse *et al.*(2007)]의 연구에 따르면 플로리다의 학교들에서 실제로 그런 일이 일어났다고 한다. 그들은 2002년 여름 F를 받은 학교들이 그 뒤 곧바로 더 좋은 평가를 받게 되었으며, 그와 같은 변화는 교육의 방침과 관행을 바꾼 데 따른 본질적인 개선의 성격을 갖는다고 평가하고 있다. A플러스프로그램에 대한 긍정적 평가는 그린[Greene(2001)]의 논문에서도 찾아볼 수 있는데, 그는 플로리다주의 학교들이 보인 성과가 진정한 의미에서의 교육의 질 향상을 뜻한다는 결론을 내리고 있다.

반면에 차크라바티[Chakrabarti(2007)]는 겉으로 드러난 성과가 진정한 교육의 질 향상을 뜻하지 않는다는 정반대의 결론을 내리고 있어 우리의 관심을 끈다. 그에 따르면 F를 받은 학교는 두 번째 F를 피하기 위해 전략적인 행동을 할 가능성이 크다고 한다. 학교에 대한 평가는 FCAT 성적이 일정한 커트라인을 넘는 학생들의 비율이 얼마인지에 의해 내려진다. 따라서 F를 받은 학교는 커트라인 바로 아래 성적을 받은 학생들에게 집중적인 교육을 하게 될 가능성이 크다. 그런 학생들의 성적이 올라가 커트라인을 넘게 됨에 따라 평가등급이 향상될 수 있기 때문이다. 말하자면 학교와 교사들의 전략적 반응에 의해 성과가 좋아진 것처럼 보인 데 불과했다는 주장인데, 이 문제는 바로 뒤에서 심도 있게 논의해 보려고 한다.

그런데 신자유주의적 교육개혁이라는 측면에서 플로리다주보다 더욱 많은 관심을 모은 것은 텍사스주다. 특히 텍사스주의 교육개혁 노력이 훌륭한 성과를 거두었다는 사실이 널리 유포되면서 ‘텍사스의 기적(Texas miracle)’이란 말까지 만들어지기에 이르렀다. 이렇게 텍사스주의 교육개혁이 언론의 화려한 조명을 받게 된 것은 주지사 부시의 대통령 출마와 상당한 관련을 갖고 있다. 자신은 텍사스주의 교육개혁을 성공적으로 이끌었고 대통령이 되면 전국의 교육을 그와 같은 방식으로 개혁하겠다고 선전하는 과정에서 언론의 집중적

(4) 이를 좀 더 정확하게 표현하면, 그 학교 학생들은 ‘사립학교로 전학했을 때 사용할 수 있는 바우처(voucher)를 받는다.’가 된다. 이렇게 전학하려는 학생을 위해 바우처를 발행하는 제도를 Opportunity Scholarship Program이라고 부른다. 플로리다주 대법원은 2006년 1월 이 프로그램에 대해 위헌 결정을 내렸다.

조명을 받게 되었기 때문이다.

텍사스의 기적이란 말이 나온 근거는 Texas Assessment Academic Skills(TAAS)라는 표준화된 시험에 기초한 교육개혁 작업이 놀랄 만한 성과를 거두었다는 주정부의 자체평가였다. 1990년에 시작된 이 시험에서 일정한 커트라인을 넘는 학생의 비율이 얼마인지에 기초해 ‘모범적(exemplary)’에서 ‘수용불가(unacceptable)’에 이르는 4등급의 평가를 받게 된다. 높은 평가를 받는 학교는 현금 보너스를 받게 되는 반면, 낮은 평가를 받는 학교는 폐교에까지 이르는 벌칙을 받도록 되어 있다. 이와 같은 강력한 유인을 부여함으로써 텍사스주의 학교들을 근본적으로 개혁하는 데 성공했다는 것이 주정부의 자체평가였다. 그 증거로 TASS 성적이 크게 향상되었을 뿐 아니라, 백인 학생과 소수인종 학생 사이의 격차도 줄이는 성과를 거두었다는 통계자료를 제시하고 있다.

그러나 텍사스 주정부가 제시한 통계자료만으로 진정한 교육개혁이 이루어졌다고 말하기는 힘들다고 지적하는 사람들이 많다. 텍사스의 기적이란 말이 나오기 시작한 직후부터 많은 교육전문가들이 이 말의 진실성을 검증하는 작업에 착수한 바 있는데, 대부분의 연구 결과가 ‘기적’이란 표현은 과장된 것임을 지적하고 있다. 텍사스주의 교육개혁 성과가 실제로는 그다지 훌륭한 것이 아니었다는 결론을 가장 설득력 있게 제시하고 있는 것은 바로 헤이니[Haney(2000)]의 논문이다. 그는 우선 백인 학생과 소수인종 학생 사이의 격차가 줄어들었다는 주장이 통계의 교묘한 조작에 의해 만들어진 허상임을 밝혔다. 이어서 주정부가 의기양양하게 내걸고 있는 TASS 평균성적의 급격한 향상도 실제로 교육의 질이 그만큼 향상되었다는 것을 뜻하지는 않는다고 주장했다.

헤이니는 만약 텍사스주 학생들의 학업성취도에 진정한 향상이 있었다면 TASS 이외의 평가지표상에도 그와 같은 변화의 징후가 나타나야 마땅하다고 지적한다. 그가 텍사스주 학생들의 Scholastic Aptitude Test(SAT) 성적을 조사해 본 바에 따르면 1990년 이래 이렇다 할 향상의 기미를 발견할 수 없었다. 사실 텍사스 학생의 SAT 성적은 1970년 이래 줄곧 전국 평균수준을 밑돌았으며, TASS 성적이 현저하게 올랐다는 1990년대에 들어오면서도 그 격차는 줄어들지 않았던 것이 사실이다. 이와 같은 분석결과에 기초해 헤이니는 TASS 성적의 향상이 교육의 질 향상을 뜻하지 않는다는 결론을 내렸다.

클라인 등[Klein *et al.*(2000)]은 헤이니와 비슷한 목적으로 또 다른 전국적 차원의 평가지표인 National Assessment of Educational Progress(NAEP)에서 텍사스주 학생들이 어느 정도의 성적을 올렸는지 조사해 보았다. 그들은 텍사스주에서 개발한 TAAS보다 전국적 규모의 시험인 NAEP가 학업성취도를 좀 더 정확하게 측정할 수 있다고 믿었다. 전국적인 규모로 실시되는 시험인 만큼 내용의 충실성이나 평가의 객관성에서 앞선다고 보았기 때문

이다. 또한 TASS는 학교와 교사의 평가에 반영되기 때문에 무슨 수를 써서든 성적을 올려야 하는 압력을 받을 수 있지만, NAEP의 경우에는 그런 압력이 없다는 점도 중요하게 고려되었다.

클라인 등이 밝혀낸 바에 따르면, 조사대상이 되는 기간 동안 전국적으로 NAEP 성적의 향상이 일어났다고 한다. 그런데 텍사스주의 경우에는 교육개혁 작업이 시작된 후에도 그 주 학생들의 성적 향상폭이 전국 평균치를 밑도는 경우가 많은 것으로 드러났다. 그 기간 동안 TASS 성적에는 괄목할 만한 향상이 있었지만, NAEP 성적상에는 그와 같은 변화가 나타나지 않았던 것이다. 뿐만 아니라 같은 기간 동안 백인 학생과 소수인종 학생 사이의 NAEP 성적 격차가 점차 더 크게 벌어진 것을 볼 수 있었다. 주정부는 TAAS 성적에 기초해 인종간 격차가 줄어들었다는 점을 대대적으로 선전했지만, 다른 지표를 보면 오히려 인종간 격차가 한층 더 벌어지고 있었던 것이다.

만약 헤이니와 클라인 등이 주장하는 것처럼 TAAS 이외의 평가지표상으로 텍사스주의 개혁 노력이 그다지 좋은 성과를 거두지 못한 것이 사실이라면 TAAS 성적의 괄목할 만한 향상은 어떻게 해석해야 할까? 헤이니가 텍사스주 교사에 대해 실시한 설문조사 결과를 검토해 보면 이 의문에 대한 해답을 얻을 수 있다. 이 설문조사에서 교사들은 TAAS가 도입된 이후 자신들이 이 시험에 특별히 대비하기 위해 많은 시간과 노력을 기울였음을 실토했다. 또한 이와 같은 시험대비 위주의 교육이 학생들의 학습능력 배양에 도움이 되기는커녕 손해를 입혔다는 점도 솔직하게 인정했다. 말하자면 학교와 교사들이 평가의 기준이 되는 시험 성적 그 자체를 위해서 전력을 기울인 결과 TAAS 성적이 향상되었을 뿐이며, 이는 진정한 학업성취도 향상과 아무 관련이 없다는 뜻이다.

물론 텍사스의 기적이 맞는 말이라고 두둔하는 교육전문가들도 있기는 하다.⁽⁵⁾ 그러나 이 문제를 분석하고 있는 대부분의 연구들이 텍사스의 기적이란 말은 과장에 지나지 않는다는 부정적인 결론을 내리고 있다. 텍사스주뿐 아니라 비슷한 교육개혁 정책을 추구한 다른 주의 성과를 평가한 결과를 보아도 결론이 대체로 비슷한 방향으로 흐르는 것을 볼 수 있다. 즉, 학교와 교사들이 해당 주에서 실시하는 표준화된 시험 성적 그 자체를 올리는 데만 관심을 갖고 있을 뿐 교육의 질 그 자체를 높이는 데는 아무런 관심을 보이지 않음으로써 진정한 개혁을 이루는 데 실패했다는 결론이 지배적인 것이다.

제2절에서 다중과업의 상황에서는 민간부문의 유인계약이 대리인의 전략적 반응에 의해 기대한 효과를 거두지 못하는 경우가 많음을 지적한 바 있다. 지금 우리가 고려대상으

(5) 예를 들어 Jay Greene(2000): "The Texas School Miracle Is for Real," *City Journal* 10을 보라.

로 삼고 있는 교육개혁의 문제 역시 전형적인 다중과업의 상황이라고 말할 수 있다. 교사들이 교육이라는 범주 안에서 수행해야 하는 구체적 과업은 여러 가지가 있기 때문이다. 이 상황에서 표준화된 시험 성적만을 기준으로 해 강력한 유인을 부과하게 되면 교사는 다른 중요한 업무에 쓸 시간과 노력을 시험 준비에 재배정하는 전략적 반응을 보일 가능성이 크다. 따라서 성과에 기초한 책임 정책은 ‘시험맞춤수업(teaching to the test)’의 유행을 가져올 가능성이 큰데, 여러 연구자들이 실제로 이와 같은 경향이 나타나고 있음을 증언하고 있다.⁽⁶⁾

그러나 시험맞춤수업은 성과에 기초한 책임 정책에 대한 전략적 반응의 사례 중 비교적 점잖은 편에 속하는 것이라고 할 수 있다. 시험 성적 그 자체를 올리는 데만 관심을 갖는 학교와 교사들이 그 정책에 어떤 전략적 반응을 보였는지를 분석한 수많은 논문들이 쏟아져 나온 바 있는데, 그것들을 읽어보면 온갖 기상천외한 방법을 모두 동원해 점수를 올리려 하고 있다는 사실을 발견할 수 있다. 이런 전략적 행위가 교육의 질 그 자체를 높이려는 노력과 아무런 관련이 없음은 두말할 나위도 없다. 표준화된 시험 성적이 향상된 것만을 보고 성과에 기초한 책임 정책이 학생들의 학업성취도를 높이는 효과를 가져왔다고 말할 수 없는 이유가 바로 여기에 있다.

시험 성적이 향상된 것처럼 보이게 만들기 위해 사용하는 방법으로 시험맞춤수업 다음으로 자주 사용된 것은 시험을 치르는 학생들의 집단을 조작하는 것이다. 특별한 사정이 있는 학생들은 시험을 치지 않아도 된다는 규정을 이용해 나쁜 점수를 받을 것으로 예상되는 학생은 미리 손을 써서 시험을 치지 않도록 만드는 수법이다.⁽⁷⁾ 컬렌-르백[Cullen and Reback(2006)]에 따르면, 텍사스주의 학교에서 흑인과 히스패닉 학생들의 시험면제율이 정상적인 수준에 비해 7%에서 14%에 이르는 폭으로 더 높아진 것을 발견할 수 있다고 한다. 이렇게 시험 성적이 나쁜 경향이 있는 집단이 미리 빠져나감으로써 평균 성적이 자동적으로 높아지는 효과가 나타날 수 있다.

필리오-게츨러[Figlio and Getzler(2002)]는 학교와 교사들에게 엄격한 책임을 묻는 플로

(6) 시험맞춤수업으로 시험 성적 그 자체를 끌어올리려는 노력이라는 형태의 전략적 반응을 논의하고 있는 논문은 앞에서 말한 Haney(2000) 이외에도 Koretz(2002), Hoffman *et al.*(2001), Jacob(2005) 등 수없이 많다. 사실 성과에 기초한 책임 정책에 대해 논의하는 대부분의 연구들이 시험맞춤교육을 거의 기정사실로 받아들이는 경향을 발견할 수 있다.

(7) 시험에서 면제되는 특별한 사정 중 가장 흔한 것은 심각한 장애를 갖고 있어 특수교육(special education)이 필요하다고 인정되는 경우다. 이 이외에도 학생이 영어에 능통하지 못하는데 스페인어 시험이 그 해에 제공되지 않는 경우, 그리고 병에 걸렸거나 결석을 한 경우도 시험면제 사항이 된다.

리다주에서 시험을 치르는 학생들의 집단을 조작하는 사례를 흔하게 볼 수 있다고 말한다. 즉, 이 정책을 채택한 후 저소득층의 자제나 과거에 성적이 나빴던 학생을 시험 면제 대상으로 분류하는 비율이 현저하게 높아지는 현상이 많은 학교에서 발견되었다는 것이다. 심지어 어떤 학교에서는 시험 성적이 좋게 보이도록 만들기 위해 학생처벌 규정까지 동원한 것으로 밝혀졌다[Figlio(2006)]. 예를 들어, 어떤 학생에게 정학처분을 내릴 때 정학기간을 일부러 늘려 시험을 칠 수 없게 만드는 방법을 쓰는 것이다. 지금 보는 것처럼 학교들은 다양한 방법을 동원해 시험을 치르는 학생들의 집단을 조작하는 것으로 드러났다.

이보다 한층 더 기상천외한 방법이 동원된 사례까지 있다. 폴리오-위니키[Figlio and Winicki(2005)]는 버지니아주의 학교들이 급식프로그램까지 활용해 시험 성적을 높이기 위해 노력한다는 사실을 발견했다. 영양학적 연구에 따르면, 포도당(glucose)의 섭취량이 많아짐에 따라 인식능력에 단기적 향상이 일어날 수 있다고 한다. 학교 당국은 이 사실을 이용해 시험기간 동안의 급식 식단을 조정해 학생들로 하여금 더 많은 포도당을 섭취하도록 만든다는 것이다. 많은 포도당을 섭취한다는 것은 칼로리 섭취량이 많아진다는 것을 뜻하는데, 나쁜 시험 성적으로 인해 제재의 위협에 직면한 학교가 많이 포함되어 있는 학군일 수록 학생들이 시험기간 중 급식에서 섭취하는 칼로리가 더 높은 경향이 있는 것으로 드러났다.⁽⁸⁾

흥미로운 점은 이와 같은 급식 내용의 조정이 시험 성적을 올리는 데 어느 정도 효과를 냈다는 사실이다. 폴리오-위니키의 분석에 따르면, 칼로리 섭취량의 증가가 통계적으로 유의한 정도의 수학 성적 향상을 가져 왔다고 한다. 한편 영어와 역사/사회 과목의 경우에는 성적 향상에 기여하는 것으로 드러났지만 통계적 유의성은 아주 작다고 한다.⁽⁹⁾ 이는 학생들로 하여금 100칼로리를 더 섭취하게 만들으로써 수학, 영어, 역사/사회 과목의 합격률을 각각 7%, 4%, 7% 높일 수 있음을 뜻한다는 것이 그들의 해석이다. 이와 같은 ‘시험맞춤급식’은 교육의 질 향상과 아무런 관련을 갖지도 못하면서 그렇지 않아도 심각한 아동비만의 문제를 더욱 심각한 차원으로 몰고 가는 결과를 가져왔다.

학교와 교사들의 전략적 반응 중에서 가장 심각한 문제가 되는 것은 점수를 높이기 위해 저지르는 부정행위(cheating)다. 얼마 전 우리나라에서 일제고사가 시행되었을 때 어느 지

(8) 보통 때의 평균 칼로리 섭취량이 740 내외인데 비해 시험기간 중에는 이것이 825의 수준으로 치솟는 것으로 나타났다.

(9) 이례적으로 과학 성적은 아주 작은 폭으로 내려가는 것으로 관찰되었다. 그러나 폴리오-위니키는 칼로리 섭취량의 증가가 전반적으로 성적 향상에 기여한다는 쪽으로 결론을 내려도 무방하다고 말한다.

역의 성적이 유달리 좋았고 이를 수상하게 여겨 조사해본 결과 부정행위를 통해 점수를 높였음이 드러난 바 있다. 미국의 학교들은 시험 성적과 관련해 우리의 학교들에 비해 훨씬 더 강한 압박을 받고 있기 때문에 부정행위의 유혹도 그만큼 클 수밖에 없다. 부정행위를 하는 것은 나쁘다고 가르쳐야 할 학교와 교사가 부정행위를 저지른다는 것은 어이없는 일이지만 이런 일이 일어날 가능성을 배제할 수 없다.

언론 보도를 보면 교육 현장에서 실제로 부정행위가 저질러지는 사례가 그리 드물지 않다는 사실을 알게 된다. *Education Weeks*는 뉴욕, 시카고, 버지니아, 캘리포니아, 메릴랜드 등지에서 다양한 형태의 부정행위가 적발되었음을 상세하게 보도하고 있다[Hoff(2000)]. 흥미로운 점은 비슷한 종류의 부정행위에 대해 어떤 지역은 교사를 처벌한 데 비해 어떤 지역은 처벌을 하지 않는 사례가 있다는 사실이다. 다시 말해, 교사들이 성적을 올리기 위해 비정상적인 방법을 쓰는데도 이를 부정행위로 간주하지 않고 용인해 주기도 한다는 뜻이다. 무슨 수단을 써서라도 성적을 올려야 한다는 압력이 크다 보니 웬만큼 비정상적인 방법을 쓰는 것은 용인해주는 풍토가 생긴 것이다.

또한 *San Francisco Chronicle*은 캘리포니아주의 51개 학교가 부정행위로 의심되는 행동을 했다는 이유로 현금 보너스 지급대상에서 제외되었다는 기사를 실고 있다[May(2000)]. 그 당시 캘리포니아주는 학생들의 시험 성적에 기초해 학교와 교사를 평가한 다음 현금 보너스를 지급하는 정책을 채택하고 있었다. 일정한 수준을 넘는 학교는 학생 1인당 67달러의 현금 보너스를 받는데, 학교는 이를 마음대로 사용할 수 있다. 그리고 좋은 성과를 올린 교사에게는 5천 달러에서 2만 5천 달러에 이르는 현금 보너스를 지급한다. 이와 같은 강력한 유인이 학교와 교사의 부정행위를 부추기는 역할을 했다는 비판이 제기되고 있는 것이다.

최근에는 조지아주 아틀란타시에서 44개 학교와 178명의 교사가 연루된 사상 최악의 성적조작 사건이 발생해 미국 국민을 충격에 몰아넣었다[한국일보(2011. 7. 7)]. 조사대상 56개 학교 중 44개 학교가 적발되었다는 사실은 조사범위를 확대할 경우 사건이 엄청난 범위로 확산될 수도 있음을 시사해 주고 있다. 심지어 2009년도 ‘올해의 교육감상’을 받았고 교육개혁가로서 전국에 이름이 알려진 전 애틀란타시 교육감의 연루설까지 흘러나오고 있는 실정이다. NCLB 정책으로 인한 성적지상주의가 얼마나 심각한 교육계의 타락을 가져왔는지를 생생하게 보여주는 사례가 아닐 수 없다.

제이콥-레빗[*Jacob and Levitt(2003)*]은 좀 더 엄격한 방법에 의해 부정행위가 존재한다는 사실을 밝혀냈다. 그들은 시험 성적에 의외의 변동이 나타나거나 학생들이 제출한 답안에 특이한 양상이 나타난다면 학교측의 부정행위가 있었다는 증거가 될 수 있다고 말한다.

부정행위가 저질러지면 어떤 한 해의 성적이 급격히 상승했다가 그 다음 해에는 상승이 거의 없거나 오히려 하락하는 현상이 나타날 수 있다. 또한 교사가 답을 가르쳐준다면 학생들이 똑같은 답을 써내거나, 쉬운 문제는 틀리면서 아주 어려운 문제는 맞추는 것 같은 특이한 양상이 나타나게 된다.

1993년에서 2000년에 이르는 기간 동안 시카고 지역의 공립학교를 대상으로 한 제이콥-레빗의 분석결과를 보면 학교와 교사의 부정행위가 분명히 존재하는 것으로 드러난다. 그들은 해마다 전체 학급의 4~5%에 이르는 학급에서 부정행위가 일어난다는 분석결과를 제시하고 있다. 물론 이 수치만을 갖고 부정행위가 널리 만연되어 있다는 결론을 내기는 힘들다. 그러나 성과에 대한 과도한 압박이 일부 학교와 교사로 하여금 부정행위까지 저지르게 만든다는 결론을 내기에는 충분한 근거가 될 수 있다. 아무리 열렬하게 성과에 기초한 책임 정책을 지지하는 사람이라 할지라도 이런 부작용까지 유발한다는 사실을 인지하게 되면 생각이 어느 정도 바뀔 수밖에 없을 것이다.

지금까지의 논의를 정리해 보면, 당초의 기대와 달리 성과에 기초한 책임 정책이 미국 교육에 획기적인 변혁을 가져다주지는 못했다는 결론에 이르게 된다. 이 정책을 지지하는 사람들은 학생들의 시험성적이 크게 올랐다는 사실을 들어 좋은 성과를 거두었다고 강변하고 있으나, 여러 사람들의 연구에 의해 시험성적의 상승이 진정한 의미에서의 교육의 질 향상을 의미하는 것은 아니라는 사실이 밝혀졌다. 시험성적에 기초해 보상과 제재를 가하는 정책은 학교와 교사로 하여금 교육의 질을 높이려는 노력은 뒷전에 밀어둔 채 시험성적 그 자체만을 올리는 데 모든 노력을 기울이게 만드는 결과를 가져 왔다.

교육의 성과를 단지 학생들의 시험성적에 의해 평가할 수 있다는 발상은 너무나도 순진한 것이다. 시험성적이 학업성취도를 정확하게 반영한다고 볼 수도 없으며, 교육은 단순하게 지식을 전달해 습득하게 만드는 것 이상의 의미를 갖고 있다. 교육의 또 다른 중요한 목표는 학생들로 하여금 주위의 사물과 사건에 대한 지적 호기심, 그리고 자신이 직면한 문제를 스스로 풀어나가는 능력을 배양하는 데 있다. 단순히 지식을 축적하는 데 그치지 않고 창의성을 발휘하도록 북돋는 데 교육의 또 다른 목표가 있다. 뿐만 아니라 행복한 삶을 일구어 가는 방법을 가르쳐 주고 좋은 시민으로서의 자질을 키워주는 것도 중요한 일이다. 이런 측면에서의 교육의 성과를 그 어떤 표준화된 시험으로도 측정할 수 없음은 두말할 나위가 없다.

2002년에 발효된 NCLB 정책 그 자체가 미국 교육 전반에 걸쳐 어떤 영향을 미쳤는지에 대한 본격적 연구는 아직 이루어진 바 없다. 통계자료의 축적이 미진한 탓에 심도 있는 분석이 어렵기 때문이다. 그러나 주정부 차원에서 실시된 성과에 기초한 책임 정책이 가져온

귀결에 비추어 보면 미국 교육의 획기적 전환점을 가져오게 될 것이라는 당초의 기대가 실현된다는 것은 매우 어려울 것으로 보인다. 비슷한 내용의 정책이 연방정부 차원에서 실시된다 해서 그 귀결이 크게 달라질 이유가 없기 때문이다.

오바마(B. Obama) 대통령은 취임한 이래 기회 있을 때마다 미국 교육이 고질적인 병폐에서 벗어나지 못하고 있는 현실을 개탄하는 발언을 계속해 왔다. 그리고 2011년의 연두연설(State of the Union Address)에서는 NCLB를 다른 법률로 대체할 것을 고려할 것이라고 시사했다. 이는 그와 같은 정책이 미국 교육을 개혁하는 데 별 효과를 거두지 못했다고 판단했음을 뜻한다. 만약 그와 같은 구상이 실현에 옮겨진다면 강력한 유인을 제공해 교육부문을 개혁한다는 접근법은 30여 년 동안의 실험기간을 끝내고 무대에서 퇴장하는 운명을 맞게 될 것이다.

4. 맺음말

인간이 유인에 반응하는 존재라는 것은 경제학의 몇 안 되는 기본법칙 중 하나다. 유인계약은 바로 이 점에 착안해 적절한 유인을 제공함으로써 대리인이 자발적으로 열심히 노력하도록 만든다는 관점에서 도입된 장치다. 그러나 유인계약은 민간부문의 기업에서조차 여러 가지 문제점 때문에 기대된 효과를 내지 못하는 경우가 많다. 인간은 주어진 유인에 조건반사적으로 반응하는 존재가 아니다. 현실의 인간은 자신의 이해관계에 따라 주어진 유인에 대해 복잡한 전략적 반응을 보인다. 따라서 아무리 정교하게 유인을 설계한다 해도 기업이 원하는 결과를 100% 얻기 힘든 것이 보통이다. 많은 기업들이 유인계약을 적극적으로 활용하지 않고 비효율적인 것처럼 보이는 고정급제도를 고수하는 주요한 이유 중 하나를 바로 여기에서 발견할 수 있다.

이렇게 민간부문에서도 좋은 성과를 거두기 힘든 유인계약이 공공부문 개혁의 수단으로 훌륭한 역할을 하리라고 기대하기는 힘들다. 공공부문의 여러 가지 특성 때문에 유인계약이 효과를 발휘하기가 한층 더 어려운 여건이 조성되어 있기 때문이다. 그러나 신자유주의적 이념을 신봉하는 사람들은 강력한 유인을 제공함으로써 잠자고 있는 공공부문을 흔들 어 깨울 수 있다는 믿음을 버리지 않았다. 1980년대 이후 신자유주의적 이념이 점차 그 영향력을 넓혀가면서 이와 같은 방식으로 공공부문을 개혁하고자 하는 움직임이 도처에서 감지되기 시작했다.

미국 주정부 차원에서 시작된 성과에 기초한 책임 정책, 그리고 이것의 연장선상에서 연방정부가 도입한 NCLB 정책은 이 신자유주의적 개혁 노력의 대표적 사례라고 할 수 있다.

이 교육개혁 프로그램이 미국 교육에 어떤 영향을 미쳤는지를 심도있게 분석해 보면 과연 신자유주의적 개혁이 공공부문의 대대적 혁신을 가져올 수 있는지에 대한 분명한 답을 얻을 수 있다. 막연한 탁상공론에 의해 그 답을 구하려 하면 몇 십 년이 가도 결론을 낼 수 없을지 모르지만, 이와 같은 실험적 상황의 귀결을 분석해 보면 이론의 여지가 없는 명확한 결론을 낼 수 있다. 이런 점에서 볼 때 미국의 교육개혁 프로그램에 대한 사례분석은 매우 중요한 의미를 갖는다.

한 마디로 말해 강력한 유인을 제공해 교육개혁을 이루려는 시도는 약속된 결과를 가져다주는 데 실패했다. 성과 평가의 기초가 되는 시험의 성적은 향상된 것으로 나타났지만, 그것이 진정한 의미에서의 교육의 질 향상을 뜻하는 것은 아니었다. 여러 가지 증거에 비추어 판단해 볼 때, 성적의 향상은 성적 그 자체를 올리려는 학교와 교사의 노력이 가져온 결과에 불과하다는 결론이 나온다. 그 과정에서 오히려 교육의 내용이 왜곡되는 부작용만 가져왔다고 볼 수 있다. 객관적인 수치로는 평가할 수 없는 측면을 많이 갖고 있는 교육을 시험 성적 수치만에 의해 개혁하려는 과욕이 그런 실망스런 결과를 가져온 것이다.

신자유주의적 방식에 의한 교육개혁 시도가 가져온 가장 비극적인 결과는 가장 깨끗해야 할 교육부문에 부정행위가 횡행하게 되었다는 점이다. 학생들의 부정행위를 적발해 엄격하게 다스려야 할 학교와 교사가 스스로 부정행위를 저지르는 어이없는 일이 벌어지게 된 것이다. 미국 아틀랜타시에서 조사대상이 된 56개 학교 중 44개 학교에서 성적을 조작한 부정행위가 적발되었다는 것은 문제의 심각성이 어느 정도에 이르고 있는지를 잘 말해 주고 있다. 이는 강력한 유인의 제공이 가져다주는 긍정적 효과를 모두 합친 것을 상쇄하고도 남을 부정적 효과다.

한 가지 흥미로운 점은 우리의 입장에서 미국에서 일어난 일들을 남의 일 보듯 할 수만은 없다는 사실이다. 우리 사회 일각에서도 미국과 비슷한 방식으로 교육개혁을 추진해야 한다는 얘기가 끊임없이 나오고 있기 때문이다. 반대하는 사람이 적지 않음에도 불구하고 현 정부가 일제고사라고 부르는 시험, 즉 국가수준 학업성취도 평가를 강력하게 추진하고 있는 이유가 과연 어디에 있을까? 아직까지 이 시험과 관련된 구체적인 청사진은 밝히고 있지 않지만 궁극적으로 미국처럼 성과에 기초한 책임 정책을 채택하기 위한 전초작업의 성격을 갖는다고 짐작해 볼 수 있다.

우리나라의 경우에는 시험 성적에 따른 유인이 별로 강하지 않은 상황인데도 벌써부터 부작용의 조짐이 보이고 있다. 예를 들어, 어떤 지역의 일제고사 성적이 유달리 좋아 이를 수상히 여긴 당국이 조사를 해본 결과 성적조작이 행해졌다는 사실을 발견한 바 있다. 또한 일부 학교에서 정규교과 시간에 일제고사를 대비한 문제풀이를 하는 시험맞춤교육이

이루어지고 있다는 실태조사 결과가 나와 있기도 하다. 더 좋은 성적을 올리기 위한 교육청과 학교들 사이의 경쟁이 점차 심해지면서 이와 같은 부작용의 사례도 급격하게 늘어날 것으로 예상할 수 있다. 미국의 사례를 반면교사로 삼아 지금이라도 근본적인 보완책을 마련해야 할 것으로 보인다.

미국의 교육개혁 실험에서 분명하게 드러난 것처럼, 강한 유인을 부여함으로써 공공부문을 개혁한다는 것은 생각하는 것만큼 좋은 방법이 아니다. 특히 유인의 제공이 마치 공공부문의 비효율성을 치료하는 만병통치약이나 되는 듯 과도한 기대를 거는 것은 결코 바람직하지 않다. 공공부문의 개혁에 오히려 더 효과를 발휘할 수 있는 방법은 공공의 이익을 위해 일한다는 자부심을 느끼게 만드는 것일지도 모른다. 유인에 의해 공공부문의 개혁을 추진하려는 시도는 공직자들의 자부심에 상처를 줄 가능성이 크다. 당근과 채찍으로 몰아감으로써 자부심을 뺏어가 버렸기 때문에 도덕의 표상이어야 할 학교와 교사가 부정행위를 저지르기에 이르렀던 것이 아닐까?

신자유주의적 개혁이 기대한 만큼의 성과를 가져오지 못한 데에는 인간 본성에 대한 정확한 이해의 결여가 한몫을 단단히 하고 있다. 요즈음 새로이 각광을 받고 있는 행태경제이론(behavioral economics)에 따르면, 인간은 단순히 자신의 이익을 합리적으로 추구하는 존재가 아니다. 공정성을 위해 때로는 자신의 이익을 서슴없이 던져버리기도 하고, 아무런 이득이 없는데도 단지 보람을 느끼기 위해서 어떤 일에 자신의 노력과 시간을 아낌없이 쏟아 붓기도 한다. 공공부문도 궁극적으로 인간에 의해서 움직여 가는 것이라면 이와 같은 인간의 본성을 정확히 파악하고 거기에 알맞은 개혁의 방식을 찾아야 그 노력이 성공을 거둘 수 있다. 이제는 신자유주의적 개혁에 대한 집착을 버리고 인간의 본성과 부합하는 참신한 개혁의 방안을 찾아 나서야 할 때다.

서울대학교 경제학부 교수
141-746 서울특별시 관악구 관악로 599
전화: (02)880-6377
팩스: (02)886-4231
E-mail: joonklee@snu.ac.kr

참 고 문 헌

Baker, George(1992): "Incentive Contracts and Performance Measurement," *Journal of*

- Political Economy*, **100**, 598-614.
- Chakrabarti, Rajashri(2007): “Vouchers, Public School Responses, and the Role of Incentives: Evidence from Florida,” *Federal Reserve Bank of New York Staff Report*, **306**.
- Cullen, Julie, and Randall Reback(2006):“Tinkering Toward Accolades: School Gaming under a Performance Accountability System,” mimeo.
- Dixit, Avinash(2002): “Incentives and Organizations in the Public Sector: An Interpretative Review,” *Journal of Human Resources*, **37**, 696-727.
- Figlio, David(2006): “Testing, Crime and Punishment,” *Journal of Public Economics*, **90**, 837-851.
- Figlio, David, and Lawrence Getzler(2002): “Accountability, Ability and Disability: Gaming the System,” NBER Working Paper Series, **9307**.
- Figlio, David, and Joshua Winicki(2005): “Food for Thoughts: The Effects of School Accountability Plans on School Nutrition,” *Journal of Public Economics*, **89**, 381-394.
- Greene, Jay(2000): “The Texas School Miracle Is for Real,” *City Journal*, **10**.
- _____(2001): “An Evaluation of the Florida A-Plus Accountability and School Choice Program,” Center for Civic Innovation at the Manhattan Institute.
- Haney, Walt(2000): “The Texas Miracle in Education,” *Education Policy Analysis Archives*, **8**.
- Hoff, David(2000): “As Stakes Rise, Definition of Cheating Blurs,” *Education Week*, June 21, 2000.
- Hoffman, James *et al.*(2001): “High-Stakes Testing in Reading: Today in Texas, Tomorrow?” *Reading Teacher*, **54**, 482-492.
- Holmstrom, Bengt, and Paul Milgrom(1991): “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design,” *Journal of Law, Economics, & Organization*, **7**, 24-52.
- Jacob, Brian(2005): “Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools,” *Journal of Public Economics*, **89**, 791-796.
- Jacob, Brian, and Steven Levitt(2003): “Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating,” *Quarterly Journal of Economics*, **118**, 843-877.
- Kane, Thomas, and Douglas Staiger(2002): “Volatility in School Test Scores: Implications for Test-Based Accountability Systems,” *Brookings Papers on Education Policy*, 235-283.
- Klein, Stephen *et al.*(2000): “What Do Test Scores in Texas Tell Us?” *Issue Paper*, Rand

Education.

- Koretz, Daniel(2002): "Limitations in the Use of Achievement Tests as Measures of Educators' Productivity," *Journal of Human Resources*, **37**, 752-777.
- Leibenstein, Harvey(1966): "Allocative Efficiency vs. "X-Efficiency"," *American Economic Review*, **56**, 392-415.
- May, Meredith(2000): "State Fears Cheating by Teachers - 51 Schools Left Off Cash Award List," *San Francisco Chronicle*, October 4, 2000.
- Prendergast, Canice(1999): "The Provision of Incentives in Firms," *Journal of Economic Literature*, **37**, 7-63.
- Rouse, Cecilia *et al.*(2007): "Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure," NBER Working Paper, **13681**.
- Wholey, Joseph(1999): "Performance-Based Management: Responding to Challenges," *Public Productivity & Management Review*, **22**, 288-307.
- Wilson, James(1989): *Bureaucracy: What Government Agencies Do and Why They Do It*, New York, Basic Books.

Abstract

The Ideal and Reality of Neo-Liberal Reform Efforts

Joon Koo Lee

It is very difficult to see an example of a country which has been remarkably successful in the task of reforming the public sector. A variety of ways to reform the public sector have been tried, but no one has proved to be a magic bullet. Recently, a new move which tries to reform the public sector by providing performance-based incentives has been in fashion. Supporters of the policies utilizing performance-based incentives argue that we can make bureaucrats work harder voluntarily by subjecting them to powerful incentives. One good example of such policies is performance-based accountability system which has been the main backbone of educational reform in the United States. Many states had been experimenting with this neo-liberal reform agenda since the 1980s, and the federal government introduced a very similar system by enacting No Child Left Behind Act. However, the introduction of performance-based accountability system has failed to revolutionize the American education. Critics point out that such a system creates incentives to raise test scores per se, not to improve student learning. It has been observed that teachers and administrators use various kinds of methods to raise test scores. They are observed even to cheat in order to raise students' test scores. In view of this somewhat disappointing results of performance-based accountability system, we can conclude that neo-liberal reform could not be a good medicine for the chronic ills of the public sector.

Keywords: Incentive contract, Performance-based incentive, Performance-based accountability, No Child Left Behind, Teaching to the test