

가중치에 대한 제약 및 분석 표본 수에 따른 DEA 효율성과 참조집단의 변화에 대한 실험연구

오동일

상명대학교 금융보험학부 부교수
(odongil@smuc.ac.kr)

DEA 모형에서는 의사결정단위가 가장 효율적으로 평가받을 수 있도록 내생적으로 상대적 가중치가 설정되므로 평가의 임의성이 배제되고 객관성이 달성된다.

본 연구에서는 가중치에 대한 제약과 분석 표본 수에 따라 DEA 효율성과 참조집합의 변화가 크다는 점과 상대적 효율성이 참조 집합의 변화와 유의한 관계를 지닌다는 점을 실험 자료를 생성하여 분석하였다.

DEA 효율성은 콤팩트글라스 함수 상의 진실한 효율성 이외에도 제약조건과 분석 표본 수에 의존하고 있음을 알 수 있었고, 제약 조건이 있는 경우와 없는 경우 모두 표본 수의 증가에 따라 과대 평가 비율은 감소하나 제약 조건이 있는 경우가 감소 폭이 더 크다는 것도 확인되었다.

또한 참조 집합의 변화를 로짓 함수를 이용하여 분석한 결과 분석 표본 수와 제약 조건에 민감한 반응을 보인다는 점과 제약 조건이 부과된 경우 참조 집합의 수는 평균에 가까이 분포하는 반면 제약 조건이 없는 경우에는 비교적 넓게 분포한다는 것도 알 수 있었다.

가중치에 무한한 탄력성을 부여하거나 분석 표본 수가 작은 경우에도 DEA 모형이 여전히 유용하지만 가중치에 대한 주관적 가치를 도입하고 분석 표본 수를 증가시킴으로써 모형의 의미는 더욱 증가될 것이다.

1. 서 론

DEA 모형은 의사결정단위의 상대적인 효율성을 측정해주는 비모수적인 방법으로 대학, 병원, 은행, 증권회사, 공기업 등 다양한 의사결정단위의 상대적인 효율성을 제공해주는 실용적 모형이다.

DEA 모형에서는 다투입·다산출 구조하에서 산출물의 가중합을 투입물의 가중합으로 나눈 비율의 최대값을 특정 단위의 효율성으로 정의한다. 그리고 투입과 산출의 가중합을 구하기 위한 가중치 승수가 모형 내에서 부여된다. 이러한 특성으로 인해 DEA 모형에서는 분석자가 인위적으로 투입물과 산출물의

가중치 승수를 부여하는 자의적 판단을 제거함으로써 분석 결과의 객관성을 증대시켜 준다.

그러므로 DEA 모형은 의사결정단위가 단 하나의 목적만을 지향하지 않는 경우 다목적 구조를 어떻게 모형 속에 반영할 것인가 하는 것과 다양한 투입물과 산출물이 가지는 가치에 대한 주관적 평가로 인한 평가결과의 임의성이라는 두 가지 면을 해결해 준다.(Charnes, Cooper, Lewin and Seiford, 1994) 그러나 가중치에 대한 탄력성을 무한정 부여함으로써 사전적으로 취득된 정보를 누락한다던가 하는 문제와 투입-산출 구조상에 가장 중요한 요소는 분석에서 누락되고 의사결정 단위가 강점을 가지는 특성의 투입물이나 산출물에 의해서

DEA 효율성 값이 좌우되는 단점이 있다.

가중치의 선정이 의사결정단위에 가장 유리하게 선정되므로 특정의 단위를 평가하기 위해서 선정된 가중치가 다른 단위를 평가하기 위해 선정된 가중치와 상당한 차이를 보일 수 있다. 이런 문제점으로 인해서 가중치 선택에 제약을 어떻게 줄 것인가에 대한 여러 연구가 있다.(Banker and Morey, 1986; Dyson and Thanassoulis, 1988; Roll, Cook and Golany, 1991; Thompson, Langemeier, Lee and Thrall 1990)

투입·산출물의 변화가 DEA 효율성에 미치는 영향과 관련해서는 변수에 대한 측정 오류에 효율성이 민감하게 반응한다는 비판과 관련한 연구(Thompson 등, 1995), 효율적으로 판정된 의사결정단위가 투입·산출 변수의 추가적인 변화에도 불구하고 효율성이 그대로 유지될 수 있는 범위가 어디까지인 가하는 모형의 안정성과 관련된 연구(안태식, 1990; Zhu, 1996; Seiford and Zhu, 1998)가 있다. 또한 3 개국 전력사업자의 효율성을 비교하는 과정에서 전력사업자 수의 차이를 조정하기 위하여 표본수를 조정한 효율성분석과 관련한 연구(Zhang and Bartles, 1998)가 있다.

위와 같은 맥락에서 본 연구에서는 전형적인 콤팩트 다글라스 생산 함수로부터 도출된 실험 자료를 이용하여 가중치에 대한 단순한 제약 조건의 부과로 인해서 DEA 효율성이 받는 영향의 정도와 DEA 효율성 결정과 밀접한 관련을 가지고 있는 참조 집단이 제약 조건의 부과로 인해서 상당한 영향을 받는다는 사실을 밝힌다. 그리고 이러한 영향의 정도는 분석 대상 표본 수와 밀접한 관련을 가지고 있다는 점을 보이고자 한다.

이를 위해 우선 DEA 모형이 가지는 일반적인 특징을 살펴보고 이를 토대로 가중치에 제약 조건을

부여하여야 할 필요성을 살펴본다. 그리고 가중치에 제약 조건이 미치는 영향을 파악하기 위해서 100개의 가상적 의사결정단위를 생성 후 분석 대상 표본수를 변화시키고 제약 조건을 추가함으로써 효율성 값의 변화를 살펴본다. 마지막으로 DEA 모형에서 참조 집단의 변화가 DEA 효율성 지표의 영향을 제거한 후에도 모형에 부과되는 제약 조건과 분석 대상이 되는 의사결정단위의 수에 의해 영향을 받는다는 점을 밝히고자 한다.

II. DEA 효율성의 특성과 제약 조건의 역할

2.1 DEA 모형 구조와 특성

DEA 모형은 경제 내에 n 개의 단위가 존재하고 각 단위는 m 개의 투입 요소를 사용하여 s 개의 산출물을 생산하는 다음과 같은 조건을 만족시키는 생산 가능 집합을 전제로 한다.

- $(x_j, y_j) \in T$ 이고 $\sum \lambda = 1, \lambda \geq 0$ 이면 $(\sum x_j, \sum y_j) \in T$
- 만약 $(x_j, y_j) \in T$ 이고 $x' \geq x$ 이면 $(x', y_j) \in T$ 이다.
또한 $(x_j, y_j) \in T$ 이고 $y' \leq y$ 이면 $(x_j, y') \in T$ 이다.
- $(x_j, y_j) \in T$ 이면 $(kx_j, ky_j) \in T$ (단, $k \geq 0$)
- $T = \cap T'$ (단, T' 는 위의 조건을 만족하는 집합)

위와 같이 정의되는 생산 가능 집합의 가정으로부터 Farrell 은 다음과 같은 투입 척도(h)를 제시하였다.(Farrel, 1957)

〈식 1〉

$$h(x, y) = \min h$$

$$s.t \quad hX \geq k \sum_j \lambda x_j, \quad y \leq k \sum_j \lambda y_j, \quad \sum_j \lambda = 1, \quad \lambda, k \geq 0$$

이 식을 변형한 후 상대형을 취하면 다음과 같은 분수계획법의 일종인 CCR 모형이 된다.

$$\langle \text{식 2} \rangle \quad \text{Max } h_0 = \frac{\sum_{r=0}^s u_r y_{r0}}{\sum_{i=1}^m v_i x_{i0}}$$

$$\text{subject to } \frac{\sum_{r=0}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1, \quad u_r, v_i > 0$$

$$j = 1, 2, \dots, n, \quad r = 1, 2, \dots, s, \quad i = 1, 2, \dots, m$$

- y_{rj} : j번째 단위가 산출한 r 번째 생산물
- u_r : r 번째 산출물에 주어진 가중치
- x_{ij} : j번째 단위가 소비한 i 번째 투입물
- v_i : i 번째 투입물에 주어진 가중치

Charnes, Cooper 그리고 Rhodes는 위의 식에 작은 양수(ϵ)를 도입하여 효율성 척도로서의 h값을 도입하여 선형계획모형을 구성하였는데, 이 선형계획모형의 상대형을 취하면 CCR비율모형으로 불리는 다음과 같은 DEA 모형의 기본형이 구해진다.(Charnes, Cooper and Rhodes, 1978)

〈식 3〉

$$h(x, y) = \min \theta - \epsilon [\sum_{i=0}^m s_i + \sum_{r=0}^s s_r]$$

$$s.t \quad \theta x_{i0} - \sum_j \lambda_j x_{ij} - s_i = 0$$

$$y_{r0} = \sum_j \lambda_j y_{rj} - s_r$$

$$(i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s,$$

$$\lambda_j \geq 0, \quad s_r \geq 0, \quad s_i \geq 0)$$

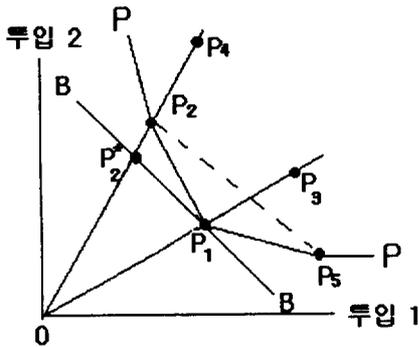
이 모형을 이용해서 해를 구하면 의사결정단위의 효율성이 구해지는데 의사결정단위가 효율적이라고 판정되기 위해서는 다음의 조건이 만족되어야 한다.(Banker, Charnes, Cooper and Swarts and Thomas, 1988)

$$\langle \text{식 4} \rangle \quad \theta = 1 \quad \text{이고} \quad \forall s_i, \quad \forall s_r = 0$$

전통적인 비율분석에서는 외생적으로 이미 결정된 가중치인 u_r, v_i 를 사용하는 것과는 달리 DEA 모형에서는 λ_j 결합에 의해서 최적 단위를 구성하고 이에 따라 특정 단위의 성과가 평가된다. 따라서 각 의사결정단위가 피평가 단위로 선택될 때마다 서로 상이한 최적의 가중치, 참조 집합 그리고 참조 집합에 대한 상대적 중요성을 개별적으로 구하게 되고, 이런 수치들의 집합은 각 의사결정단위마다 상이하다. 따라서 DEA 모형은 특정의 의사결정단위를 가장 잘 묘사할 수 있는 참조 집단이 선택되어 효율성이 평가됨으로써 특정의 의사결정단위가 효율적이 되기 위해 필요한 달성 가능한 실천 방안을 제시해 줄 수 있다.(Charnes, Cooper, Lewin and Seiford, 1994; Sherman, 1984; Seiford, 1995)

두 가지의 투입물을 사용하여 하나의 산출물을 생산하는 등생산량 곡선과 예산선으로 이루어진 〈그림 1〉을 살펴봄으로써 DEA 모형의 특성을 보다 잘 알 수 있다.

〈그림 1〉 등생산량곡선과 DEA 효율성



(단, \overline{BB} : 예산선, \overline{PP} : 등생산량곡선)

우선 점 P_1 과 P_2 는 모두 등생산량곡선 상에 위치하므로 〈식 1〉로 주어지는 파렐척도의 관점에서 효율적이다. 만약 투입물의 단위 원가가 각각 v_1^* , v_2^* 라고 하면 예산선의 기울기는 $-v_1^*/v_2^*$ 인 직선 \overline{BB} 이다.

〈식 2〉를 살펴보면 목적 함수의 분모 부분이 $\overline{BB} = \sum_{i=1}^n v_i x_{i0}$ 으로 이는 평가 대상 단위 0이 산출을 위해서 소비한 예산이다. P_1 은 등생산곡선 뿐만 아니라 예산선상에도 존재하므로 기술적인 측면과 배분적(가격적) 측면에서 모두 효율적이다. 그러나 P_2 는 등생산량곡선 상에는 위치하나 예산선상에는 위치하지 않으므로 기술적 측면에서는 효율적이거나 배분적 측면에서는 비효율적이다. 한편, P_3 은 기술적인 측면에서는 비효율적이거나 P_1 과 같은 투입 산출의 배합 구조를 가지고 있으므로 배분적 측면에서는 효율적이다. 그러나 P_4 는 등생산량곡선내의 점으로 등생산량곡선의 접선과 원점을 지

나는 확장 선이 만나는 선상에 존재하지 않으므로 배분적 측면에서도 비효율적이다. 따라서 P_4 의 전체적인 효율성은 기술적인 효율성과 배분적 효율성의 곱으로 표현 가능하다. 〈식 3〉의 DEA 모형에서 P_4 의 전반적 효율성과 기술적 효율성은 다음과 같은 관계를 가진다.

전체적 효율성 = 기술적 효율성¹⁾ × 배분적 효율성

$$OP_2^*/OP_4 = OP_2/OP_4 \times OP_2^*/OP_2$$

P_2 의 예에서와 같이 기술적인 측면에서는 비효율적인 단위가 배분적 측면에서 효율적일 수도 있고 그렇지 않을 수도 있음을 알 수 있다. 그리고 〈식 3〉에서 나타난 것처럼 P_3 의 참조집합은 P_2 와 P_5 으로 구성된다. P_1, P_2, P_5 는 프론티어상에 존재하므로 각자가 스스로의 참조집합이 된다. 그리고 P_4 의 참조집합은 또 다른 점들로 구성될 것이다. 그러므로 DEA 모형의 효율성은 다음과 같은 특징을 가진다. (Thompson, Langemeier, Lee, Lee, and Thrall, 1990)

특정한 의사결정단위만을 참조집합으로 선택해서 구한 DEA 효율성은 모든 단위를 비교해서 구한 상대적 효율성이 아니라 국지적인 상대적 효율성을 나타낸다. 또한 DEA 효율성은 현재의 투입·산출 자료를 효율적인 것과 비효율적인 것으로 구분해주므로 효율적인 단위라 하더라도 그 효율성은 절대적인 효율성이 아니고 의사결정단위가 최선의 노력을 기울여서 달성할 수 있는 최적의 효율성을 나타내는 것이 아니다.

극단적으로는 의사결정단위가 다수의 투입 산출 요소 중 특정한 단 하나의 투입·산출물에 대해서

1) 기술적 효율성은 다시 순수한 기술적 효율성과 규모의 효율성의 곱으로 표현 가능하다. (Banker, Charnes, Cooper, Swarts and Thomas, 1988)

만 모든 다른 단위보다 우월한 지위에 있게 된다면 효율적인 것으로 판명되어 효율성 지표는 1을 나타내게 된다. DEA모형은 특별히 타 단위에 비하여 효율적인 요소에 대해 높은 가중치를 부여하고 다른 투입물이나 산출물에 대해서는 낮은 가중치를 부여하는 특성이 있다. 즉 특이점이나 극단치의 경우에는 자동적으로 효율적인 것으로 분류될 수 있는데 이는 그 단위가 수행하는 모든 다른 활동들을 무시한 결과를 초래한다.

$v_{xij} / \sum_{i=1}^m v_{xij}$ 은 의사결정단위 j 의 투입물 i 가 다른 투입물에 대해 가지는 상대적인 중요성을 나타낸다고 볼 수 있다. DEA 모형에 의해 구해진 가중치가 의미를 가지려면 중요하다고 판단되는 투입물과 산출물에 대한 가중치가 0 아닌 값을 가져서 총비용이나 총효익에 적어도 일정 부분을 설명할 수 있어야 한다. 만약 산출물을 고정시켜 놓고 생각한다면 특정 요소가 예산의 일정 부분을 설명할 수 있어야 한다.(Chaparro, Jimenez and Smith, 1997)

투입물이나 산출물중 0의 가중치를 가지는 것이 있다면 이 요소는 효율성 측정에 반영되지 않았다는 것이고 특정의 요소를 반영하지 않은 상태에서 결정된 효율성은 그 의미가 반감된다. 역으로 0의 가중치가 부여된 투입물이나 산출물이 정말 중요하지 않은 요소라면 이런 요소들을 분석에 포함시켜 모형을 구성한 것은 모형 설정의 오류라고 할 수 있다.

2.2 가중치에 대한 제약 조건의 필요성

이와 같은 DEA 모형의 특성은 가상적 가중치인 v_i , u_r 또는 λ_j 에 대한 모형의 자유로운 선택에 기인한다. 그런데 이런 자유로운 선택은 위에서 살펴본 특성과 함께 다음과 같은 문제점을 지닌다.(Banker

and Morey, 1989; Chaparro, Jimenez and Smith, 1997)

이차적으로 중요한 요소들이 단위의 효율성 지표를 결정하게 된다. 만약 분석에 사용되고 있는 투입물이나 산출물의 중요성이 동일하지 않다면 중요한 투입물과 산출물에 0의 가중치가 주어진 단위가 상대적으로 효율적이라고 판단하는 것은 비상식적이다. 가중치에 대한 전적인 탄력성은 그 단위가 가장 중요하게 고려하여야 하는 투입물과 산출물에 내재된 비효율성을 포장하면서 별로 중요하지 않은 투입물과 산출물의 효율적인 사용을 강조하는 잘못된 결과를 초래한다.

분석에서 중요한 요소가 무시될 수 있다. 단위를 분석함에 있어서 어떤 투입물과 산출물은 평가에서 고려되어지지 않을 수 있다. 그 결과, 단위의 상대적인 효율성이 투입물과 산출물을 모두 다 같이 고려하였을 때의 결과와는 상이하다. 분석을 진행하는 과정에서 모형 내에 포함되어야 할 투입물과 산출물을 그 중요도에 따라서 신중하게 결정하였음에도 불구하고 DEA 분석 결과 이 요소들에 대한 가중치가 0으로 주어진다면 변수 선택 과정에서 그 변수의 선택이 잘못 되었거나 분석의 결과가 잘못된 것이므로 결과를 수용하기가 어렵다.

가중치에 완전한 탄력성을 부여하는 DEA 모형의 암묵적인 가정은 피평가 단위는 분석 시에 반드시 고려해 주어야 하는 자신만의 고유한 목적과 환경을 가지고 있다는 것이다. 그러나 일반적으로 DEA 모형은 동질적인 피평가 단위를 분석하는데 사용하는 도구로 알려져 있으므로 각 단위가 사용하는 투입물과 산출물에 각각 상이한 가중치를 부여하고 그 중요성이 서로 다르다고 가정하는 것은 받아들이기 힘들다. 비록 DEA 모형에서 각 단위의 특이한 환경적 요소와 특성을 반영해서 각 단위

에 주어지는 가중치에 어느 정도의 탄력성을 부여하는 것이 바람직할 수 있다 하더라도 각 단위마다 비교 집단이 다르게 되고 가중치의 값이 상이하다는 것은 설득력이 약하다. 가중치에 대한 완전한 탄력성은 각 단위가 비슷한 생산 구조와 투입물에 대한 유사한 가격 그리고 비슷한 생산물을 산출하는 상황으로 볼 때 상식에 어긋난다.

마지막으로 어떤 경우에는 투입물과 산출물에 대한 가중치와 관련된 정보가 있을 수가 있다. 이런 경우 상대적인 효율성의 도출과 관련된 정보를 사용하는 것이 더 의미가 있다.(Roll, Cook, and Golany, 1991) 이런 정보를 이용함으로써 분석되는 단위의 상이한 산출물에 대한 상대적인 중요성과 그 산출물을 생산하기 위한 투입물의 상대적 기회비용에 대한 가치판단을 수행할 수 있다. 제약식을 가하는 것이 DEA 모형이 가지는 자유로운 객관성에 대한 비난의 대상이 될 수 있지만 무제한 DEA 모형의 결과로부터 도출되는 가중치는 가격 비효율성에 대한 판단을 전혀 하지 않는 결과를 초래한다. 그러나 가격 정보가 있는 경우에 이것을 무시하는 것은 받아들이기 어렵다.

의사결정단위 j 가 투입물 i 와 관련하여 지출하는 총비용은 $v_i x_{ij}$ 이다. 그러므로 DEA 모형에서 이 투입물을 포함시킨다는 것은 이 요소가 비용의 상당 부분을 설명할 수 있다는 것을 가정하는 것이다. 무제한 하의 DEA 모형에서 구해진 가중치가 이 단위의 비용을 설명하기에 충분하지 못한 요소라고 드러나면, 이 투입물이 분석에 포함되어서는 안되겠지만 만약 모형에 포함시키는 것이 합리적이라면 이 요소에 대한 가중치(v_i)의 값이 증가되어야 한다. 이는 결국 가중치에 제약을 가해야 한다는 점을 의미한다. 물론 가중치에 대한 명시적인 판단은 효율성 예측이 투입물과 산출물의 상대적인

중요성과 관련하여 정의되어야 한다. 공공 분야와 같이 산출물 선정이 어렵고 상대적인 가중치 부여는 더 어려운 경우에는 제약 조건을 부과하는데 어려움이 있을 수 있다. 이런 경우에는 DEA 결과를 해석함에 있어 주의를 요한다.

의사결정단위의 개별적인 목적과 환경을 반영해 준다는 점과 가중치에 대한 정확한 정보를 얻기가 힘들다는 점을 고려한다면, 가중치가 어느 정도 탄력성을 가지는 것은 바람직하지만 앞에서 살펴본 DEA 모형의 특성으로 인해 완전한 탄력성은 각 단위의 배분적 비효율성을 감출 수 있다. 가중치에 대한 정확한 정보가 없는 경우에는 가중치가 가지는 값의 범위에 대한 제약을 가하는 방법을 고려해 볼 수 있다. 즉 모형 내에서 가중치가 변화될 수 있는 한계를 정함으로써 평가 대상이 되는 의사결정단위가 더 이상 비효율적인 투입을 감출 수 없게 되고 가장 중요한 투입물과 산출물에 대한 가중치가 중요성이 떨어지는 투입물과 산출물의 가중치보다 더 큰 값을 가지게 된다. 가중치의 한계를 부여하기 위해서는 <식 2>나 <식 3>에 추가적인 제약식을 부과하면 된다. 모든 투입물과 산출물의 가상적 승수에 하한 값과 상한 값을 부여한다면 단위는 더 이상 자신이 비효율적으로 운영하는 요소에 대해서는 0의 값을 부여하고 자신이 효율적으로 운영하는 요소에 대해서는 최고의 가중치를 부여할 수 없게 된다. 그러므로 제약 모형에서는 추가된 제약식이 구속적 역할을 수행하지 못하는 경우에만 동일한 결과가 나오므로 제약 모형에서의 효율성 지표의 값은 무제한 모형에 비하여 같거나 작게 될 것이다.

이상과 같은 관점에서 DEA의 장점을 살리면서 주관적인 범위를 정하기 위해 가중치에 제약 조건을 부과하는 모형이 사용된다. 가중치에 제약을 가

함으로써 가중치를 찾는 영역의 범위가 줄어들고 이에 따라 각 단위의 효율성 값은 증가하지 않고 감소하게 된다.

극단적으로는 만약 DEA 모형에 가중치의 탄력성을 전혀 주지 않는다면 이 모형은 위에서 언급된 대로 투입물과 산출물의 가격에 동일한 가중치를 두고서 산출물의 가중 합을 투입물의 가중 합으로 나누는 전통적인 비율 분석이 된다. <그림 1>에서 이런 척도는 다른 단위에 대해 상대적으로 측정되는 전반적인 효율성과 동일하다. 그러므로 가중치에 더욱 강한 제약 조건을 부과할수록 효율성 척도는 상대적인 효율성을 측정하는 지표로부터 전반적인 효율성을 측정하는 지표로 변화하게 된다.

다음 <표 1>에서 알 수 있는 바와 같이 DEA 모형의 장점은 가중치가 무한한 탄력성을 가진다는 점이다. 이로 인해서 평가 단위마다 상이한 자유로운 가중치를 가진다. 가중치에 사전적으로 값을 부여하지 않음으로써 비효율적으로 측정된 단위가 "자신이 비효율적이라고 측정된 것이 투입물과 산출물에 주어진 잘못된 주관적인 가중치 때문이다." 라고 주장할 수 없게 된다. 자유롭게 가중치가 주어지는 상황에서도 비효율적으로 측정된 단위의 경우 비효율성이 더욱 부각되는 장점이 있다. 투입물 사이의 시장가치에 대한 정보를 얻을 수 있거나 산

출물간의 선호도가 주어진 경우에는 가중치의 범위에 제한을 가할 필요가 있다. 만약 가중치를 자유롭게 부여하는 경우에는 특정의 단위는 적절하다고 판단되는 어떤 단위들에 의해서만 평가된다. 더군다나 가중치를 자유롭게 부여하는 경우에 산출은 낮고 투입은 많아서 부적격한 단위임에도 불구하고 그러한 부적격을 감출 수 있게 된다. 이런 경우에는 무제약 하의 DEA 모형을 이용해서 가중치를 구한 후 다시 요소의 한계를 설정하여야 한다.

III. 실증분석

가중치에 대한 제약과 분석 표본 수에 따라 DEA 효율성이 민감하게 영향을 받을 수 있다는 점과 참조 집합의 크기가 DEA 효율성 이외에도 제약 조건 유무나 분석 대상 표본 수와도 밀접한 관계를 가지고 변동한다는 점을 실험 설계를 통하여 알아본다.

3.1 표본 설계

여기서는 진실한 생산 함수가 콥다글라스형의 모양을 나타낸다고 가정하고 자료를 생성한다. 물론

<표 1> 가중치에 대한 제약이 없는 DEA모형과 제약이 있는 DEA모형의 비교

분류	특성	효율성의 종류	참조 집합	모형의 탄력성	평가의 주관성
무제약의 DEA모형	양의 제약 조건만 있음	기술적 효율성만 측정	특정의 단위로 한정	모형의 탄력성이 큼	주관적 요소 고려하지 않음
제약하의 DEA모형	위의 조건 + 가중치에 대한 제약조건추가	기술적 효율성 + 배분적 효율성 측정	특정의 단위로 한정되거나 축소가능성 존재	모형의 탄력성 유지	주관적 요소 일부 고려
전통적 비율분석모형	가중치에 대한 제약이 경험적으로 결정됨	배분적 효율성 측정	모든 단위가 참조집합 됨	모형이 경직됨	매우 주관적임

DEA 모형은 생산 함수의 모양이 알려져 있지 않은 경우에 더욱 유용한 모양이긴 하지만 생산 함수의 모양을 전혀 모르는 경우에는 DEA 효율성 지표를 비교할 수 있는 수단이 없으므로 가장 다루기 쉬운 함수형을 선택하였다.

DEA 모형이 가중치로부터 받는 영향의 정도와 참조 집합의 변화를 보기 위하여 다음과 같은 조건을 만족하는 100개의 실험 자료를 생성한다. 우선 규모에 대한 보수가 일정한 콤팩트글라스 생산 함수를 구성한다.

$$\langle \text{식 6} \rangle Y_j = \beta \prod X_{ij}^{\alpha_i} \eta_j, \quad \sum_{i=1}^m \alpha_i = 1$$

위 식의 양변에 로그를 취하여 정리하면 다음과 같은 함수 모양을 얻을 수 있다.

$$\langle \text{식 7} \rangle y_j = b + \sum_{i=1}^m \alpha_i x_{ij} + e_j$$

(단 $y_j = \ln(Y_j)$, $x_{ij} = \ln(X_{ij})$, $e_j = \ln(\eta_j)$)

생산물을 산출하기 위해서 사용되는 투입 변수는 가능한 한 서로 독립적이고 산출에 기여하는 정도가 큰 변수를 선정하는 것이 좋다. 그래야만 다중공선성의 문제도 없고 선택된 변수의 설득력도 높아질 것이다. 그러나 대부분의 실증적인 연구에 의

하는 경우 보다 현실적인 모형을 구성하기 위해서 변수들간의 상관관계를 비교적 높은 것으로 관찰되고 있으므로 변수들간의 상관관계를 높게 잡고 있다.²⁾ 여기서는 4개의 투입물을 이용해서 하나의 산출물을 생산하는 경우를 분석하되 네 가지 투입요소를 상관관계가 높은 투입 요소와 상관관계가 낮은 투입 요소로 배분해서 실험 변수를 구성한다.

〈식 2〉의 변수 e_j 는 교란항으로 이 변수의 크기에 따라 원하는 산출량보다 더 낮은 수준의 생산이 이루어질 수 있다. 따라서 변수 e_j 는 변수 x_{ij} 와는 상관관계가 없고 (-1, 0) 사이에서 움직이는 균등분포를 하는 확률 변수로 한다.³⁾ 80개의 e_j 는 Excel의 RAND함수에 의해서 생성하고 나머지 20개의 표본은 인위적으로 $e_j = 0$ 을 부여함으로써 총 100개의 표본 중 20%는 효율적인 단위로 구성하였다. 이렇게 인위적으로 20%의 의사결정단위를 효율적으로 설정한 이유는 제약식의 부과와 분석 대상 표본 수의 변화에 따른 DEA 효율성 값의 변화를 측정하기 위함이다. 이런 조건을 만족시키는 실험 자료 100개를 생성한 결과 각 분석에 사용될 변수간의 분산공분산행렬은 〈표 2〉와 같다.

위와 같은 분산공분산 구조를 취하는 (X_i, e_i)

〈표 2〉 독립변수와 잔차항간의 분산공분산행렬

변수	X ₁	X ₂	X ₃	X ₄	η
X ₁	1	0.1726(0.9096)	-0.0102(-0.1152)	0.0972(0.7815)	-0.0045(-0.047)
X ₂	0.1726(0.9096)	1	-0.0168(-0.087)	0.1870(0.6877)	-0.0098(-0.0482)
X ₃	-0.0102(-0.1152)	-0.0168(-0.087)	1	-0.0138(-0.1090)	0.0022(0.0225)
X ₄	0.0972(0.7815)	0.1870(0.6877)	-0.0138(-0.1090)	1	-0.0000(-0.0005)
η	-0.0045(-0.047)	-0.0098(-0.0482)	0.0022(0.0225)	-0.0000(-0.0005)	1

()안의 수치는 상관계수

2) Beasley(1995)는 대학 학과 투입물간의 상관관계가 0.66에서 0.92임을 보이고, Chaparro, Jimenezr and Smith(1997)는 투입물간의 상관관계를 0.85로 두고 있다.

3) 이와 같은 생성을 통해서 η(기술적 효율성)는 .37 ~ 1 사이에 분포하게 만들.

를 생성한 후 4 개의 투입물이 y_i 생산에 기여하는 정도를 동질적인 것으로 가정하고 α 의 값을 1/4로 두며, β 는 계산의 편의를 위해서 1로 둔다. <식 7>을 이용하여 가장 효율적인 경우의 산출물 y_i^* 와 비효율이 고려된 실제 관찰된 산출물 y_i 의 값을 구한다. <식 7>에서 사용된 y, x 의 값에 지수를 취한 후 이 값을 DEA 모형의 투입 산출 자료로 이용한다. 그리고 η 는 DEA 모형과 비교되는 진실한 효율성 지표이다.

DEA 효율성 지표가 분석 대상 표본 수에 매우 민감한 영향을 받는다는 점을 규명하기 위해서 처음 구성된 100개의 표본을 다시 50개씩 묶은 표본 2개와 20개씩 묶은 표본 5개로 하위 표본을 구성하였다. 그리고 각 표본은 다시 제약 조건이 부과된 경우와 제약 조건 부과되지 않은 경우의 2가지로 나눈다.

따라서 전체적으로는 $(100 + 50 \times 2 + 20 \times 5) \times 2 = 600$ 개의 표본이 존재하게 되어 600개의 선형 계획 모형이 구성되고 이 모형의 해를 구함으로써 DEA 효율성을 구한다. 가중치에 대한 제약 조건으로는 투입 요소에 대한 불완전한 가격 정보를 반영 해주는 $v_1 \leq \beta_{1j} v_j$ 의 제약 조건을 추가하고 $\beta_{11} = 1, \beta_{12} = 1, \beta_{13} = 2, \beta_{14} = 3$ 으로 둔다. 이 제약 조건은 단지 첫 번째의 투입 요소의 가상적 가중치가 다른 투입 요소에 비해서 상대적으로 크다는 것만을 기술하고 있으므로 강한 형태

의 제약 조건이라 할 수 없다. 이렇게 약한 조건을 도입한 이유는 가중치에 대한 약한 제약 하에서도 DEA 효율성이 영향을 받을 수 있고 이에 따라 무 제약 하의 효율성과 제약 하의 효율성이 상당한 차이를 보일 수 있다는 점을 보이기 위함이다.

3.2 실증 분석 결과

3.2.1 제약 조건이 있는 경우의 DEA 효율성 변화

우선 DEA 효율성 지표의 값과 실험에 사용한 콤팩트글라스 함수 상의 효율성 값을 비교해 보면 효율성 값의 크기에 있어서는 서로가 상당한 차이를 보이고 있다. 그러나 양자 사이의 상관 계수는 0.919로서 두 값 사이의 움직임이 매우 밀접하게 연관되어 있다. 이는 DEA 모형에서 도출된 효율성 값을 다른 효율성 측정치와 그 값의 크기의 절대치를 비교하는 것은 의미가 없고 상호간의 증가 또는 감소하는 경향, 값의 비례적 증분, 효율적인 단위와 비효율적 단위를 구분할 수 있는 능력의 정도 등과 관련해서 해석되어야 함을 의미한다. 그리고 DEA 효율성이 비록 상대적인 효율성을 제시해 주지만 절대적 효율성 지표와 매우 밀접한 상관 관계를 가지고 움직이고 있고 그 관계 또한 $p = 0.0000$ 수준에서 유의하므로 DEA 효율성 지표를 절대적 효율성의 대용치로 사용해도 좋을 것이다. 예측된 효율성 척도에 의해서 의사결정단위를

<표 3> DEA 효율성 지표와 절대적 효율성간의 비교

쌍체비교	평균	t통계량	상관계수
DEA 효율성지표값(θ)	75.7	85.68*	0.919*
콤팩트글라스함수상의 절대적 효율성값(η)	71.3	(d.f = 599)	(n=600)

* p 값 = 0.0000 < 0.05

서열화한 결과가 의사결정단위의 진실한 효율성과 깊은 관련을 맺고 있다는 사실은 분석가들이나 정책 결정자가 성과 평가 도구로서 DEA 모형을 사용할 수 있는 근거를 제시해 준다.

위의 <표 3>에서 θ 와 η 간에 밀접한 관계가 유지되고 있음을 알았으므로 이제 가중치의 제약 조건 DEA 효율성에 어느 정도의 영향을 미치는지 알아보기 위하여 회귀 모형을 설정한다. 우선 DEA 효율성이 콤팩트글라스 함수 상에 주어진 효율성과 매우 밀접한 관련을 맺고 있으므로 이 변수를 설명 변수에 추가하고 가중치에 대한 제약 조건을 두 번째 설명 변수로, 그리고 분석 대상의 크기를 세 번째의 설명 변수로 추가한다. 그런데 표본 설계에서 언급된 바와 같이 가중치에 대한 제약 조건과 분석 대상 표본의 크기는 범주형 변수이므로 더미 변수를 포함한 다음과 같은 회귀 모형을 구성한다.

$$\langle \text{식 8} \rangle \quad y_i = \beta_0 + \sum_{j=1}^m \beta_j x_{ij} + \varepsilon_i$$

DEA 효율성 값: y_i

콤팩트글라스 효율성 값 (x_1): $0 \leq x_1 \leq 1$ 인 정량변수

가중치 제약 조건(x_2):

$$\begin{cases} x_2 = 0, & \text{제약조건이 없는 경우} \\ x_2 = 1, & \text{제약조건이 있는 경우} \end{cases}$$

분석대상 표본의 수(x_3):

$$\begin{cases} x_3 = 0, & \text{만약 } n = 100 \\ x_3 = 1, & \text{만약 } n = 50 \\ x_3 = 2, & \text{만약 } n = 20 \end{cases}$$

우선 위의 <식 7>과 같은 선형회귀모형의 타당성을 알아보기 위하여 다중공선성과 자기 상관을 알아본 결과 분산팽창계수(VIF)는 1이고 더빈-왓슨 통계량(D-W)은 2.06이므로 통계적 관점에서 볼 때 다중공선성이나 자기 상관 문제는 무시할 수 있다. 또한 회귀분석 결과 결정계수가 0.854이고 유의확률 0.00으로 선택된 독립변수가 전체적인 관점에서 모형 전반을 설명하는 정도가 높고 모형은 전반적으로 타당하다고 볼 수 있다.

회귀분석에 도입된 설명 변수의 개별적인 유의성은 <표 5>와 같다.

<표 5>에서는 DEA 효율성 값이 상관분석에서 확인된 바와 같이 콤팩트글라스 함수로부터 도출된 효율성(여기서는 실험 모형을 구성하였으므로 진실한 효율성으로 추정할 수 있음)에 매우 강한 영향을 받는다는 것을 보여준다. 그리고 DEA 모형에

<표 4> 선형회귀모형의 전반적 타당성

모형	R ²	F 값	p 값	D-W통계량	VIF
$y_i = \beta_0 + \sum_{j=1}^m \beta_j x_{ij} + \varepsilon_i$	0.854	1164.6	0.000	2.06	1.0

<표 5> 설명 변수의 유의성 검정

변수의 종류	β 의 계수	t 값	p 값
상수항	-212.08	-42.19	0.000
콤팩트글라스효율성값	308.27	58.75	0.000
제약식의 유무	-1.56	-2.37	0.020
분석 표본 수	2.48	6.08	0.000

부과된 제약 식과 분석 표본 수에 의해서도 강한 영향을 받는 것으로 나타난다. 회귀분석의 결과 제약 식의 유무와 관련된 편회계수는 (-)부호를 나타내고 분석 표본 수의 편회계수는 (+)의 부호를 나타내고 있다. 이는 제약식이 도입될수록, 또한 가중치에 대한 제약이 더욱 강해질수록 DEA 효율성은 감소한다는 것과 분석 대상 표본수가 감소할수록 DEA 효율성 값은 크게 증가한다는 것을 보이고 있다. 보다 정밀한 연구는 표본 수를 크게 증가시키고 표본 수를 계속적으로 변동시켜 가면서 보다 정밀한 시뮬레이션을 해 보아야 하겠으나 표본 수가 20개인 경우와 50개인 경우 그리고 100개인 경우 각 각의 효율성 값이 통계적으로 유의한 차이를 보인다는 점은 표본 수가 적은 경우에 DEA 효율성을 해석함에 있어 주의를 필요로 한다는 점을 지적할 수 있다.

위 <표 5>는 단계적 회귀분석법에 의하여 콤팩트라스 효율성 값이 DEA 효율성 값을 결정하는 의미 있는 요인이라는 것을 밝힌 후 여기에 다시 제약 조건과 분석 대상 표본 수를 추가하여 회귀분석한 결과이다. 따라서 전체 표본 중 영향을 미치는 강한 요인으로 이미 절대적 효율성을 모형 내에 고려한 상태이므로 이 영향을 제거한 후에도 제약 조건과 표본수가 DEA 효율성 결정에 영향을 미친다는 점은 매우 의미 있는 결과라 할 수 있다. 엄격한 제약식이 도입될수록 DEA 효율성 지표는 하락한다. 이는 효율적이라고 판단되는 단위의 수가 전체적으로 감소를 의미한다. 따라서 주관적인 가중치나 근거있는 제약식을 도입하여야 함에도 불구하고 이런 요소들은 제외하고 분석하는 경우에는 효율성 지표의 값이 과대평가 되는 경향을 지속적으로 나타날 것이다. 그리고 DEA 효율성이 분석 대상이 되는 표본수가 감소하는 경우에 더욱 증가한

다는 것은 DEA 모형의 특성상 분석 대상 표본 수의 감소는 비교 대상이 되는 참조 집단의 감소를 의미하게 되고 참조 집단의 감소에 따라 의사결정 단위 스스로 만이 자기 자신의 유일한 참조 집합을 구성하게 될 확률이 증가하고 이로 인해서 특정의 DEA 효율성 지표의 값이 증가한다는 것을 의미한다. 따라서 통상적으로 행해지는 작은 표본 단위를 통한 분석 결과만 가지고 특정 단위가 효율적이라고 판정하는 데는 주의가 필요하다. 회귀분석 결과 유의한 것으로 판명된 변수들의 변화에 따른 DEA 효율성 지표의 변화를 <표 6>을 통해 살펴보자.

<표 6>를 통해서 제약 조건 있는 경우의 DEA 효율성이 무제약 하에서의 DEA 효율성에 비하여 과대평가를 하는 비율이 낮음을 알 수 있다. 이 경우 효율적이라고 판단된 단위 수에 있어서는 그 차이가 크게 나타나고 있으나 진실한 효율성에 대한 과대평가가 정도의 비율에 있어서는 양자 사이의 차이가 크지 않은 것을 알 수 있다. 그러나 앞서 살펴보았듯이 진실한 효율성에 대한 과대평가 비율 또한 통계적으로 의미 있는 차이를 보이고 있음을 주목할 필요가 있다.

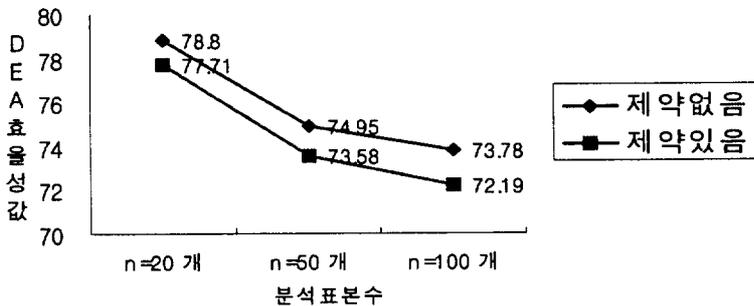
제약 조건 부여된 경우에는 100개의 표본을 분석한 결과 실험 집단의 구성과 정확하게 일치하게 전체 표본 중 20개의 의사결정단위를 효율적이라고 분류하였고 진실한 효율성에 대한 과대평가 비율도 1%에 그치고 있으므로 가중치에 대한 제약 조건의 부과는 DEA 효율성의 질을 상당히 개선시킨다고 볼 수 있다. 이는 제약 조건 없는 경우에 100개의 표본을 구성하더라도 27개 단위가 효율적인 것으로 분류되어 35%의 과대평가를 보이고 DEA 효율성 지표의 평균값도 진실한 효율성에 비하여 3% 과대평가 비율을 보이는 것과는 대조적이다.

〈표 6〉 제약 조건과 표본수에 따른 DEA 효율성의 변화

항목 제약 유무	분석 대상 표본의 크기	각 표본 하에서 DEA 효율적이라고 판정된 단위의 총수(A)	효율적이라고 과대 평가된 단위의 비율 (B)	효율적이라고 과대 평가된 단위의 감소비율 (C)	각 표본 하에서 DEA 효율성의 평균값 (D)	진실한 효율성에 대한 과대평가 정도비율(E)	진실한 효율성에 대한 과대평가치의 감소비율(F)
제약 조건이 없는 경우	20개씩 총5묶음	39개 (9+8+7+6+9)	95%	23%	78.80	10%	4.54%
	50개씩 총2묶음	30개 (14+16)	50%	10%	74.95	5%	1.90%
	100개 한 묶음	27개	35%	-	73.78	3%	-
제약 조건이 있는 경우	20개씩 총5묶음	34개 (6+8+7+6+7)	70%	29%	77.71	9%	5.50%
	50개씩 총2묶음	24개 (12+12)	20%	20%	73.58	3%	1.94%
	100개 한 묶음	20개	0%	-	72.19	1%	-

* 진실한 효율성 값의 평균 = 71.33 * 실험에 사용된 효율성이 1 인 단위의 수 = 20개

〈그림 2〉 제약 조건, 분석 표본 수와 DEA 효율성의 관계



〈표 6〉과 〈그림 2〉에서 알 수 있는 바와 같이 DEA 효율성은 분석 대상 표본 수가 증가하는 경우에는 그 값이 체계적으로 감소해서 진실한 효율성에 가까이 접근하게 되고 표본 수의 증가에 따른 개선의 정도는 감소하는 경향을 보임을 알 수 있다.⁴⁾

그리고 효율적이라고 과대 평가된 단위의 감소 비율을 살펴보면 〈그림 2〉에서와 같이 제약 조건 없는 경우와 제약 조건 있는 경우 모두 표본 수의 증가에 따라 과대평가 단위의 비율이 급격히 감소 하기는 하나 제약 조건 있는 경우가 더욱 빠르게

4) Zhang과 Bartels(1998)은 투입물이 bivariate lognormal 분포를 한다고 가정하고 가중치의 효과를 고려하지 않은 그들의 연구에서 DEA 효율성이 기술적 효율성의 평균, 분산, 투입물간의 상관관계와 관계없이 이와 같은 체계적인 관계가 성립함을 보이고 있다. 그들은 표본수에 따른 영향을 줄이기 위해서 "size-n" 방법을 사용하여 3개국 전력 사업자의 생산성을 분석하였다.

감소한다. 제약 조건 있는 경우는 20단위에서 50단위로 증가함에 따라 과대 평가된 단위의 29%가 감소하고 50단위에서 100단위로 증가함에 따라 과대평가된 단위의 20%가 추가적으로 감소하여 결국은 100단위의 경우에는 과대 평가된 효율적 단위는 존재하지 않게 된다.

한편, 진실한 효율성에 대한 과대평가치의 감소 비율의 경우에도 제약 조건 없는 경우와 제약 조건 있는 경우 모두 표본수의 증가에 따라 평균적 효율성 지표의 값이 급격히 하락하기는 하나 제약 조건 있는 경우가 더욱 빠르게 하락한다. 제약 조건 있는 경우는 20단위에서 50단위로 증가함에 따라 과대평가된 정도의 비율이 5.50%가 감소하고 50단위에서 100단위로 증가함에 따라 과대평가된 정도의 비율이 1.94%가 추가적으로 감소하여 결국은 100단위의 경우에는 평균적 효율성 지표의 값이 72.19로 단지 1% 정도만의 과대평가 비율이 존재하게 된다.

3.2 DEA 효율성과 참조 집단 변화

제약 조건의 도입과 분석 대상 표본 수가 DEA 효율성에 영향을 미친다는 점이 규명되었으나 과연 DEA 모형에서 특정의 의사결정단위를 평가하는 경우 사용되는 참조 집합이 DEA 효율성 지표의 영향을 제거한 후에도 제약 조건이나 분석 대상 표본 수에 의해서도 영향을 받을 것인가 하는 점을

규명할 필요가 있다. 따라서 로짓 모형을 이용하여 참조 집합 결정요인을 살펴본다. 로짓 모형을 사용한 이유는 이 모형이 회귀분석 시에 설명 변수들이 정규분포를 따라야 한다던가 설명 변수간의 공분산 구조가 동질적일 것을 요구하지 않는다는 점 그리고 단 두 개의 값만을 가지는 목적 변수와 설명 변수 사이의 인과관계를 규명하는데 도움이 되기 때문이다.

DEA 모형의 효율성 값에 따라 참조 집합을 구성하는 램다의 수에 변화가 초래하는 지를 검증하기 위해 변화를 살펴본 결과 램다는 1에서 4 사이의 분포를 하므로 램다의 수가 1,2인 경우와 3,4인 경우의 두 개의 집합으로 나누어 분석을 수행한다.⁵⁾ 램다의 수가 작다는 것은 DEA 효율성 분석 과정에서 사용된 참조 집합의 수가 작다는 것이고 램다의 수가 크다는 것은 참조 집합이 그 만큼 많다는 것이다. 그러므로 종속변수인 y 를 다음과 같이 정의한다.

$$y = \begin{cases} y=1, & \text{만약 } \lambda \text{의 갯수} = 1, 2 \\ y=2, & \text{만약 } \lambda \text{의 갯수} = 3, 4 \end{cases}$$

그리고 $\frac{P(y=1/x)}{P(y=2/x)} = \exp(\beta_0 + \sum \beta_i x_i)$ 의 함수를 구성한 후 양변에 자연로그를 취하여 램다의 수를 나타내는 변수 y_i 의 변화를 선택 변수인 공변량 x_i 로 설명하고자 한다. 그런데 이 함수를 그냥 사용하기에는 불편하므로 이 식을 변형하면

$$E(y) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2)} = p \text{가 된다.}$$

5) 종속변수(λ)가 1,2,3,4의 네 개의 값을 가지나 독립변수인 제약조건 유무와 분석표본 수의 크기는 모두 범주형 변수로 각각 (0,1), (0,1,2)의 값만을 가지므로 독립변수의 변화에 따른 종속변수의 값의 유의한 차이를 검증하기 어려울 뿐 만 아니라 효율적인 단위의 경우 종속변수의 개수가 1로 나오므로 일반적인 회귀식만으로 분석하기에는 어려움이 있다. 또한 본 연구에서 보이고자 하는 것은 참조집합의 변화를 회귀식으로 구하고자 하는 것이 아니라 참조집합의 변화의 방향이 독립변수에 의해 어떤 영향을 받는가 하는 점에 있으므로 로짓 모형이 바람직하다고 판단되었다. 종속변수 1,2,3,4를 (1,2)와 (3,4)로 나눈 것은 임의적인 구분이라기 보다는 참조 집합의 수가 작은 경우와 큰 경우로 구분한 것으로 정보량은 상실될 지 모르나 범주형 변수로 처리하기 쉽다.

〈표 7〉 범주형 변수의 재코드화

	값	분석표본수	파라메타 코드화	
			(1)	(2)
x_1	0	300	1	0
	1	300	0	1
x_2	0	200	1	0
	1	200	0	1
	2	200	0	0

이 식의 양변에 자연로그를 취하여 정리하면 분석에서 사용할 수 있는 다음과 같은 로짓 모형이 된다.

$$\langle \text{식 8} \rangle \ln\left(\frac{p}{1-p}\right) = \beta_0 + \sum \beta_i x_i$$

β_0 : 상수항

x_1 : 계약조건유무(더미변수: 0, 1)

x_2 : 분석표본수(더미변수: 0, 1, 2)

x_3 : DEA효율성값(정량변수)

독립변수 x_1 과 x_2 는 로짓 분석 과정에서 〈표 7〉과 같은 변수로 재분류된다.

로짓 모형은 최우추정법에 의하여 모형을 추정하므로 적합도를 측정해 주는 -2LL를 살펴보아야 한다. -2LL 값이 상수만이 포함된 경우의 값이 808.4이고, 상수항에 분석하고자 하는 변수를 추가로 포함하는 경우에는 657.1로 값이 감소하므로 모형에서 선택된 변수가 모형 추정의 불확실성을 제거

하는데 기여하고 있으므로 추가된 설명 변수가 의미를 가짐을 알 수 있다.

또한 모형의 전반적인 적합성을 의미하는 $\chi^2 = 32.86$ 으로 유의확률은 0.000으로 $\alpha = 0.05$ 보다 작으므로 귀무가설 $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ 은 기각되고 대립가설인 “적어도 하나의 $\beta_j \neq 0$ ”이 선택된다. -2LL과 χ^2 두 가지 통계량의 결과를 참조할 때 모형은 통계적 관점에서 적합성이 높다고 할 수 있다. 로짓 모형에 의한 회귀 계수의 설명력은 다음과 같다.

〈표 8〉을 보면 분석에 도입된 변수는 모두 유의수준 10%에서 0 아닌 람다의 수 예측에 도움이 된다는 사실을 알 수 있다. 특히 분석 표본의 수나 DEA 효율성 지표에 대해서는 매우 강한 관계를 지니고 있다. 또한 제약 식이 증가할수록 0 아닌 참조 집합의 개수도 증가한다는 사실도 알 수 있다.

우선 DEA 효율성이 커질수록 0 아닌 람다의 개수가 감소한다. 이는 DEA 효율성이 커지는 경우

〈표 8〉 로짓 모형에 의한 설명 변수의 검정

변수 종류	계수	표준편차	Wald통계량	유의확률
상수항	2.0418	0.3802	28.8372	0.0000
계약 조건유무(1)	0.3380	0.1913	3.1231	0.0772
분석표본수(1)	1.3573	0.2408	31.7582	0.0000
분석표본수(2)	0.7929	0.2398	10.9304	0.0009
DEA효율성	-.0458	0.0048	92.5261	0.0000

는 그 자신이 효율적 프론티어 상에 존재할 가능성이 커지게 되므로 자기 자신이 스스로의 참조 집합으로 선택되거나 그 단위의 투입 산출의 구조가 타 단위와 달라 특이하므로 다른 준거 기준이 없어서 자기 자신만으로 준거집단을 구성하는 경우가 있을 수 있다. 그리고 효율성이 낮은 단위의 경우에는 효율적 프론티어를 구성하는 2 이상의 단위에 의해서 평가될 것이고 이런 단위들과 비교했을 때 성과가 낮은 것으로 밝혀진다. 그런데 비효율성의 정도가 증가할수록 참조 집합이 증가하는 경향이 있다는 점은 그 동안 밝혀지지 않은 DEA 모형의 실증적 특성이라고 할 수 있다. 즉 효율성이 낮은 단위는 그와 유사한 투입 산출 구조를 가지고 있는 여러 단위들과 비교되므로 참조 집합의 수가 증가하고 참조 집합 내의 단위들과 비교했을 때 성과가 낮다는 점이다. 이는 DEA 모형의 평가 결과가 비효율적인 단위에 대해서는 매우 객관적인 결과라는 것을 의미한다. 적어도 비효율적으로 밝혀진 단위들은 가장 유사한 구조를 가진 단위들과 비교한 결과 투입이나 산출 쪽에서 잉여 분이나 부족 분 발생 결과 비효율적인 것으로 판정되었으므로 평가의 신뢰성이 객관성이 보장된다. 제약 조건의 유

무 또한 참조 집합 수에 영향을 미친다. 가중치에 제약 조건 부과되는 경우는 제약 조건을 만족하는 단위 중에서 투입-산출 구조가 유사한 단위가 선정되므로 참조 집합의 범위가 그 만큼 제한된다. 그리고 자기 단위만의 뛰어난 투입 산출 요소만으로 효율성을 평가하게 되면 그 요소에 대한 가중치만 증가하고 다른 요소에 대한 가중치는 감소하게 되므로 제약 조건을 만족시키지 못한다. 따라서 최초에 선택된 가중치의 조정이 있게 되고 이에 따라 참조 집합의 구성 단위도 변화하게 된다.

마지막으로 분석 대상이 되는 표본 수의 증가는 DEA 효율성에 매우 중요한 영향을 미치는 것을 알 수 있다. 표본 수가 증가할수록 비교 집단의 수가 증가하게 되고 자신만의 독특한 투입-산출 구조에 의해서만 평가되는 기회가 감소하므로 DEA 효율성이 감소하게 되고 동시에 0 아닌 람다의 수가 증가한다.

〈표 8〉에 나타난 것처럼 람다 수의 증가와 관련된 로짓 모형은 다음과 같이 나타낼 수 있다.

$$\ln\left(\frac{p}{1-p}\right) =$$

$$2.0418 + 0.338\text{제약조건유무}(1) + 1.3573\text{분석표본수}(1)$$

〈표 9〉 제약 조건과 분석 표본 수에 따른 참조 집합의 변화

항목	분석 대상 표본크기	참조 집합(0 아닌 λ_j)의 개수				평가에 사용된 참조 집합 수의 평균	참조 집합 갯수의 증가 비율
		1 개	2 개	3 개	4 개		
제약유무	20 개	42	33	22	3	1.89 개	18.5 %
	50 개	35	20	34	11	2.24 개	10.2 %
	100 개	31	14	35	20	2.47 개	-
제약 조건 없는 경우	20 개	38	41	21	0	1.85 개	12.4 %
	50 개	31	32	37	0	2.08 개	11.5 %
	100 개	22	26	37	0	2.32 개	-

$+0.7929$ 분석표본수(2)- 0.0458 DEA효율성 값

여기서 p 는 람다의 수가 상대적으로 적을 (1개 또는 2개) 확률을 나타내고 $(1-p)$ 는 람다의 수가 상대적으로 클(3개 또는 4개) 확률을 나타낸다. 로짓 분석 결과와 관련하여 참조 집합 수의 변화를 살펴보면 다음 <표 9>와 같다. 제약 조건 없는 경우에는 참조 집합이 4개인 경우가 표본 수가 20개, 50개, 100개인 경우에 각각 3개, 11개, 20개에서 가중치에 대한 제약이 추가되는 경우에는 표본 수에 관계없이 모두 0개가 된다. 이렇게 줄어든 0 아닌 λ_j 의 수는 참조 집합의 개수가 2개나 3개인 경우로 주로 변경되었다. 그 결과 제약 조건에서의 참조 집합의 개수의 평균이 제약 조건 없는 경우에 비하여 체계적으로 감소하게 된다.

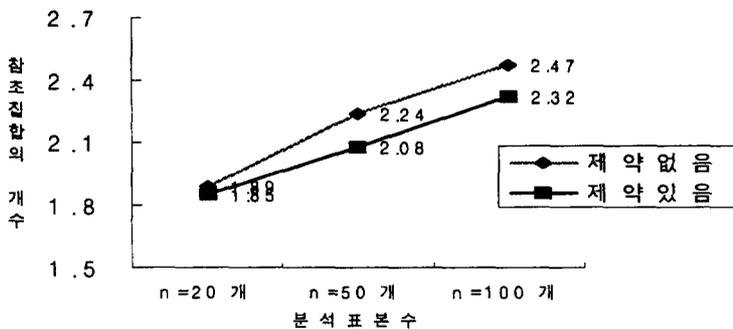
로짓 분석에서 확인한 바와 같이 분석 대상 표본 수가 증가할수록 참조 집합의 개수가 증가하는데 분석 대상 표본이 20개에서 50개로 증가하는 경우에는 참조 집합 개수가 증가하는 비율이 각각 18.5%와 12.4 %로 나타나지만 50개에서 100개로 증가하는 경우에는 그 비율이 각각 10.2 %와 11.5%로 나타나 체계적인 감소를 나타내지는 않고 있다. 그리고 표본 수의 증가에 따른 참조 집합

개수의 감소 폭은 무제약 하의 경우가 18.5%에서 10.2 % 로, 제약 조건 있는 경우에는 12.4%에서 11.5%로 변해 무제약 하의 경우가 더욱 크다. 한편 제약 조건 없는 경우에는 참조 집합이 1개인 경우가 표본 수에 따라 42개, 35개, 31개에서 제약 조건 있는 경우에는 각각 38개, 31개 22 개로 변한다.

그리고 다음 <그림 3>에서 알 수 있는 바와 같이 제약 조건 있는 경우와 없는 경우의 참조 집합의 개수의 차이는 표본 수가 작을 때는 거의 없으나 표본 수가 커지면 참조 집합의 개수의 차이가 나타나게 되고 그 차이는 표본 수가 50개인 경우나 100개인 경우나 거의 균등하게 유지되고 있음을 알 수 있다.

마지막으로 제약 조건 없는 경우에는 참조 집합의 개수가 1개부터 4개까지 폭넓게 분포하나 제약 조건 부과된 경우에는 참조 집합의 개수가 2개와 3개에 보다 많이 집중되어 있고 참조 집합의 개수가 1개인 경우는 상대적으로 줄어들고 참조 집합의 개수가 4개인 경우는 전혀 나타나지 않는 특징을 보인다. 그러므로 제약 조건 없는 경우에는 참조 집합의 수가 넓게 분포하는 반면 제약 조건 있는 경우에는 참조 집합의 수가 조밀한 분포를 한다

<그림 3> 제약 조건, 분석 표본 수에 따른 참조 집합의 개수



고 볼 수 있다.

IV. 결 론

DEA모형은 평가 대상이 되는 의사결정단위가 가장 효율적으로 평가받을 수 있도록 상대적 가중치를 설정하므로 주관적인 가중치가 배제되고 평가의 객관성이 유지된다. 그러나 이차적으로 중요한 요소가 효율성을 결정할 수 있고, 중요한 요소가 무시될 수 있고 사전 정보가 누락될 수 있는 단점도 있다. 그리고 무엇보다도 의미 있는 제약 요소가 무시됨으로써 효율성 값이 과대 평가되거나 왜곡될 수 있다. 본 연구에서는 DEA모형의 특성에 따라 가중치에 제약 조건이 부과됨으로써 효율성이 영향을 받을 수 있다는 점과 참조 집합이 변화되며 그에 따라 상대적인 효율성 값도 심각하게 영향을 받는다는 점을 콤팩트클래스 생산함수라는 이상적인 함수를 도입하여 실증 분석하였다. 그리고 상대적 효율성이 참조 집합의 변화와 매우 유의한 관계를 지닌다는 점도 규명하였다. DEA 효율성 값은 콤팩트클래스 함수 상의 진실한 효율성 값의 효과를 제거한 후에도 제약식의 유무와 분석 표본 수에 의해서도 설명이 가능하다. 이는 제약식이 없고 분석 대상 표본 수가 작은 경우에는 DEA 효율성이 급격히 증가하는 현상을 보이고 있음을 의미한다.

그리고 효율적이라고 과대 평가된 의사결정단위의 감소 비율은 제약 조건이 있는 경우와 없는 경우 모두 표본 수의 증가에 따라 과대평가 단위의 비율이 급격히 감소하나 제약 조건 부과된 경우가 더욱 감소폭이 크다는 것도 확인하였다. 본 실험에서는 표본 수가 50개이상인 경우에는 진실한 효율

성에 대한 과대평가 비율이 5% 이내로 감소하는 것도 확인할 수 있었다.

참조 집합의 변화를 로짓 함수를 이용해서 분석한 결과 분석 표본 수의 증가에 따라 참조 집합의 수가 증가하며, 제약 조건이 있는 경우 참조 집합의 수가 감소하며 DEA효율성이 증가할수록 참조 집합의 수가 감소한다는 점을 알 수 있었다. 이는 대체적으로 그런 경향이 볼 수 있다는 것이 아니라 통계적으로 검정 가능하다는 점에서 더욱 의미가 있다.

제약 조건이 있는 경우에는 참조 집합의 수가 2개, 3개 등으로 평균치에 가까이 조밀하게 분포하고 있는 반면 제약 조건이 없는 경우에는 참조 집합의 수가 1개, 4개 등으로 폭넓게 분포하는 특징도 알 수 있었다.

비록 가중치에 무한한 탄력성을 부여하거나 표본 수가 작은 경우라도 DEA모형은 실무적인 차원에서 매우 유용한 평가 기법임에는 틀림없다. 이에 추가하여 가중치에 대한 제약과 분석 대상 표본 수에 따른 DEA효율성과 참조집합의 변화를 고려해서 분석한다면 성과 평가 수단으로서 DEA 모형을 적용하고 해석하는데 보다 의미 있는 결과를 얻을 수 있을 것이다.

참 고 문 헌

- 안태식(1990), "효율성 평가방법으로서의 DEA: 가설검정시의 민감도 분석", **한국회계학회 춘계발표회**.
- 오동일(2000), "투입 요소에 범주형 변수가 포함된 DEA 모형의 설계 및 사례 분석", **생산성은집**, 생산성학회, 2000년 2월호.
- 오동일(2000), "효용함수를 반영한 가치효율성분석 : 시스

- 템 개발 프로젝트 성과평가”, *고객만족경영연구*, 고객만족경영학회, 제 2권 1호.
- 오동일(2000), “참조점 방법을 이용한 DEA 모형의 프론티어 탐구”, *산학기술성공학회지*, 산학기술성공학회, 제 1권 1호.
- Ahn, T., A. Charnes, and W. W. Cooper(1988). “Efficiency Characteristics in Different DEA Models.” *Socio-Economic Planning Science* 22, 253-257.
- Banker, R. D., A. Charnes, W. W. Cooper, J. Swarts and D. Thomas(1988), “An Introduction to Data Envelopment Analysis with Some of Its Model and Their Uses, *CCR Report* 619, The University of Texas at Austin.
- Banker, R. and R. C. Morey(1986), “The Use of Categorical Variables in Data Envelopment Analysis.” *Management Science* 32(12) 1613-1627.
- Banker, R. D., and R. M. Thrall(1992), “Estimation of Returns to Scale Using Data Envelopment Analysis.” *European Journal of Operational Research* 66, 74-84.
- Banker, R. and R. C. Morey(1989), “Incorporating Value Judgements in Efficiency Analysis”, *Research in Governmental and Nonprofit Accounting* 5, 245-267.
- Chaparro, F., P. Jimenez, J. S., and P. Smith(1997), “On The Role of Weight Restrictions in Data Envelopment Analysis”, *Journal of Productivity Analysis* 8, 215-230.
- Charnes, A., and W. Cooper., and E. Rhodes(1978), “Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research* 2(6), 429-444.
- Charnes, A., W. W. Cooper, A. Y. Lewin and L. Seiford(1994), *Data Envelopment Analysis, theory, methodology and Applications*, Kluwer Academic Publishers.
- Charnes, A., W. W. Cooper, A. Y. Lewin and L. Seiford(1994), *Data Envelopment Analysis, theory, methodology and Applications*, Kluwer Academic Publishers.
- Charnes, A., W.W. Cooper, and L. Seiford, and J. Stutz(1983), “Invariant Multiplicative Efficiency and Piecewise Cobb-Douglas Envelopments.” *Operations Research Letters* 2, 101.
- Charnes, A., W.W. Cooper, and R. L. Thrall(1986), “Classifying and Characterizing Efficiencies and Inefficiencies in Data Envelopment analysis.” *Operations Research Letters* 5, 105-110.
- Clark, J. A(1988), “Economies of Scale and Scope at Depository Financial Institutions: A Review of the Literature.” *Economic Review*, Federal Reserve of Kansas City, Kansas City, Sep./Oct., 16-33.
- Dyson, R. G., and E. Thanassoulis(1988), “Reducing Weight Flexibility in Data Envelopment Analysis.” *Journal of the Operational Research Society* 39(6), 563-576.
- Farrell, M. J(1957), “The Measurement of Productive Efficiency.” *Journal of the Royal Statistic Society Series A* 120, 251-281.
- Joe Zhu(1996), “Robustness of efficient DMUs in data envelopment analysis”, *European Journal of Operational Research* 90(3), May, 451-460.
- Roll, Y., W. D., Cook, and B. Golany (1991), “Controlling Factor Weights in Data Envelopment Analysis.” *IIE Transactions* 23(1), 2-9.
- Seiford, L. M(1995). “A DEA Bibliography (1978-1994).” In A. Charnes, W. W. Cooper, A. Lewin, ed. L. M. Seiford, (eds.), *Data Envelopment analysis: Theory, Methodology and Applications*, Kluwer Publishers.
- Seiford, L. M., and J. Zhu(1998), “Stability Regions for maintaining efficiency in data envelopment analysis”, *European Journal of Operational*

- Research* 108(1), July, 127-139.
- Seiford, L. M., Thrall(1990), "Recent Developments in DEA: the Mathematical Approach to Frontier Analysis." *Journal of Econometrics* 46, 7-38.
- Sherman, H. D(1988), "Data Envelopment Analysis as A New Managerial Audit Methodology, Test and Evaluation", *Auditing* 4(1), 35-53.
- Thompson, R. G., P. S. Dharmapala, and R. M. Thrall(1993), "Importance for DEA of Zeros in Data, Multipliers and Solutions." *Journal of Productivity Analysis* 4(4), 379-390.
- Thompson, R. G., P. S. Dharmapala, and R. M. Thrall(1995), "sensitivity analysis of efficiency measures with applications to KANSAS farming and ILLINOIS coal mining", *Data Envelopment analysis: Theory, Methodology and Applications*, Kluwer Publishers.
- Thompson, R. G., L. N. Langemeier, C. Lee, E. Lee, and R. M. Thrall(1990), "The Role of Multiplier Bounds in Efficiency Analysis with Application to Kansas Farming." *Journal of Economics* 46, 93-108.
- Yun Zhang and R. Bartels(1998), "The Effects of Sample Size on the Mean Efficiency in DEA with an Application to Electricity Distribution in Australia, Sweden and New Zealand", *Journal of Productivity Analysis* 9(3), 187-204.
- Valter, B.(1999), "A note on robustness of the efficient DMUs in data envelopment analysis", *European Journal of Operational Research* 112(1), Jan, 240-244.

An experimental study on the behaviour of DEA efficiency scores and reference sets with the restrictions in virtual weights and sample size variations.

Dong-II O*

Abstract

DEA is value free and does not require specification or knowledge of prior weights or prices for inputs or outputs. If there is some informations on value judgements on virtual weights, we can improve the quality of DEA results.

This study investigates the relationships between DEA efficiency scores, weight restrictions and data sizes by simulation experiments. Also this paper interests in the behaviour of reference sets and which factors influence the change of reference sets.

Some results of this study are as follows.

Overestimation in efficiency score arises when using unbounded DEA, but this phenomenon can be reduced to some extent with additional restrictions in virtual weights and increased sample sizes.

After eliminating the effects of the assumed real efficiency scores, efficiency scores systematically changed by restriction in weights and sample sizes.

The number of reference sets is reduced as sample sizes decreased, but the degree of reduction rate is smaller in sample with restrictions.

Key words: DEA, performance evaluation, weight restrictions, reference sets

* Associate Professor, Dept. of Finance and Insurance, SangMyung University.
330-180, AhnSeo-Dong San 98-20, Chung Nam, SangMyung University, Korea.