

# 공익법인에 대한 제3자 등급평가는 기부금 수입을 변화시키는가?\*

육근호(단독저자)  
부산외국어대학교 경영학부  
(yook@bufs.ac.kr)

본 연구에서는 NPO(공익법인) 감시기관이 부여한 제3자 등급(별점 평가)을 조사하여, 그러한 등급의 증가 또는 감소가 NPO의 기부금 수입에 독립적인 영향을 미치는지를 회귀불연속설계(Regression Discontinuity Design: RDD) 접근법을 사용하여 분석하였다.

분석결과를 요약하면 다음과 같다. 첫째 NPO에 대한 제3자 등급(평점)이 기부금 의사결정에 유용한 정보가 된다는 가설에서는, 모든 등급의 증가가 기부금에 우호적인 영향을 미친 것이 아니라, 일부의 범위(별 등급 2개에서 3개로 증가)에서만 지지되었다. 둘째 규모에 따른 등급과 기부금 크기의 관계는, 대규모 NPO의 경우에 별 등급 2개에서 3개로 등급의 상승이 임계점(90점)에서 유의한 기부금 수입의 증가(jump)가 나타난 것을 제외하고는 소규모NPO의 경우에는 등급의 증가가 모두 임계점에서 기부금의 증가 즉 회귀불연속이 발생하지 않았다. 셋째 개인기부자와 기업기부자의 등급에 대한 반응이 상이할 것이라는 가설에서는, 개인기부가 임계점 90에서 등급의 처리효과가 기업 및 단체기부에 비해 두 배 이상 강하게 나타났다. 넷째 NPO에 대한 제3자 등급 정보와 웹 사이트 공시 정보가 언론보도에 비해 상대적으로 기부금 크기와 강한 관련성을 갖는다는 것이 확인되었다. 또한 추가분석에서는 매년 연속적으로 별 3개 등급을 받은 NPO가 그렇지 않은 NPO에 비해 기부금 수입에 매우 강한 양의 영향을 미친다는 것을 보여주었다. 본 연구의 시사점은 제3자 등급평가의 향상이 NPO의 기부금 수입에 양의 영향을 미치는 것은 분명하지만, 무엇보다 실제 평점의 절단점에서 변화하게 되는 별 등급 정보가 기부 의사결정에 핵심적 역할을 담당한다는 것이다. 향후 연구에서는 최근 신용평가 제도가 등급에서 점수제로 전환되는 것처럼 별 등급을 보완할 수 있는 다양한 평가방안이 제시되어야 할 것이다.

주제어: 기부금, 공익법인(NPO), 회귀불연속, 별 등급, 평점

## 1. 서론

국세통계연보에 의하면 2018년에 우리나라 기부금 규모는 12조 9천억원에 이르며 GDP 대비 0.8% 정도를 차지하고 있다. 이중 NPO(의무공시 공익법인)은 2018년 국세청 공시 데이터 기준으로 6조5천억을 기부 받았으며 총기부액은 증가하는 추세에 있다. 이와 같이 기부자가 기부를 늘리면 자선단체 감시기관(Watchdog) 및 이해관계자에 의한 감시·감독의 필요성이 확대되고 NPO로 하여금 수탁 책

임을 이행하도록 압력을 증대시키게 된다(Primoff, 2012). 미국의 경우 Sarbanes-Oxley 규정 준수 문제로 인해 NPO 법인의 기부금에 대해서도 실제로 이전보다 더 많은 재정적 감독이 요구되고 있다(Hatfield, 2018).

그런데 NPO의 활동은 기업과 시장 간의 관계와 같이 그 활동의 성과를 측정하고 성과의 피드백을 얻을 수 있는 메커니즘이 존재하지 않아 NPO 활동의 가치 판단은 개별 법인의 판단에 맡기게 된다. 이것은 어떤 면에서 NPO 활동이 독선적으로 이어질 위험성을 내포하고 있다고 할 수 있다. 따라서 NPO

최초투고일: 2020. 1. 11

수정일: (1차: 2020. 2. 6)

게재확정일: 2020. 2. 7

\* 본 연구는 2020년 부산외국어대학교 교내지원연구비에 의해 수행되었으며 본 연구를 위해 자료를 제공해 준 (재)한국가이드스타에 감사드린다.

가 설명책임을 충실히 이행하고 있는지에 대해서는 개인이나 기업 기부자가 다양한 매체를 통해 획득한 정보에 의해 확인하고 싶어 한다.<sup>1)</sup> 그러나 일반적으로 기부자는 자신이 지원하고자 하는 자선 단체가 합법적인지 여부를 판단하고 단체를 선택할 수 있는 전문 지식이 부족하다. 반면에 NPO는 자신들의 부정적인 정보는 가능한 억제하고 긍정적 평판 및 조직에 대한 호의적 정보를 홍보할 수 있는 자원을 보유하고 있다. 이것은 기부자와 NPO 간의 정보 비대칭으로 이어져 대리인이 비효율적으로 조직의 자원을 배치하거나 자신의 목표를 추구하기 위해 자원을 의도적으로 배치하는 문제가 발생한다(Sloan, 2009).

이런 현상에 대응하여 NPO의 수준을 판별할 수 없는 기부자는 소위 '레몬 문제'(Akerlof, 1970)를 해결하는 수단으로 자선단체의 감시기관이 제공하는 제3자 평가인 등급 정보를 사용하게 된다.<sup>2)</sup> 즉 기부자는 감시기관이 제공하는 신뢰할 수 있는 정보(주로 등급 또는 평점)를 기반으로 NPO를 평가하고 선택하여 기부 여부를 결정할 수 있으며, 이 때 등급 부여는 기부 결정과 관련하여 NPO의 재무 정보 등을 해석해 줌으로써 기부 시장에서의 실패를 완화시키는 기능을 담당하게 된다.

감시기관의 평가 결과인 등급 및 순위는 교육분야, 의료분야, 비영리단체 등 다양한 분야에서 나타나는 글로벌 설명책임 확장의 한 가지 현상이며 지속적으로 확산되는 추세에 있다(Espeland and Sauder, 2007). 특히 국내외를 막론하고 많은 추문이 이어

지는 비영리 부문에서 이런 경향이 심화되고 있다(Gibelman and Gelman, 2001).

그런데 영리단체와는 달리 NPO가 제공하는 서비스 및 상품에 대한 제3자 평가의 영향은 상대적으로 덜 알려져 있으며 특히 자선 기부에 대한 제3자 평가의 효과를 조사한 연구는 한정되어 있다(Yörük, 2016). 또한 기부자가 기부금을 결정하는 동안 등급을 심각하게 인식하는지 여부는 선행연구에서 명확하지 않으며, 등급과 기부 사이에 연관성을 분석한 선행연구의 결과 또한 혼재되어 나타났다.

이에 본 연구에서는 NPO와 기부자들 간의 관계를 대리인 이론으로 파악하여 대리인 이론의 렌즈를 통해 정보-기반 규제 논리의 즉 자선단체 감시기관이 제공한 정보를 통한 NPO 규제의 효용성과 타당성을 검토한다(Prakash and Gugerty, 2010). 그것은 감시단체의 등급 부여가 낮은 거래 비용으로 관련 정보에 접근할 수 있는 능력을 제공하고 사용자가 이 새로운 정보를 바탕으로 기부에 관한 행동을 변경하려는 의지에 공헌하기 때문이다(Szper and Prakash, 2011).

구체적으로 본 연구에서는 첫째 (재)한국가이드스타가 부여한 등급(별점 평가)을 조사하여 그러한 등급의 증가 또는 감소가 NPO 기부금 수입에 대해 독립적인 영향을 미치는지를 분석한다. 둘째 상대적으로 규모가 작고 신생 NPO에 기부할 때 기부자들이 독립적인 제3자 등급에 더 의존할 가능성이 높다는 가설을 검증한다. 셋째 개인 만족감이나 기부 행동

1) 예를 들면 기부자가 직접 관련 사이트를 방문하거나 가족과 친구의 추천을 받거나 자선단체 감시기관의 등급(ratings)이나 인증(Charity Seal) 정보를 활용하여 기부 대상기관을 선정한다(Hatfield, 2018).

2) Charity Watchdog(자선단체 감시기관)은 NPO에 대한 정보, 검토 및 평점을 제공하는 비영리 감시단체에 적용되는 일반적인 용어이다. 그들은 일반적으로 자선 감시단이라는 우산 아래에서 함께 모이기는 하지만 서로 다른 임무를 수행하고 기능도 상이하며 NPO가 필요한 자금을 모금하는 능력에 상이하게 영향을 미친다. 대표적인 감시기관으로 American Institute of Philanthropy, Charity Navigator, the Better Business Bureau's Wise Giving Alliance 등이 있으며 우리나라에서는 (재)한국가이드스타가 있다.

자체에 기쁨을 느끼거나 개인 기부자에 비해 전략적 마케팅 기회를 중시하는 기업 기부자는 완전하고 면밀한 회계 검토를 요구하므로 평가기관의 정보를 더 많이 사용하는지를 검토한다(Cnaan, Jones, Dickin, and Salomon, 2011). 넷째 기부자의 기부 의사결정에 영향을 미치는 중요 요인으로서 제3자 등급이나 순위 이외에 NPO 웹 사이트 검색과 미디어 보도 정보가 있다. 본고에서는 이들 세 가지 요인이 기부자의 의사결정에 미치는 정보의 효과 즉 정보 내용의 상대적 유용성을 분석하여 기부자들이 선호하는 정보를 확인하고자 한다. 이와 같이 NPO 규모와 기증자의 분류(개인 기부, 기업 및 단체 기부)에 따른 기부금에의 영향, 제3자 등급 정보와 웹 사이트 공식 정보 그리고 검색에 의한 각종 미디어 보도와 같은 비간섭 척도(Unobtrusive measure)를 활용하여 기부금 수입에 대한 영향력을 실증적으로 비교하는 것은 국내외를 막론하고 현재까지 처음 시도되는 연구로서 상당한 의의가 있을 것이다.

특히 본 연구는 NPO가 수취한 기부금에 대한 등급 부여 효과의 측정을 위해 회귀불연속설계(Regression Discontinuity Design: RDD) 방법을 사용하였다. 이것은 기부에 대한 제3자 평가의 영향을 확인한 선행연구가 대부분 OLS 추정 또는 패널 데이터 모형에 의존했으며(Gordon, Knock, and Neely, 2009; Szper and Prakash 2011), 이로 인해 외생변수나 시간에 따라 변하는 기부특성 때문에 기부 활동에 대한 등급 효과의 편향된 추정치를 제공할 가능성이 있기 때문이다.<sup>3)</sup> 이런 문제를 해결하기 위해 본고에서는 NPO에 대한 특정 등급보다 약간 높거나

나 낮은 평점이 기부금 성과에 미치는 영향을 정교하게 비교하여 경계선 주변에서의 등급 효과를 식별 및 추정할 수 있는 RDD를 사용한다(Grant, 2010; Yörük, 2016).

## II. 선행연구 검토와 가설설정

### 2.1 NPO 감시기관(Watchdog)의 등급 부여

기부자는 자선 등의 비영리활동을 지원하기를 원하지만 그렇게 할 수 있는 전문 지식이 없다. 반면에 NPO는 자선 관련 프로젝트를 수행할 수 있는 전문 지식을 보유하고 있지만 그렇게 할 수 있는 자원이 부족하다. 이것은 기부자와 NPO 간의 정보 비대칭으로 이어져 대리인이 비효율적으로 조직의 자원을 배치하거나 자신의 목표를 추구하기 위해 자원을 의도적으로 배치할 수 있게 한다. 실제로 기부 스캔들로 상당한 언론의 관심을 받은 일부 NPO는 행정 보상 및 비대한 관료제를 위한 간접비에 과도한 금액을 지출한다는 것이 밝혀졌다(Gibelman and Gelman, 2004). 이런 현상에 대응하여 NPO의 수준을 판별할 수 없는 일부 기부자는 이것을 '레몬 문제'라고 생각하고(Akerlof, 1970) NPO에 대한 기부금을 축소하거나(O'Neill, 2009), NPO에 대한 자신들의 강한 영향력을 행사하고자 하였다. 그런데 기부자의 목소리를 강조하려면 NPO의 내부 운영에 관한 정보가 필요하지만 그러한 정보에 대한 접근은 쉽지 않

3) 예를 들어 실제로 개인 신용평가점수가 1점 차이로 등급이 떨어져 대출에 불이익을 받는 문제를 개선하기 위해 최근 정부는 신용평가 등급제(1등급-10등급)에서 1000점 만점의 점수제로 변경하였다. 기부금의 경우에는 특정 NPO에 대한 전체 평점이 60에서 61로 증가하게 된다면 기부금에 미치는 영향은 매우 미미할 것으로 예상되지만 등급은 3등급에서 4등급으로 변경되어 버리기 때문에 등급을 기준으로 기부를 결정하는 일반 기부자들에게는 상당한 영향을 미치게 된다(Yörük, 2016).

다. 이에 자선단체의 감시기관(watchdog)이 기부 결정과 관련하여 NPO의 재무 정보를 해석해 줌으로써 기부 시장에서의 실패를 완화시키는 기능을 담당하게 되었다. 그리고 기부자는 감시기관이 제공하는 신뢰할 수 있는 정보(특히 등급 또는 순위)를 기반으로 NPO를 평가하고 선택하여 기부금을 결정할 수 있다.

물론 NPO 감시기관이나 평가기관에서 부여하는 등급 또는 평점 시스템에도 한계가 있다. 무엇보다 등급은 주로 프로그램의 내용 또는 품질보다는 재무 정보를 기반으로 하기 때문에 결과적으로 등급은 복잡한 조직기능에 비해 다소 단순한 정보를 제공하는 경향이 있다(Espeland and Sauder, 2007). 또한 NPO는 내부 관행을 변경하지 않고 합법성을 유지하기 위해 등급결정에 중요한 요소에 맞추어 재무 수치 등을 보고하는 방법을 바꿈으로써 등급을 향상시키는데 집중할 가능성이 있다.

그러나 이런 한계에도 불구하고 감시기관에 의한 등급시스템은 NPO와 기부자들에게 지속적으로 확대 적용되고 있다(Szper, 2013). 그것은 첫째 등급 시스템은 불완전한 정보 및 정보 비대칭 문제를 최소화한 기부자 관점에서 완화하기 위한 수단이 되기 때문이다(Sloan, 2009). 둘째 자선단체 평가기관이 높은 등급을 준 자선단체에 기부하는 것은 비용-효과적으로 쓰일 것을 보장받는 일이므로 기부에 대해 회의적인 사람들의 태도를 변화시킬 대안이 된다(Singer, 2015). 셋째, 이러한 등급 평가 시스템은 감시기관 스스로 정량적 요소 이외에 정성적 요인을 가미하여 평가시스템을 개선함으로써 지속가능한 실

무를 발전시킨다. 예를 들어, BBB(Better Business Bureau's Wise Giving Alliance) 표준은 재무실적뿐만 아니라 지배구조, 감독 및 측정 효과를 포함하며 Charity Navigator는 기존의 재정 면에서의 건전성에 더하여 2011년부터 설명책임과 투명성 지표를 추가하여 자선 활동 실적을 분석한다.<sup>4)</sup> 마지막으로 등급은 평가기관이 자선단체에 관한 지식을 통합할 수 있게 한다. Charity Navigator는 섹터별(교육기관, 종교단체, 공공의료, 문화재단 등)로 평점 기준을 달리하는데 예를 들면 자원봉사자가 많은 자선단체는 관리비가 다른 단체에 비해 적게 소요되므로 이런 경우에는 해당 단체가 높은 등급을 받는데 필요한 관리비의 절단점(cutoff point)을 다른 섹터에 비해 상대적으로 낮게 책정한다는 것이다(Ling and Neely 2013).

이와 같이 각 자선단체에 대한 평가기관의 등급 및 순위부여는 현재와 미래의 기부자에게 자선단체의 재무 운영, 안정성, 소비 행태에 대한 정보를 제공하여 기부자의 합리적 의사결정에 공헌하므로 등급에 대한 수요는 증가하는 추세에 있다.

## 2.2 등급과 기부금의 관계에 대한 선행연구

일반적으로 공신력 있는 기관을 통해 평가·공시되는 등급 정보는 해당 조직에 대한 객관적인 정보를 제공함으로써 조직과 정보 이용자들 간의 정보비대칭을 완화시키고 자본시장 내 자원이 효율적으로 배분되도록 유도하는 역할을 담당한다(Tang, 2009). 예를 들어 기업의 신용평가정보는 채권시장에서의

4) 기존의 재정 면에서의 건전성에 더하여 2011년부터 설명책임과 투명성 지표를 추가하였다. 최근에는 Charity Navigator가 GuideStar, Classy, GlobalGiving, ImpactMatters와 협력하여 2,200개의 등급을 부여 받은 자선단체와 등급을 받지 못한 12,300개 자선단체에 대한 임팩트-관련 평가 정보(보고서)를 발행하고 있는데(2018년 기준) 이것은 빈곤, 소득불평등, 교육, 의료, 환경, 문화예술 등 사회문제를 해결하는 조직에서 이뤄낸 사회적 성과 혹은 영향력을 지표화해 기부자에게 제공하는 것이다.

정보비대칭을 줄이고 투자자와 채권자의 의사결정에 직접적인 영향을 미치게 된다. 에너지 효율의 증가를 유발할 수 있는 에너지효율등급 인증을 받은 건축물의 에너지효율등급이 높을수록 아파트 가격에 긍정적인 영향을 주는 것으로 나타났다(이동성·문태훈 2016). 또한 온라인에서 수많은 정보원로부터 접하게 되는 정보의 방향성을 나타내는 온라인 리뷰나 평점은 도서매출이나 영화산업의 수익에도 상당한 영향을 미친 것으로 파악되었다.

그러나 영리단체와는 달리 NPO가 제공하는 서비스 및 상품에 대한 제3자 평가의 영향은 상대적으로 덜 알려져 있으며 특히 자선 기부에 대한 제3자 평가의 효과를 조사한 연구는 한정되어 있다(Yörük, 2016). 그리고 등급과 기부 사이에 연관성을 분석한 선행연구를 살펴보면 결과가 혼재되어 나타난다. 먼저 등급의 긍정적 효과를 조사한 연구를 보면 Tinkelman(1999)은 개인을 제외한 기업 등의 조직 기부자는 부적절한 등급에 부정적으로 반응하는 것을 발견했으며 Chhaochhari and Ghosh(2008)는 가장 높은 등급의 자선 단체가 가장 낮은 등급의 자선 단체보다 16% 많은 자선 기부를 받았다는 사실을 발견했다. Sloan(2009)은 기부자가 긍정적인 평가에만 반응한다고 보고했으며 이와 유사하게 Charity Navigator가 평가한 405개의 자선단체의 데이터를 사용하여 Gordon et al. (2009)은 Charity Navigator가 수여하는 별 수가 증가하면 자선 단체에 대한 기부금이 증가한다는 사실을 알게 되었다. Charity Navigator가 평가한 자선단체의 더 큰 표본을 사용하여 Grant(2010)는 기부자가 등급이 저조한 자선단체에 대해 기부를 줄인 반면에, 별 하나 등급의 증가가 자선단체가 받은 기부금의 3~8% 증가와 관련이 있음을 실증했다. Finch(2014)의 설문 조사에 따르면 응답자의 55%가 기부 여부를 결정

할 때 자신이 관심을 갖는 자선단체의 등급을 살펴 보며 이 중에서 93%는 웹 사이트에 등급 표시가 없었을 때 자선 기부에 부정적인 영향을 미치며 72%는 등급이 기부금 크기에 유의적 영향을 미친다고 보고 하였다. 회귀불연속설계(RDD)를 사용하여 Yörük(2016)은 Charity Navigator 등급 시스템에서 추가 별점의 기부자 기여에 대한 영향이 자선 단체의 크기와 현재 등급의 함수임을 설명하였다. 소규모 자선 단체의 경우, 2점에서 3점, 3점에서 4점으로 별점이 증가하면 기부한 금액이 약 28% 증가하였다. Brown, Meer and Williams(2017)에서는 자선 단체에 관한 정보를 다양하게 하고 피실험자들에 기부할 자선단체를 선택하도록 하는 실험을 실시하였다. 결과는 이러한 등급이 특정 자선단체를 선정할 때 조건부로 기부 행위를 다소 늘리는 것으로 요인으로 나타났다.

한편 이와는 반대로 등급의 부정적 효과를 조사한 연구를 보면 Frumkin and Kim(2001)은 총비용에서 관리비가 차지하는 비율이 낮게 나타나 비용 효율적이라고 평가되는 NPO는 비효율적으로 보이는 조직보다 기부금 수입이 나아지지 않았다는 것을 발견했다. Szper and Prakash(2011)는 워싱턴 주의 자선단체를 대상으로 하여 자선단체에 대한 등급과 기부자의 기부금 간에는 아무런 관련이 없음을 확인하였다.

이상의 연구를 종합하면 대체로 NPO에 대한 평가 평점이 높을수록 우호적인 기부자 반응을 유도하여 NPO의 기부금 수입에 긍정적 영향을 미쳤다는 것을 알 수 있다. 따라서 감시기관이 부여한 등급이 하락한 NPO는 기부금이 감소하는 반면 높은 등급을 부여받은 NPO는 기부금이 증가하게 될 것이라고 예상하여 다음 가설 1을 설정하였다.

가설 1: NPO에 관한 등급(평점)이 높을수록 기부금 수입이 증가할 것이다.

### 2.3 자선단체의 규모, 등급, 기부금의 관계

기부자가 기부할 NPO를 선택할 때 가장 우선적으로 고려하는 두 가지 요소는 NPO의 크기와 등급인데 자선단체의 크기에 따라 등급이 기부금에 미치는 영향은 잘 알려져 있지 않다. 일반적으로 기부자들은 NPO의 명성을 그 단체의 품질에 대한 지침으로 사용하며 조직 명성의 대리변수로 규모를 사용하는데(Tinkelman, 1999), 이 방법에 의하면 잘 알려지지 않은 소규모 또는 지역 자선단체에 대해서는 기부자들이 선호하지 않는 결과를 가져오게 된다(Meer, 2014).

Gordon(2009)은 조직의 규모가 등급 변화와 공적인 직접 기부금 사이의 관계에 영향을 미칠 수 있는지를 검증하였으나 소규모, 중규모 또는 대규모 세 가지 집단에서 단체의 크기가 기부금 변화에 미친 직접적 영향을 발견하지 못했다. 이는 평가기관인 Charity Navigator로부터 등급을 받기 위해 선택된 자선단체들이 등급을 받지 못한 단체보다 상대적으로 더 크기 때문에 Charity Navigator의 등급을 활용하는 기부자들에게는 자선단체의 규모는 중요한 요인으로 고려되지 않기 때문이라고 설명했다. 한편 Luca(2011)는 제품이 비교적 알려지지 않고 고객들이 품질을 쉽게 관찰할 수 없는 경우에 전통적인 형태의 조직에 관한 평판(명성)보다 온라인 소비자 리뷰를 선호한다고 하였다. 이 조사 결과를 원용하면 소규모 NPO일수록 기금 모금 또는 광고를

통해 자신들의 품질을 알리는 데 필요한 자원이 상대적으로 제한적이므로 따라서 잠재적 기부자는 소규모 NPO에 기부할 때 제3자 평가에 더 많은 가치를 부여할 것으로 추정할 수 있다. Yörük(2016)에서는 상대적으로 알려지지 않은 작은 NPO에 대한 정보가 적기 때문에 기부자들은 이런 단체에 대해 기부할 때에는 Charity Navigator와 같은 감시기관에서 제공하는 독립적인 제3자 평가에 의존할 가능성이 높다고 주장하였다. 가설2에서는 이런 연구 결과를 반영하여 NPO의 규모별(소규모, 대규모)로 등급이 기부금 수입에 미치는 영향이 어떻게 다른지를 검증한다.

가설 2: NPO의 규모에 따라 기부금 수입에 제3자 평가가 미치는 영향이 상이할 것이다.

### 2.4 기부자 유형, 등급, 기부금의 관계

개인기부자와 기업 및 단체기부자는 기부를 하는 동기가 상이하다. 먼저 개인 기부자는 무수히 다양한 이유로 기부하지만 본 연구와 관련이 있는 내재적 동기를 위주로 보면 이타주의(43.3%), 개인적 신념(14.9%), 종교적 신념(7.9%)의 순으로 기부를 하는 것으로 나타났다(통계청 기부통계). 이들은 대부분 불우한 사람들을 돕는 개인적인 만족감이나 본인이 기부한다는 행동 자체에 기쁨을 느끼는 “따뜻한 빛(warm glow, 내적 기쁨)”에 의해 동기부여된다(Harbaugh, Mayr and Burghart, 2007).<sup>5)</sup> 이 “따뜻한 불빛(warm glow)”이 기부자의 지배적인 동기로 작용할 때 감시기관이 제공하는 객관적이

5) 이타적 행동으로 인하여 다른 사람의 안녕과 행복이 증진될 때 얻게 되는 기부자의 측은지심(惻隱之心)을 ‘따뜻한 빛(warm glow, 내적 기쁨)’이라고 한다. 특히 소액 기부자의 경우, 이들은 기부의 효과와는 상관없이 본인이 기부한다는 행동 자체에 만족과 기쁨을 느끼는 소위 “따뜻한 빛(warm glow, 내적 기쁨)” 기부자라고 할 수 있다(Singer, 2015).

고 완전한 정보를 얻는 것은 불필요하다고 생각할 수 있다(Cnaan et al., 2011).

이에 비해 기업 기부자는 개인기부자와 완전히 다른 접근 방식을 취하는데 이들은 기부를 발표하는 보도 자료, 언론에 게재된 기부금 전달 사진, 스폰서십 기부금과 같은 기부에 의해 발생하는 마케팅 기회를 중시하는 전략적 기부를 결정할 가능성이 크다(The Staff of Entrepreneur Media and Corbin Collins, 2017). 따라서 기업 기부자는 개인 기부자보다 완전하고 면밀한 회계 검토를 요구하므로 평가기관의 정보를 더 많이 사용할 것으로 기대된다. 이런 점을 반영하여 본고에서는 개인기부금과 기업 및 단체기부금 수입에 제3자 평가가 미치는 영향이 상이하다는 것을 가설3을 통해 검증한다.

가설 3: 개인기부금과 기업기부금에 제3자 평가가 미치는 영향이 상이할 것이다.

## 2.5 등급과 온라인 정보원의 상대적 유용성

등급이나 평점 이외에 기부자의 기부 결정에 영향을 미치는 중요 요인으로는 온라인 검색 엔진(구글 등)과 해당 NPO의 웹사이트 등이 있다. 시장조사 기관인 Harris Interactive가 정기적으로 조사하는 옴니버스 설문인 Donor Pulse 2008년도 조사 결과를 보면 해당 NPO에 관련된 정보를 얻기 위해 Google 등의 검색 엔진을 이용한다는 비율이 58.9%로 가장 많으며 다음으로 해당 조직의 웹사이트를 이용한다는 비율이 56.9%를 차지했다(Cnaan et al. 2011). 온라인 매체를 통한 기부 정보의 수

집에 대해 Google이 설문조사한 결과를 보면 NPO의 영향을 이해하는 데 사용되는 대부분의 상위 정보원은 디지털 매체(온라인 비디오, 비영리조직 웹사이트 및 검색)이며, 모바일에서는 검색(search)이 가장 유용한 정보원(87%)이며 NPO 웹사이트는 75%로 2위를 차지했다.<sup>6)</sup> 미국 연방거래위원회(Federal Trade Commission)의 소비자 정보 사이트에서도 자선단체에 기부하고 싶은 경우에 다음 세 가지 접근법을 제시했다. 첫째는 재난구호나 아동 보호와 같은 관심 있는 목적 대상과 “최고 등급의 자선단체”와 같은 문구로 기부자가 원하는 단체를 온라인으로 검색하는 것이다. 둘째는 자선단체의 웹사이트를 기부자들이 지원하고자 하는 프로그램 또는 기부금 사용 방식에 대해 어떤 정보를 제공하는지 확인하는 방법이다. 셋째는 자선단체를 분석하는데 도움이 되는 평가기관을 활용하는 것이다.<sup>7)</sup>

이와 같이 등급 이외에 기부금 수입에 영향을 미치는 인터넷 검색(미디어 노출)과 웹 사이트 이용에 관한 선행연구를 살펴보면 다음과 같다. 먼저 미디어 매체의 정보 효과가 기부금에 미치는 효과를 보면 Jen and Croson(2009)에서는 방송 기금 모금 캠페인을 설정한 실험실 기법을 사용해 사회적 정보가 여론을 조종하여 개인 기부금을 증가시키는 것으로 나타났다. Besiou and McCormick(2013)에서도 재난 대응 프로그램과 관련된 미디어의 양이 많을수록 해당 프로그램에 대한 기부 금액이 증가하였다고 주장하였다. Yörüük(2012)은 미디어(대중 매체)라는 중요한 도구를 통해 자선 단체가 잠재적인 기부자에게 그들이 제공하는 활동과 서비스를 알릴 수 있다고 하였다. 다음으로 웹 사이트 이용과 기부

6) Google에서는 설문을 통해 기부자들이 온라인으로 현금을 기부하는 양식(Patterns)을 조사하였다.

<https://www.pinterest.com/pin/323625923194050487/>

7) <https://www.consumer.ftc.gov/articles/0074-giving-charity>

금의 관계를 보면 높은 수준의 웹 공시를 구현하는 조직은 정보 공시 빈도 증가, 대중 신뢰 증진 및 기부금 증가와 같은 호순환의 장점을 경험한다고 하였다(Blouin and Lee, 2015). Saxton, Neely, and Guo(2014)도 웹 공시는 기부자에게 유용한 정보와 신호를 제공하기 때문에 비영리조직 웹 사이트에 공개된 품질 정보에 대해 기부자가 기부금을 기꺼이 지불한다는 것을 실증하였다.

본고에서는 이상의 연구결과를 종합하여 NPO에 관한 감시기관의 등급 부여와 인터넷 검색을 통한 미디어 노출 및 웹 사이트 이용이 기부금 수입에 미치는 상대적 영향력이 어떻게 다른가를 <가설 4>를 통하여 실증하고자 한다.

가설 4: 제3자 등급, 미디어 노출 및 웹 사이트 이용이 기부금 수입에 미치는 영향이 상이할 것이다.

또한 추가분석에서는 매년 연속적으로 별 3개 등급을 받은 NPO가 그렇지 않은 NPO에 비해 기부금 수입이 증대되었는지, 그리고 기부금수입 등 변수들의 극단치 조정을 반영한 분석에 관해서도 실증하고자 하였다.

### III. 데이터와 연구방법

#### 3.1 데이터 수집

본 연구에서는 (재)한국가이드스타가 국세청으로

부터 자료를 제공받아 정리한 데이터와 이를 근거로 재무안정성 및 효율성 12개, 책무성 및 투명성 12개 항목을 토대로 산출된 평가결과 정보를 사용하였다. 전자에는 공익법인의 사업유형, 기부금단체 유형, 자산보유, 수입금액과 필요경비, 그리고 세무확인과 회계감사 현황 등이 포함되어 있으며, 후자에는 100점 만점의 평가점수와 이를 별점(star ratings)으로 표시한 별1점~별3점의 정보가 있다.

(재)한국가이드스타에서 공익법인에 대한 평가를 실시할 때는 공시자료를 불성실하게 작성한 법인(비용이 대부분 0의 값 등)과 기부금이 0으로 나타난 법인 등을 제외하고, 그 다음에 이의신청 등의 의견과 법적 제외사항 등을 보다 세부적으로 확인하여 최종 평가대상 공익법인을 확정한다. 그리고 법적 강제력이 있는 평가가 아니므로 해당 단체에서 평가제외의 요청이 있으면 평가대상에서 제외시키는데 지금까지 평가대상은 평균적으로 전체 공익법인의 약 4~5%정도에 해당한다.

실증분석에 사용한 NPO에 관한 미디어보도는 네이버 상세검색과 언론진흥재단의 기사통합검색시스템인 KINDS를 이용하였다. 해당 NPO의 웹 사이트 관련 자료는 해당 조직의 홈페이지 등을 통해 수집하였으며 기타 정보는 Kisline웹서비스와 국세청 공익법인 데이터베이스를 활용하였다. 이런 과정을 거쳐 최종적으로 확정된 공익법인 105개(312법인/년)의 연도별 사업유형과 법인 유형을 보면 다음 <표 1>과 같다.<sup>8)</sup>

#### 3.2 연구 모형

먼저 NPO가 받는 자선 기부금에 대한 등급(평

8) 연도별 표본 수를 구체적으로 보면 2016년 105개, 2017년 105개, 2018년 102개로 구성되어있다. 2018년도에 3개 단체가 의무공시가 아닌 자율공시를 하여 데이터에서 제외되었다.

〈표 1〉 데이터의 연도별 자료

연도	사업유형					법인유형				
	교육	학술 장학	사회 복지	문화	기타	재단 법인	사단 법인	인가 단체	공공 기관	기타 기관
2016년	3	18	41	15	28	62	26	3	3	11
2017년	3	18	41	15	28	61	26	3	3	12
2018년	3	17	41	14	27	58	24	2	3	15
합계	9개 2.9%	53개 17.0%	123개 39.4%	44개 14.1%	83개 26.6%	181개 58.0%	76개 24.4%	8개 2.6%	9개 2.9%	38개 12.2%

점)의 영향을 추정하는 것으로 시작한다. 선행연구에 이어 데이터 세트의 패널 구조를 이용하여 시간-불변의 산업과 연도 고정효과를 통합한 다음 모형을 추정한다.

$$Y_{it} = \beta_0 + \beta_1 RATING_{it-1} + \beta_2 CONTROL_{it-1} + \gamma_i + \mu_t + \epsilon_{it} \quad (1)$$

#### Dependent Variables

$GIVING_i$  = 기부금 수입의 자연로그

$GIVING1$  = 개인 기부금의 자연로그

$GIVING2$  = 기업 기부금의 자연로그

#### Independent Variables

$RATING$  = 등급 범위: 3 star  $\geq 90$ 점,  $80 \leq 2$  star  $> 90$ ,  $70 \leq 1$  star  $> 80$ , 0 star  $< 70$ 점

$SCORE$  = 실제 평점 (최저 54점 최대 100점)

#### Control Variables

$AGE$  = 조직의 운영연수(설립이후의 운영기간);

$ASSETS$  = 총자산의 로그;

$ADMIN$  = 관리비/총비용;

$PRICE$  =  $1 / \{1 - (\text{기금모금비용} + \text{관리비용}) / \text{총비용}\} = 1 / (\text{프로그램 비용} / \text{총비용})$

$\gamma_i$  = 산업 고정효과(주요 다섯 가지 각각의 법인유형에 대한 더미변수)

$\mu_t$  = 연도 고정효과(각 회계연도에 대한 더미변수)

$\epsilon_i$  = 관찰치 i에 대한 임의의 오차항

식 (1)에서  $Y_{it}$ 는 자선단체 i가 t기에 수취한 총 기부금액의 자연로그를 취한 값이다. 주요 관심계수인  $\beta_1$ 는 전기의 종합적 등급이 당기의 기부금 총액에 미치는 영향을 나타낸다. Chhaochharia and Ghosh (2008) 등에 따라 당기에 자선단체가 기부한 금액은 전기의 성과와 등급에 따라 다르므로 전기의 통제변수 세트를 사용한다. 이러한 변수는 CONTROL로 표시되며 여기에는 AGE, ASSETS, ADMIN 및 PRICE 변수가 포함된다.<sup>9)</sup> 먼저 ADMIN(사업관리비 비율)은 사업관리비가 총비용에서 차지하는 비율이며, PRICE 변수는 총사업비 중에서 사업수행(프로그램) 비용이 차지하는 비율의

9) 기부금 의사결정에 외부감사 여부가 주요 영향변수이지만 3년간 315개 샘플 중에서 9개 단체를 제외하고는 모두 외부감사 자료를 제출했으며 비적정 감사의견이 없어 통제변수에 포함하지 않았으며, 법정기부금 또는 지정기부금 여부에 따라 공제한도가 달라지므로 기부 의사결정에 영향을 미칠 수 있으나 전체 표본 중에서 약 1/3은 법률에 의한 모금액이 없으며 나머지도 모두 법률에 의한 모금, 개인기부, 기업 및 단체기부, 기부물품 등이 혼재되어 있어 통제변수에서 제외하였다.

역수로서<sup>10)</sup> 1달러의 자선 서비스를 행하기 위해 들어가는 비용 즉 기부자가 1달러의 기부성과(산출물)가 나타나기 위해 NPO에 지급해야 하는 금액(가격)을 의미한다(Roberts, Smith, and Taranto, 2006). NPO의 규모는 자산의 자연로그 값이며 AGE는 NPO가 면세 자격을 얻은 이후의 운영년수에 대해 자연로그로 측정한다(Gordon et al. 2009). 고정산업 효과는  $\gamma_i$ 로 표현되며  $\mu_t$ 는 고정연도효과를 나타낸다.  $\varepsilon$ 은 임의의 오차항으로 독립적이고 동일하게 분포된다고 가정한다.

다음으로 본 연구에서는 가설을 검증하기 위해 회귀불연속설계(RDD: regression discontinuity)를 사용하여 NPO(공익법인)에 대해 (재)한국가이드스타가 부여한 등급에서 별첨 한 개의 증가가 NPO가 수령한 기부금 금액에 어떠한 영향을 미치는지 분석한다.

RDD란 어떤 연속 변수(Z) 값이 특정 절단점(cutoff)<sup>11)</sup> 보다 높으나에 따라 처리집단(T=1)에 할당되거나 통제집단(T=0)에 할당될지가 정해지는 현상을 이용함으로써 프로그램 효과를 추정하는 방법이다. 이 방법은 연속변수 Z(본고에서는 평점)를 따라 결과(기부금)를 보러 가다 보면 절단점 값 근처에서는 Z의 값은 거의 변하지 않음에도 불구하고 절단점의 한쪽은 처리집단(실험집단), 다른 한쪽은 통제집단(비교집단)에 할당되어 버리는 현상이 생겨난다. 예를 들어, 전체 평점이 90에서 91로 증가하면 기부금에 작거나 또는 통계적으로 유의한 영향을 미치지 않을 것으로 예상할 수 있다. 이 두 가지 평

점에 대한 등급은 모두 별3개 등급에 해당하며 매우 우수한 평가 집단에 위치하는 것이다.<sup>12)</sup> 반면 실제 평점이 89에서 90으로 증가하면 기부 수준에 상당한 영향을 미칠 수 있다. 그 이유는 평점 1점의 증가로 인해 별 등급(star rating)이 2등급에서 3등급으로 변경되기 때문이다.

이 절단점(90점)의 바로 양쪽 옆의 평점(89와 90)을 비교하면 이 두 자선단체는 상당히 유사한 품질의 조직이라는 것을 알 수 있으며, 유일하게 다른 것은 단지 처리집단(별2개 등급)이 되었는지 통제집단(별3개 등급)이 되었는지 하는 것뿐이며, 이 두 집단 간에 결과(기부금)를 비교하면 그 차이가 처리효과(Treatment effect) 즉 등급 차이의 효과라고 할 수 있다. 환언하면 Y축을 기부금 수준, X축을 평점으로 하여 각 데이터의 점에 따라서 선을 그어 가면 Z=90의 곳에서 연속성이 없어지게 되는데, 이(Y축 방향의) 간격(비연속)이 프로그램의 효과 그 자체가 된다는 것이다. 이 두 극한 기댓값의 차이가 배정점(c)에서 평가사업(등급)의 효과 즉 국지적(Local) 평균처리효과(ATE)를 의미한다(우석진, 2016). 그러므로 <그림 2> 등에서 이러한 RDD의 ‘불연속 또는 처리효과’는 단순한 절편의 차이를 의미하는 것이 아님을 알 수가 있다.

실증 분석을 통해 사용된 기본 RD 모형은 다음과 같다.

$$Y_{it} = \alpha + \delta T_{it-1} + f(X_{it-1}) + \zeta CONTROL_{it-1} + \mu_{it} + \nu_{it} \quad (2)$$

10) 총비용=사업수행비+모금·관리비 이므로 PRICE=1/{1-(모금비+관리비)/총비용}=1/{(사업수행비/총비용)}이 된다. 이 경우 사업수행비와 모금·관리비는 상호 수학적 연관성을 갖는다. 따라서 사업수행비용이 총비용에서 차지하는 비율의 역수(PRICE)를 취한다(Gordon et al. 2009; Saxton et al. 2014).

11) 등급 수혜의 기준이 되는 절단점(cutoff point)은 문맥에 따라 배정점, 임계점 등으로 적절하게 사용된다.

12) (재)한국가이드스타에서는 평점이 100점 만점으로 할 때 90점 이상은 별3개 등급, 80 이상 90 미만은 별2개 등급, 70이상 80 미만은 별1개, 70 미만은 0 등급을 부여하고 있다.

$$\forall x_i \in (c-h, c+h)$$

$\alpha$  = 등급 변수 통제 후의 처리집단에 있어서의 결과 (기부금)의 평균값

$T_i$  = 관찰치  $i$ 가 처리집단에 배정되면 1 아니면 0

$X_i$  = 관찰치  $i$ 에 대한 개별 평점, 절사점(cutoff)이 중심에 위치

$CONTROL_i$  = 통제변수로서 AGE, ASSETS, ADMIN, 및 PRICE 변수

$\mu_i$  = 연도 고정효과

$\nu_i$  = 관찰치  $i$ 에 대한 임의의 오차항.

여기에서  $Y_{it}$ 는 문제의 결과(기부금 수준)이고, 처리변수  $T$ 는 이항변수로서 전기의 특정 자선단체 등급이 전체 등급의 임계값보다 크면 1이고 그렇지 않은 경우 0의 값을 취한다. 이 변수에 대한 계수  $\delta$ 는 주요 관심계수이며 절단점에서 프로그램의 한계 영향을 나타내는 것으로 전기의 자선단체 등급에서 별점 1개 증가가 당기에 받은 기부금액에 미치는 영향을 측정한다. 통제함수  $f(x_i)$ 는 배정점  $c$ 의 양쪽 측면에 대한 연속적인 실행변수(강제변수)의 평활함수이고 다차항과 교호항을 포함한다.<sup>13)</sup> 식(2)에서  $X_i$ 는 실행변수이고,  $h$ 는 배정점  $c$ 의 위나 아래의 근거리에서 있는 관찰치로서 대역폭(bandwidth)이라고 한다.  $X_{t-1}$ 은 전기( $t-1$ 기)의 실제 평점과 컷오프 평점 간의 차이이다.

기부에 영향을 줄 수 있는 다른 관찰 가능하거나 관찰 불가능한 NPO의 특성이 종합 등급(별1개, 별2개, 별3개)의 임계값 근처에서 유사하게 남아 있기 때문에, 이런 측면에서 RD 분석의 가정은 종합 등급의 임계값에서 수취한 기부금의 변화는 전적으로 별 등급의 변경으로 인한 것임을 의미한다. 이것은

별 등급을 제외하고는 실제 등급의 임계값에서 NPO의 다른 관찰 가능한 특성에 중요한 변화가 없어야 함을 의미한다.

강제변수의 평활함수(smoothing function)를 정확하게 모형화 하는 것은 RD 설계를 구축할 때 발생하는 중요한 문제이다. 평활함수  $f(x_i)$ 는 등급 변수와 결과의 관계를 나타내는데 어떤 것이 데이터에 가장 적합한지를 결정하기 위해 다양한 함수 형태를 테스트하여 편향(bias)을 최소화할 수 있다. 즉 강건성을 테스트하기 위해 강제변수의 1차(선형), 2차, 3차 다항식 형태의 다양한 모형을 추정하고 또한 여기에 처리변수와 상호작용하는 항을 더한 모형이 RD디자인의 모수 분석에 사용된다. 대표적으로 3차 다항식과 상호작용(cubic interaction)을 포함하는 강제변수의 평활함수 모형을 보면 다음과 같다.

3차 방정식

$$Y_i = \alpha + \delta_0 \cdot T_i + \delta_1 \cdot X_i + \delta_2 \cdot X_i^2 + \delta_3 \cdot X_i^3 + \nu_i \quad (3)$$

3차 상호작용

$$Y_i = \alpha + \delta_0 \cdot T_i + \delta_1 \cdot X_i + \delta_2 \cdot X_i^2 + \delta_3 \cdot X_i^3 + \delta_4 \cdot X_i \cdot T_i + \delta_5 \cdot X_i^2 \cdot T_i + \delta_6 \cdot X_i^3 \cdot T_i + \nu_i \quad (4)$$

실제 RD모형의 실증분석에 있어서는 NPO의 전반적인 평가가 컷오프 등급보다 일정한 점수 이상 낮거나 높은 샘플을 제한하는데 그것은 RD 추정치는 이 대역폭의 선택에 민감하기 때문이다. 그런데 지금까지 대역폭의 선택에 대한 광범위하게 인정된 기준이 비모수(non-parametric) 접근법에는 없으

13) 평점 변수(범위: 최저 60점에서 최고 100점)는 연속 변수로서 한국가이드스타가 각 NPO단체에 부여한 평점을 몇 등급(예를 들어 2등급 또는 3등급)에 할당(배정)할지 여부를 결정하는 값이다. 강제변수, 실행변수 또는 할당변수라고도 한다(Jacobs and Zhu 2012).

므로 본고에서는 모수(parametric) 접근방식을 사용하여 전체 데이터 세트에 맞는 등급 변수와 결과 사이의 최적 함수 형태를 찾는 데 초점을 맞춘다. 비모수 추정법은 절단점(cutoff) 부근의 데이터에 의한 추정량을 얻을 수 있으며 직감적으로도 알기 쉽다는 장점이 있다. 그러나 국지적 선형회귀((local linear regressions)로 대표되는 비모수 전략은 대부분의 경우 분석에 사용되는 표본 크기가 작기 때문에 통계적 검증력이 제한되는 단점이 있다(Jacob and Zhu 2012). 이에 본 연구에서는 표본 크기의 한계를 고려하여 처리효과의 추정에 이용 가능한 모든 데이터를 사용하고 비모수 접근방식보다 잠재적으로 더 높은 정밀도를 제공할 수 있는 모수 접근법을 채택한다. 구체적으로 본 연구에서는 처리효과에 대한 더 정확한 추정치를 유도하기 위해 공변량의 벡터(즉 *CONTROL* 변수)를 추가하는 반모수(semi-parametric)모형 방정식을 사용한다.

끝으로 식 (5)에서 등급이나 평점 이외에 기부자의 기부 결정에 영향을 미치는 중요 요인인 온라인 검색 엔진(미디어 노출)과 해당 NPO의 웹사이트 이용의 상대적 중요성을 파악하고자 한다. Cnaan et al. (2011) 등의 조사 결과를 보면 기부자들이 기부하고자 하는 NPO에 관련된 정보를 얻기 위해서는 Google 등의 검색 엔진을 이용한다는 비율이 가장 높고 다음으로 해당 NPO의 웹 사이트 활용, 평가기관의 등급이나 순위가 사용되는 것으로 나타났다. 회귀식(5)에서는 등급 이외에 인터넷 검색(미디어 노출)과 해당 NPO의 웹 사이트 이용이 기부금 수입에 미치는 상대적 영향력이 어떻게 다른가를

실증하고자 한다.

$$Y_{it} = \beta_0 + \beta_1 RATING_{it-1} + \beta_2 WEB_{it-1} + \beta_3 COVG_{it-1} + \beta_4 CONTROL_{it-1} + \gamma_k NPO_i + \delta_j YEAR_{it} + \epsilon_{it} \quad (5)$$

여기서 *RATING<sub>t</sub>*은 한국가이드스타가 부여한 등급(평점), *WEB<sub>t</sub>*는 해당 NPO 웹 사이트의 공시 수준 그리고 *COVG<sub>t</sub>*는 검색엔진을 활용한 NPO의 미디어 노출 빈도를 나타낸다. 전자인 웹 사이트에서 제공하는 공시정보는 Saxton et al.(2014)에서와 같이 성과 정보(4점 만점), 연차보고서 공시 여부(1점 만점), 재무공시(3점 만점)의 세 가지로 측정한다.<sup>14)</sup> 전체 웹 사이트 공시 수준은 이 세 가지로 측정된 값을 모두 더한 것이며 최소 0에서 최대 8점이다. 후자인 미디어 노출 빈도는 과거 2년 동안(t-1, t-2) 언론진흥재단의 기사통합검색시스템인 KINDS와 네이버 검색엔진을 사용하여 해당 NPO의 법인명과 대표자를 웹 크롤링으로 수집하고 이 중에서 긍정적인 기사 횟수를 취사선택하여 이를 바탕으로 미디어 노출 빈도를 측정하였다.

## IV. 실증분석 결과

### 4.1 기술통계와 상관계수

본 연구에 사용된 변수의 기술통계는 <표 2>에 나

14) 온라인 재무정보 공시 지수는 각 조직 웹 사이트의 수입지출명세서, 재무제표, 감사보고서 항목의 합계로 정의한다. 예를 들면 세 항목 모두 웹 사이트에 공시되면 3점, 하나의 항목만 나타나면 1점으로 측정한다. 성과 정보에는 조직의 미션, 비전, 가치, 목표에 관련된 정보가 포함되는데 네 가지가 모두 공개되면 4점, 한 가지만 공개되면 1점으로 측정한다. 연차보고서 공시는 해당 보고서를 웹 사이트에 공개하면 1, 그렇지 않은 경우 0의 값을 부여한다.

〈표 2〉 기술적 통계량

	n	통계량	기부금	개인 기부금	단체 기부금	총자산	운영 년수	프로그램 경비	사업 관리비
전체 표본	312	평균	6,960	5,462	2,860	18,007	19	7,538	1,723
		표준편차	22,272	18,911	6,821	29,408	15	21,381	3,331
		최대값	156,590	128,354	55,065	195,198	70	159,334	18,275
		최소값	17	1	4	6	3	4	1
3 star	112	평균	15,960	11,743	5,737	23,120	18	15,890	2,209
		표준편차	34,430	27,726	9,831	29,740	14	31,992	4,224
2 star	160	평균	2,160	1,072	989	17,008	19	3,494	1,391
		표준편차	6,855	4,315	2,279	31,181	15	7,892	2,656
1 star	30	평균	973	411	1,063	9,990	16	1,244	1,183
		표준편차	1,711	876	2,019	18,955	12	2,123	2,197
대규모	158	평균	13,004	9,405	4,543	34,056	20	14,413	2,842
		표준편차	29,996	24,481	8,522	34,731	15	28,096	4,273
소규모	154	평균	675	363	491	1,852	17	1,033	596
		표준편차	844	593	718	1,372	15	1,387	1,188

주) 운영년수를 제외한 모든 변수는 100만원 단위이다.

등급(평점) 범위: 3 star ≥ 90점, 2 star ≥ 80 but < 90점, 1 star < 80점

와 있다. 연말 기준 기부금 수입은 최소 17백만 원에서 최대 156,590백만 원이며 평균 6,960백만 원이다. 총자산은 최소 6백만 원에서 최대 195,110백만 원이며 평균 자산은 18,000백만 원으로 나타났다. NPO들은 비용의 81%가 프로그램 서비스 제공에 지출했으며, PRICE 즉 기부자가 1원의 임무 관련 산출물에 대해 평균 2.67원, 중앙값 1.26원, 최대 14.7원의 비용을 지불할 수 있음을 보여주고 있다.

또한 전체 표본에서 대부분의 NPO단체는 별 2개와 3개로 평가된다는 것을 보여준다. 별 3개와 별 2개의 통계량 차이는 운영년수를 제외하고 기부금 수입( $t=4.18$ ), 총자산( $t=1.67$ ), 프로그램 경비( $t=4.01$ ), 사업관리비( $t=1.82$ ) 모두에서 유의적인 차이가 발생했다. 별 2개와 별 1개의 통계량 차이도 모든 항목에서 유의하게 나타났다. 이것은 별 3개 조직은 등

급이 낮은 조직보다 기부금 수입을 비롯한 자산과 모금비용 등 운영 규모가 크고 이로 인한 경제적 효과가 우수할 가능성이 더 높은 것을 의미한다. 대규모와 소규모 조직의 통계량 차이는 운영년수를 제외하고는 모든 항목에서 유의수준 0.01 수준에서 대규모 조직이 크게 나타났다.

〈표 3〉에는 NPO에 관한 등급(평점)이 높을수록 기부금 수입이 증가할 것이라는 가설 1을 검증하기 위한 기본 모형 식 (1)에 대한 결과가 제시되어 있다. 먼저 NPO의 회계비율 변수가 기부금 크기에 미치는 영향을 보면 운영년수(AGE)와 자산 규모(ASSETS)는 기부금에 모두 긍정적이고 유의적인 반면 사업관리비(ADMIN)는 부정적 영향을 미치고 있다. 가격 변수(PRICE)는 통계적으로 유의하지 않았다. 가설 1에서 예상한 바와 같이 별 등급과 평점은 모두 기

〈표 3〉 평점과 등급이 기부금 수입에 미치는 영향

설명변수 \ 모형	모형 1	모형 2	모형 3
Constant	4.579(0.459)***	3.343(0.58)***	3.339(0.916)***
SCORE			0.041(0.016)***
RATING		0.536(0.197)**	
AGE	0.011(0.006)*	0.010(0.006)*	0.011(0.006)*
ASSETS	0.358(0.049)***	0.341(0.049)***	0.346(0.049)***
ADMIN	-3.863(0.440)***	-3.229(0.471)***	-3.309(0.491)***
PRICE	0.007(0.011)	0.008(0.010)	0.007(0.010)
IND	Yes	Yes	Yes
YEAR	Yes	Yes	Yes
N	312	312	312
F-value	46.93	39.40	41.21
Adj. R <sup>2</sup>	0.384	0.394	0.405

주) '유의수준: '\*\*\*' 0.01 '\*\*' 0.05 '\*' 0.1' 그리고 ( ) 내는 표준오차

부금 수입에 긍정적 영향을 미쳤으며 통계적으로 유의하다. 다만 별 등급이 평점보다 기부금과의 관련성이 상당히 강하게 나타났는데(별 등급의 계수는 0.53이며 평점의 계수는 0.04), 이것은 기부자들이 기부 의사결정을 위해 평점을 사용하기 보다는 별 등급에 대한 선호도가 높다는 것을 시사한다. 별 등급이 기부자들에게 보다 유용한 정보원이 되는 것은 전체 평점에 비해 별 등급과 같은 정성적 등급이 비교적 쉽게 해석할 수 있기 때문일 것이다(Yörük 2016). 실제 평가기관인 한국가이드스타에서도 전체 평점보다 별 등급을 많이 홍보하고 있는 실정이며 웹사이트의 메인 페이지에서도 “별 3개 공익법인”은 공시하지만 NPO의 실제 전체 평점은 보고하지 않고 있다.

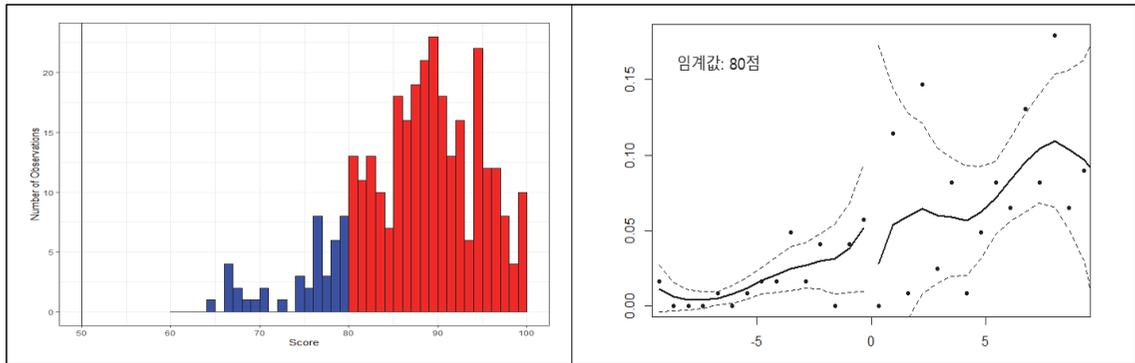
#### 4.2 RD 분석 결과

본장에서는 이와 같이 NPO에 대한 특정 등급보

다 약간 높거나 낮은 평점이 기부금 성과에 미치는 영향을 정교하게 비교하여 경계선 주변에서의 등급 효과를 식별 및 추정할 수 있는 RDD를 적용한 결과를 제시한다.

##### 4.2.1 예비분석 결과

RDD의 타당도를 검토하기 위해서는 효과를 추정하기 이전에 그래픽 분석을 통해 실행변수에 대한 히스토그램이나 밀도검사(density test)가 일반적으로 사용된다. RD 결과로 이동하기 전에 두 가지 표준 유효성 검사(Imbens and Lemieux 2008)를 제시한다. 첫째, 그래픽 분석의 목적은 실행변수의 분포가 절단점의 양쪽이 유사한지를 확인하기 위한 것인데 SCORE 변수는 연속성과 평가점수의 절단점(0)에서 좌우 대칭이 유사한 정규성에 가까운 분포를 보이고 있다. 〈그림 1〉의 왼편은 먼저 전체 범위에 대한 강제변수의 히스토그램을 양쪽 2%(-1~+1)



주) 이 그림은 실제 평점의 밀도가 임계점에서 연속적인가를 분석한 것이다. 왼쪽 패널은 2016년-2018년 전체 범위에 대한 실제 평점의 히스토그램을 1% 간격으로 표시했다. 오른쪽 패널의 그래프는 실제 평점의 밀도에 불연속성이 있는지 여부에 대한 McCrary (2008) 테스트 결과를 보여준다. 여기서 실선은 밀도함수, 점선은 95% 신뢰구간임.

〈그림 1〉 등급(평점) 변수의 유효성 검사

의 간격으로 나타낸다. 여기서 임계점 90점 전후에서는 임계수준 주위에 NPO 수가 이산적으로 차이가 관찰되나 임계점 80점 전후에서는 좌우 대칭을 보이고 있다. 임계점 90점에서도 경계선 오른쪽 등급(평점)을 받은 NPO 수가 감소하기 때문에 정책(강제)변수를 조작하여 등급 향상 정책에 참여했을 가능성은 낮아 보인다.

둘째 〈그림 1〉의 오른쪽은 McCrary(2008)가 제시한 밀도검사를 이용하여 강제변수(SCORE)의 조작으로 인한 임계 등급 주변의 NPO들의 무작위 분류(sorting) 가능성을 검증한 결과이다. 이 그래프에서 절단점을 보면 0으로 표시되어 있는데, 이전에 배정점의 값(예, 90점)을 각 관측치가 가지는 실행변수의 값에서 차분하여 중심화(centring)하였고 결과적으로 분석의 배정점은 0이 된 것이다. McCrary의 밀도검사 결과를 보면 전체 표본의 경우에는 불연속 추정치는 -1.308이고 유의수준(p) |T|: 0.091)로서 강력하지는 않으나 평점에 대한 NPO의 무작위 분류가 존재한다는 귀무가설은 기각되지 않는다. 즉 경계선 근처에서 밀도함수의 차이가 발생한다고 보기

어렵다고 할 수 있다. 또한 본 연구에서 사용하고 있는 자료가 매년 모든 동일한 NPO를 대상으로 규칙적으로 이루어지는 공적기관의 공개 패널자료이므로 자료에서의 조작의 가능성은 적다고 판단할 수 있다.

셋째 RDD가 성립하기 위한 주요 가정 중 하나는 특성 변수들의 경계선 근처에서의 연속성이다. 이것은 등급의 처리효과가 NPO가 가진 특성들의 불연속성에 기인한 것인지를 확인하기 위해서 이다. 〈표 4〉에서 RDD에 의해 4가지 특성 변수들을 종속변수로 하여 분석한 결과를 보면 대부분의 경우 불연속이 발생하지 않았다. 다만 전체 표본을 대상으로 분석한 경우 총자산은 유의수준 10% 수준에서 불연속이 발견되나 이것도 별 등급을 세분한 표본에서는 불연속은 나타나지 않았다. 즉 각 공변량들은 배정점에서 매끄럽게(smoothness) 이어지고 있어 기부금 수입의 불연속성이 발생한 이유는 정책 효과 때문이지 NPO 특성에 있어서 불연속이 존재하기 때문은 아닌 것으로 판단된다.

〈표 5〉에서는 별 등급이 1개, 2개 또는 3개 주위의 NPO에 대해 RD를 각각 추정하기 위해 강제변

〈표 4〉 배정점 전후에서의 통제변수(NPO의 특성)의 연속성 검증

	AGE	ASSETS	ADMIN	PRICE
전체 표본	0.898 (3.335) [312]	1.010* (0.502) [312]	-0.061 (0.048) [312]	2.285 (1.957) [312]
Around 3 star	0.193 (3.403) [218]	0.601 (0.422) [218]	-0.031 (0.043) [218]	0.311 (0.344) [217]
Around 2 star	0.874 (0.491) [94]	0.519 (0.642) [94]	-0.102 (0.078) [94]	-2.024 (3.841) [94]

주) 통제변수의 정의는 연구모형을 참조하기 바란다.  
 각 셀의 상단은 RD의 회귀계수, ( ) 내는 표준오차, [ ] 내는 표본수를 나타낸다.  
 \*유의수준: \*\*\*\* 0.01 \*\*\* 0.05 \*\* 0.1 그리고 ( ) 내는 표준오차

〈표 5〉 별 등급별 평점과 기부금의 관계

	전체	3 star	2 star	1 star
RDD(모수)	0.426(0.431)	1.455(0.403)	-0.543(0.505)	0.668(1.284)
다중선형회귀	0.404(0.429)	1.447(0.376)	-0.521(0.486)	0.592(1.102)
RDD(통제변수)	0.433(0.456)	1.475(0.481)	-0.615(0.406)	0.672(1.329)

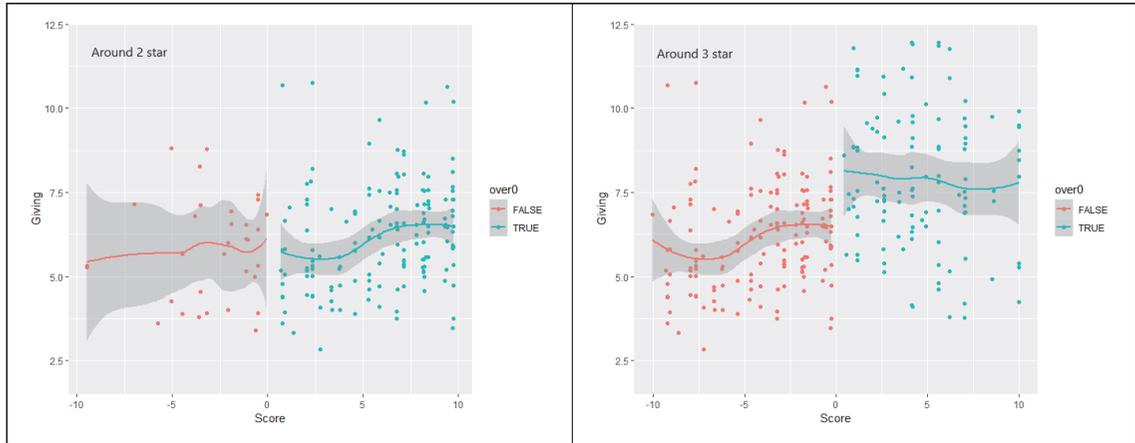
RDD 모형:  $Y_{it} = \alpha + \delta T_{it-1} + f(X_{it-1}) + \nu_{it}$   
 다중선형회귀 모형:  $Y_{it} = \beta_0 + \beta_1 X_{it-1} + D2 + \epsilon_{it}$  (D2는 더미로서 통제집단은 0 처리집단은 1)  
 RDD(통제변수): 기본 RDD 모형에 통제변수와 더미변수를 포함한 것이다.  
 검증 결과 값의 원편은 회귀계수 오른쪽은 표준오차를 나타낸다.  
 3 star는 2 star는 1 star는 각각 임계값 90점, 임계값 80점, 임계값 70점후의 데이터를 의미한다.

수의 3차 다항식을 포함하는 모형의 결과를 제시한다. 별 2개에서 3개로 한 개의 별 등급 증가(임계값 90점)는 자선 기부금 1.4배의 증가와 관련이 있다. 단, 이것은 임계값 90 전후에서 일정한 대역폭 내에서 별 등급의 증가로 인한 효과를 의미한다.<sup>15)</sup> 한편 별 1개에서 2개로 증가한 경우에는 임계점에서 유의미한 기부금 수입의 증가는 나타나지 않았다.

이 결과를 그래프로 확인해 보면 〈그림 2〉의 패널

B는 결과 변수의 예측치를 나타내며, 별 등급 2개에서 3개로 별 등급의 상승이 임계점에서 통계적으로 유의한 기부금 수입의 증가(jump)가 있음을 보여준다. 별 등급 1개에서 2개로 상승하는 경우에는 임계점에서 기부금 수입의 불연속이 나타나지 않으며 별 등급이 0에서 등급 1로 높아지는 경우에도 역시 유사한 결과가 발생한다. 이 결과는 별 등급이 1개, 2개 3개에 평점이 약간 못 미치는 NPO들이 별 등급을

15) 이것은 대역폭(bandwidth)을 임의로 지정하지 않고 RDD 소프트웨어에서 최적 대역폭( $|D| \leq 7.78$ )을 자동 설정하여 계산한 결과이다. 본고에서 대역폭을  $|D| \leq 3$ 이나  $|D| \leq 4$ 로 설정하여 추정했으나 유의한 결과가 도출되지 않았다. 이는 충분하지 않은 표본의 크기에 기인한다고 판단된다.



주) 여기서 3 star는 임계값 90점 전후의 평점이며 2 star는 임계값 80점 전후의 데이터를 의미함  
 각 점은 0.2점 간격의 평점을 갖는 모든 관측치(NPOs)의 평균을 나타낸다.  
 실선은 컷오프 평점(각각 80점과 90점)의 양쪽에 있는 실제 관측치에 의거한 3차 다항식이다.

<그림 2> 배정점(cutoff point) 전후에서 제3자 등급과 기부금 크기의 관계: 전체 표본

1개 상승시키기 위해 관리한다면 기부금이 3-5% 증가할 수 있다는 Grant(2010)의 결과와,<sup>16)</sup> 소규모 NPO를 제외하면 별 등급의 증가가 기부금에 미세하게 영향을 미치거나 유의한 영향을 미치지 않는다는 Yörük(2016)의 결과와는 다소 차이가 있다. 이것은 제3자 등급 평가기관(한국가이드스타)에서 별 3개 등급 위주로 평가결과를 공개하는 점과 무관하지 않는 것으로 생각되며 해당 NPO들도 자신들의 웹 페이지에 별 3개 등급이 아닌 경우에는 가능한 등급 정보 자체를 공시하지 않는 현상과도 관련이 있다고 할 수 있다.

다음으로 <표 6>에서는 기부자들은 상대적으로 덜 알려진 소규모 NPO에 대한 정보가 적기 때문에 이런 단체에 기부할 때에는 기부금 감시기관에서 제공하는 독립적인 제3자 평가에 의존할 가능성이 높다는 가설 3에 대한 분석 결과를 보여준다. 먼저 임계

점(80점, 90점)과 NPO 규모별로 구분하여 강제변수(SCORE)의 3차 다항식을 포함하는 모수 모형의 결과를 살펴본다. 첫째 대규모 NPO에서 임계점 90점의 전후에서 별 등급이 2개에서 3개로 상승하면 기부금 액수의 1.4배 증가와 관련이 있음을 나타낸다. 이 효과는 평점 변수의 선형 다항식과 2차 다항식 모형에서도 통계적으로 유의하고 긍정적 결과를 산출하였다. 또한 비모수 모형과 다중선형회귀 모형에서도 유사한 결과를 보여 주었다. 둘째 대규모 NPO에서 임계점 80점의 전후에서 별 등급이 1개에서 2개로 상승하면 기부금 증가에 대한 유의한 효과가 나타나지 않았다. 이 결과는 별 1개에서 별 2개로 늘어난 것은 NPO가 조악한(poor) 품질 등급에서 평균적인 등급으로 품질을 높일 수 있었다는 것을 암시하지만, 여전히 좋은(very good) 등급의 많은 NPO에 훨씬 못 미치는 평가를 받고 있기 때문

16) RDD를 적용한 Grant(2010)의 연구에서도 별 등급 0에서 1로 증가한 경우에는 기부금에 유의한 영향을 미치지 않았다.

〈표 6〉 NPO 규모별 등급(star ratings) 차이와 기부금의 관계

모형 \ 규모	대규모(3 star)	대규모(2 star)	소규모(3 star)	소규모(2 star)
RDD회귀(모수)	1.437(0.401)	-0.543(0.505)	0.429(0.409)	-0.877(0.537)
RDD(통제변수)	1.462(0.481)	-0.521(0.486)	0.728(0.498)	-0.767(0.603)
다중선형회귀	1.183(0.388)	-0.719(0.448)	0.135(0.380)	-0.807(0.422)

주) RDD 모형:  $Y_{it} = \alpha + \delta T_{it-1} + f(X_{it-1}) + \nu_{it}$

MLR 모형:  $Y_{it} = \beta_0 + \beta_1 X_{it-1} + D2 + \epsilon_{it}$  (D2는 더미변수로서 통제집단은 0 처리집단은 1)

RDD(통제변수)는 기본 RDD 모형에 통제변수와 더미변수를 포함한 것이다.

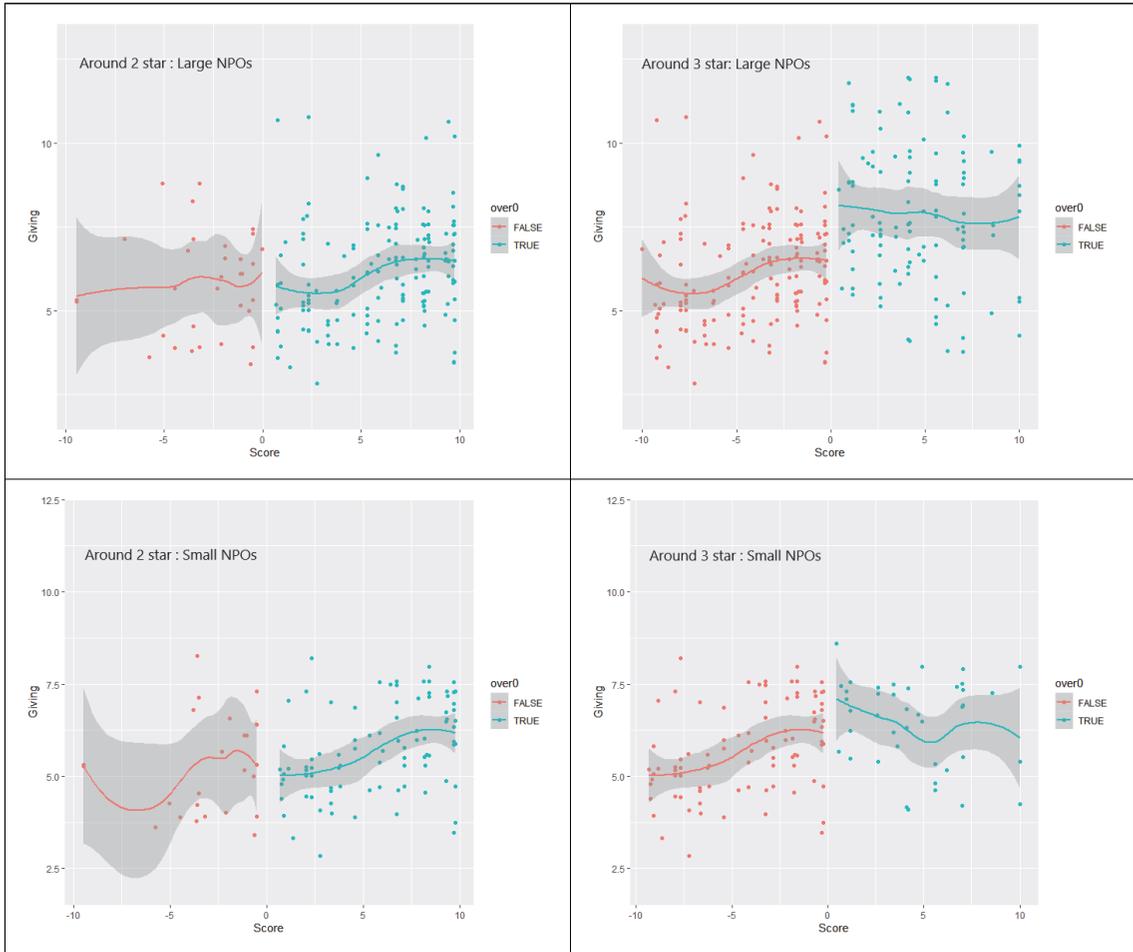
검증 결과 값의 원편은 회귀계수 오른쪽은 표준오차를 나타낸다.

에 놀라운 결과가 아니다. 즉 별 등급이 1개에서 2개로 증가하는 것은 여전히 해당 NPO가 허용 가능한 수준 이하라는 신호를 기부자에게 보낸 것이다. 셋째 소규모 NPO에서는 모든 임계점(80점, 90점) 전후에서 별 등급의 증가가 기부금 증가에 아무런 영향을 미치지 않았다. 소규모 NPO에 기부하는 자는 상대적 홍보를 많이 하는 대규모 NPO에 비해 대중의 인지도는 낮아도 자신들의 기부목적에 맞는 단체를 선택하여 기부한다는 행동 자체가 지배적인 동기로 작용하므로 감시기관이 제공하는 등급 정보를 얻는 것은 불필요하다고 생각할 수 있다.

이 결과를 그래프로 확인해 보면 〈그림 3〉의 패널에서 대규모 NPO가 별 1개에서 2개로 증가한 경우와, 소규모 NPO에서 별 1개에서 2개 그리고 별 2개에서 3개로 증가한 경우는 모두 임계점에서 기부금의 증가(jump) 즉 회귀불연속이 발생하지 않았다. 이에 비해 대규모 NPO의 경우에 별 등급 2개에서 3개로 별 등급의 상승이 임계점(90점)에서 통계적으로 유의한 기부금 수입의 증가(jump)가 있음을 보여준다. 이것은 본 연구의 대규모 NPO에서는 개인 기부금 비중이 기업(단체) 기부금 보다 평균 2배 이상 높기 때문에 상대적으로 NPO 정보 수집에 취약한 개인 집단이 객관적인 등급정보에 더욱 의존한

결과라고 해석할 수 있다. 즉 기부자들이 대규모 NPO에 기부할 때 상대적으로 높은 별 등급(3개)만이 기부금 수입에 긍정적인 신호 가치를 부여한다는 것을 알 수 있다. 이는 대규모 조직은 등급을 강조하는 언론의 관심 등에 더 취약할 수 있다는 것을 반증한다고도 할 수 있다(Szper and Prakash, 2011).

다음으로 기업이나 단체기부자는 개인 기부자보다 완전하고 면밀한 회계 검토를 요구하므로 평가기관의 정보를 더 많이 사용할 것이라는 가설 3을 검증한 결과를 살펴본다. 먼저 〈표 7〉에서는 나타나지 않지만 일반선형회귀의 분석결과에서는 등급(평점)이 기부금에 미치는 영향이 개인기부금의 경우에는 계수가 0.787((p) |T|: 0.02) 기업(단체)기부금의 경우에는 1.146((p) |T|: 0.0001)로 나타났다. 이에 비해 RDD 모형의 경우에는 개인기부가 임계점 90에서 별 2개 등급이 별 3개로 상승하면 2.3배의 기부금 증가(jump)를 가져온다. 단체기부도 임계점 90에서 등급의 처리효과 유의하게 상당히 긍정적으로 나타났다. 다만 처리효과 크기는 개인기부가 단체기부에 비해 2배 정도 강하게 나타났다. 이와 같이 예상과는 달리 개인기부금에 있어서 등급의 처리효과가 상대적으로 크게 나타난 것은 본 연구대상 NPO의 경우 개인들이 자신들이 기부할 단체에 대



주) 여기서 3 star는 임계값 90점 전후의 평점이며 2 star는 임계값 80점 전후의 데이터를 의미함  
 각 점은 0.2점 간격의 평점을 갖는 모든 관측치(NPOs)의 평균을 나타낸다.  
 실선은 컷오프 평점(각각 80점과 90점)의 양쪽에 있는 실제 관측치에 의거한 3차 다항식이다.

〈그림 3〉 NPO 규모별 등급(star ratings)의 차이로 인한 처리 효과

한 정보 부족을 평가기관이 제공한 등급(평점)으로 보완했다는 것을 알 수 있다. 이에 비해 기업과 단체 기부자들은 등급(평점) 보다 자신들의 전문 능력과 자원을 활용하여 기부 결정에 활용한 것으로 추정할 수 있다.

마지막으로 NPO에 관한 감시기관의 등급 부여와

인터넷 검색을 통한 미디어 노출 및 웹 사이트 이용이 기부금 수입에 미치는 상대적 영향력이 상이할 것이라는 가설 4를 검증한 결과가 〈표 8〉에 제시되어 있다. 분석결과를 보면 NPO에 대한 제3자 등급 정보와 언론보도, 웹 사이트 공시는 모두 기부금 크기에 긍정적 영향을 미치는 것으로 확인되었다. 특

〈표 7〉 개인기부와 단체기부의 평점 처리효과 비교

모형 \ 규모	개인 (3 star)	개인 (2 star)	단체 (3 star)	단체 (2 star)
RDD회귀(모수)	2.385(0.573)	-0.345(0.763)	1.135(0.424)	-0.808(0.701)
다중선형회귀	2.061(0.722)	-0.486(0.682)	1.235(0.453)	-0.744(0.621)
RDD(통제변수)	2.402(0.699)	-0.803(0.701)	1.225(0.466)	-0.803(0.717)

주) RDD(통제변수)는 기본 RDD 모형에 통제변수와 더미변수를 포함한 것이다.  
 여기서 3 star는 임계값 90점 전후의 평점이며 2 star는 임계값 80점 전후의 데이터를 의미함  
 검증 결과 값의 원편은 회귀계수 오른편은 표준오차를 나타낸다.

〈표 8〉 기부금 수입의 영향요인에 관한 상대적 유용성

설명변수 \ 모형	모형 1	모형 2	모형 3	모형 4
Rating	0.542(0.150)***			0.441(0.143)**
Website		0.528(0.061)***		0.469(0.061)***
COVG			0.17(0.04)***	0.101(0.032)**
AGE	0.010(0.009)*	0.010(0.009)*	0.011(0.010)	0.009(0.011)
ASSETS	0.342(0.049)***	0.221(0.041)***	0.32(0.05)***	0.201(0.052)***
ADMIN	-3.313(0.492)***	-3.642(0.391)***	-3.601(0.433)***	-2.991(0.420)***
PRICE	0.007(0.008)	0.005(0.009)	0.008(0.010)	0.009(0.011)
IND	Yes	Yes	Yes	Yes
YEAR	Yes	Yes	Yes	Yes
N	312	308	312	308
Adj. R <sup>2</sup>	0.405	0.502	0.416	0.526

주) '유의수준: '\*\*\*' 0.01 '\*\*' 0.05 '\*' 0.1' 그리고 ( ) 내는 표준오차

히 등급 정보와 웹 사이트 공시 정보가 언론보도에 비해 상대적으로 기부금 크기와 강한 관련성을 갖는다는 것을 알 수 있다.

#### 4.3 추가분석 결과

제3차 등급 평가가 기부금 수입에 미치는 영향을 분석한 모형에 대한 강건성을 확보하기 위해 몇 가지 추가분석을 실시하였다. 먼저 데이터의 이상치에

의한 결과의 변화 여부를 조사한다. 기부금과 평점에 대해 표본의 상위 1%와 5% 또는 하위 1%와 5% 내에 있는 NPO를 제외하고 설정한 모형을 다시 추정한 결과 기부금에 미치는 별 등급의 효과는 보다 강화되었다. 대표적으로 〈표 2〉의 모형3에서 상하 5%의 이상치를 제거했을 때 평점이 기부금에 미치는 회귀계수가 0.101(p > T=0.001)로 높아졌으며 RDD 분석 모형에서도 임계값 90점 전후의 별 등급의 변화가 기부금에 미치는 효과도 강화되었다.

다만 계수의 부호가 바뀌거나 비유의적인 결과가 유의하게 변경된 경우는 나타나지 않았다.

그리고 연구결과의 신뢰성을 확보하기 위해 기부금에 미치는 별 등급의 효과를 추정함에 있어서 전기의 통제변수 이외에 당기의 통제변수에 의해서도 차별적인 영향을 받는지를 추가 분석하였다. 추정 결과는 당기의 통제변수 세트를 사용한 경우에도 기부금에 미치는 효과가 유사하게 나타났다.<sup>17)</sup>

끝으로 매년 연속적으로 별 3개 등급을 받은 NPO가 그렇지 않은 NPO에 비해 기부금 수입이 증대되었는지에 관한 추가분석을 실시하였다. 3년 연속 별 3개 등급을 받으면 3점, 2년 연속이면 2점, 한 번도 연속해서 별 3개를 받지 않으면 0점으로 환산했다. 대상 관측수를 보면 2018년 기준 샘플 수 105개 NPO중에서 3년 연속해서 별 3개를 받은 NPO는 15개이며 한 번도 받지 못한 NPO는 42개 이었다.<sup>18)</sup> 회귀분석 결과는 계수 0.708(t value=4.23)으로 나타나 연속적으로 별 3개 등급을 받은 단체일수록 기부금 수입에 매우 강한 양의 영향을 미친다는 것을 보여주고 있다.

## V. 결론

본 연구에서는 NPO(공익법인) 감시기관이 부여한 제3자 등급(별점 평가)을 조사하여, 그러한 등급의 증가 또는 감소가 NPO의 기부금 수입에 독립적인 영향을 미치는가를 회귀불연속설계(Regression

Discontinuity Design: RDD) 접근법을 사용하여 분석하였다.

분석결과를 요약하면 다음과 같다. 첫째 NPO에 대한 제3자 등급(평점)이 기부금 의사결정에 유용한 정보가 된다는 가설 1은 모든 등급의 증가가 기부금에 우호적인 영향을 미친 것이 아니라, 일부의 범위(별 등급 2개에서 3개로 증가)에서만 지지되었다. 둘째 규모에 따른 등급과 기부금 크기의 관계는, 대규모 NPO의 경우에 별 등급 2개에서 3개로 별 등급의 상승이 임계점(90점)에서 유의한 기부금 수입의 증가(jump)가 나타난 것을 제외하고는 소규모의 경우에는 등급의 증가가 모두 임계점에서 기부금의 증가(jump) 즉 회귀불연속이 발생하지 않았다. 셋째 개인기부자와 기업기부자의 등급에 대한 반응이 상이할 것이라는 가설 3은 예상과는 달리 개인기부자 임계점 90에서 등급의 처리효과가 기업 및 단체기부에 비해 두 배 이상 강하게 나타났다. 넷째 가설 4에서는 NPO에 대한 제3자 등급 정보와 웹사이트 공시 정보가 언론보도에 비해 상대적으로 기부금 크기와 강한 관련성을 갖는다는 것이 확인되었다. 또한 매년 연속적으로 별 3개 등급을 받은 NPO가 그렇지 않은 NPO에 비해 기부금 수입에 매우 강한 양의 영향을 미친다는 것을 보여주었다.

본 연구의 시사점은 첫째 전반적으로 제3자 등급의 증가가 NPO의 기부금 수입에 양의 영향을 미치는 것은 분명하지만, 무엇보다 실제 평점의 절단점(cutoff)에서 변화하게 되는 별 등급 정보가 기부 의사결정에 핵심적 역할을 담당한다는 것이다. 다만 이러한 별 등급의 상승효과는 상대적으로 높은 별

17) 예를 들어 <표 3>의 모형3을 적용한 결과 평점(SCORE)의 계수가 0.058(표준오차 0.016)으로 유사하게 나타났고 통제변수 4가지도 모두 계수의 크기와 부호가 통계적으로 상이하지 않았다.

18) 2017년 사업년도 기준 2018년 발표된 별3점 NPO(공익법인)은 143개 이었으나 대상 분석기간 4년간 재무정보와 기타 조직 특성 정보가 누락되지 않은 법인은 105개(312개 년 법인) 이다.

등급의 변화 즉 별 등급 2개에서 3개로 상승하는 경우에 한정된다. 둘째 향후 실제 근소한 평점 차이로 별 등급 3개를 받지 못하는 NPO들의 불만을 해소하기 위해서는 최근 신용평가 제도가 등급에서 점수제로 전환되는 것처럼 별 등급을 보완할 수 있는 다양한 방안이 강구되어야 한다는 것이다.

본 연구의 결과는 기부금 시장에서 등급 정보의 유용성을 실증한 것에 의의가 있으며 기부자들이 상대적으로 대규모 NPO의 등급 품질에 관한 신호에 긍정적으로 반응한다는 발견은 NPO 관리자들의 발전방향에 중요한 전략적 함의를 갖는다. 나아가 제3자 등급을 통하여 NPO의 품질과 기부자와의 관계를 강화하는 것은 사회에 긍정적인 이익을 가져다준다는 근거를 마련했다는 점이다. 또한 본 연구의 결과는 객관적으로 제공되는 NPO의 품질에 대한 등급 정보는 종국적으로 기부 제도에 관한 신뢰성과 투명성 및 효율성을 개선하는데 유용한 정보로서 활용될 수 있다. 즉 기부자들을 더 높은 품질의 NPO로 안내할 수 있고 또한 NPO 스스로 기부자의 수요 변화와 요구사항에 대응해 부실한 단체는 폐업하거나 자발적으로 품질을 개선하도록 유도할 수 있게 된다.

이상의 연구 공헌에도 불구하고 본 연구는 몇 가지 한계가 있다. 첫째 무엇보다 공익법인 중에서 제3자 등급을 받은 단체의 수가 많지 않아 결과를 일반화하는 것이 쉽지 않다. 둘째 NPO의 정보 공시와 독립된 등급의 관계를 이해하고 제3자 등급에 대한 NPO의 대응을 조사하는 것이 필요하다. 셋째 제3자 등급과 NPO의 다른 특성(공익사업 유형, 지역 특성, 재단 또는 사단법인, 회계감사 여부 등)을 함께 비교할 수 있는 분석모형 개발이 향후의 연구과제이다.

## 참고문헌

- 우석진(2016), "학업성취도 미달 학생이 보통 이상 학생의 학업성취도에 미치는 친구효과 추정," **노동경제논집**, 39(3), 75-97.
- 이동성·문태훈(2016), "에너지효율등급이 아파트 가격에 미치는 영향에 관한 연구," **환경정책**, 24(4), pp. 25-42.
- Akerlof, G. A.(1970), "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*. The MIT Press. 84(3), pp.488-500.
- Besiou, M. and J. McCormick(2013), "The Impact of Media on Donations." 3rd International HumLogWorkshop, Essen.
- Blouin, M. and R. Lee(2015), "A signaling theory approach to nonprofit web disclosure," *Journal of Academy of Business and Economics*, 15(1), pp.7 - 20.
- Brown, A., J. Meer and J. Williams(2017), "Social distance and quality ratings in charity choice," *Journal of Behavioral and Experimental Economics*, 66, pp.9-15
- Chhaochharia, V. and S. Ghosh(2008), "Do Charity Ratings Matter?" Unpublished Manuscript, Florida Atlantic University.
- Cnaan, R. A., Jones, K., Dickin, A. & M. Salomon (2011), "Nonprofit watchdogs: Do they serve the average donor?" *Nonprofit Management and Leadership*, 21, pp.381-397.
- Espeland, W. and M. Sauder(2007), "Rankings and Reactivity: How Public Measures Recreate Social Worlds," *American Journal of Sociology*, 113(1), pp.1-40
- Finch, J.(2014), "Survey: Do Ratings From Watchdog

- Groups Impact Giving Decisions?" The Able Altruist. <http://able-altruist.softwareadvice.com/>
- Frumkin, P. and M. T. Kim(2001), "Strategic Positioning and the Financing of Nonprofit Organizations: Is Efficiency Rewarded in the Contributions Marketplace?" *Public Administration Review*, 61, pp.266 - 275.
- Gibelman, M. and S. Gelman(2001), "Very Public Scandals: Nongovernmental Organizations in Trouble," *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 12(1), pp.49-66
- Gordon, T.P., C.L. Knock, and D. Neely(2009), "The Role of Rating Agencies in the Market for Charitable Contributions: An Empirical Test," *Journal of Accounting and Public Policy*, 28, pp.469-484.
- Grant, L. E.(2010), "The Response to Third-Party Ratings: Evidence of the Effects on Charitable Contributions," Unpublished Manuscript, University of California Santa Barbara.
- Harbaugh, W. T., Mayr, U., and D. R. Burghart (2007), "Neutral Responses to Taxation and Voluntary Giving Reveals Motives for Charitable Donations." *Science*, 316, pp.1622-1625.
- Hatfield, P.(2018), "How Helpful are Charity Watchdog Ratings?" *Journal of Accounting and Finance*, 18(2), pp.71-88.
- Imbens, G.W. and T. Lemieux(2008), "Regression discontinuity designs: A guide to practice." *Journal of Econometrics*, 142(2), pp.615-635.
- Jacob, R. and P. Zhu(2012), "Practical Guide to Regression Discontinuity," MDRC working paper.
- Jen, S. and R. Croson(2009), "A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods," *Economic Journal*, 119 (540), pp.1422-1439
- Ling, Q and D. G. Neely(2013), "Charitable Ratings and Financial Reporting Quality: Evidence from the Human Service Sector," *Journal of Public Budgeting, Accounting & Financial Management*, 25(1), pp.69-90.
- Luca, M.(2011), "Reviews, reputation, and revenue: The case of Yelp.com, Harvard Business School," Working Paper No.12-016.
- O'Neill, M.(2009), "Public Confidence in Charitable Nonprofits." *Nonprofit and Voluntary Sector Quarterly*, 38(2), pp.237-269.
- McCrary, J.(2008), "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 142, pp.698 - 714.
- Meer, J.(2014), "Effects of the price of charitable giving: evidence from an online crowdfunding platform." *Journal of Economic Behavior and Organization* 103, pp.113-124 .
- Prakash, A. and M. Gugerty(2010), "Trust but verify? Voluntary regulation programs in the non-profit sector," *Regulation & Governance*, 4 (1), pp.22-47.
- Primoff, W.(2012), Fiduciary financial management in nonprofit organizations. *The CPA Journal*, pp.48-57.
- Roberts, A. A., P. Smith, and K. Taranto(2006), "Evaluating charitable spending: Marginal versus average performance measures," Working paper. Boston College, University of Texas at San Antonio and Georgetown University.
- Saxton G. D., D. G. Neely, and C. Guo(2014), "Web disclosure and the market for charitable contributions," *Journal of Accounting and*

- Public Policy*, 33(2), pp.127-144.
- Sloan, M.F.(2009), "The Effects of Nonprofit Accountability Ratings on Donor Behavior," *Nonprofit and Voluntary Sector Quarterly*, 38, pp.220 - 236.
- Singer, P.(2015), "*The Most Good You Can Do: How Effective Altruism Is Changing Ideas About Living Ethically*," Yale University Press.
- Szper, R. and A. Prakash,(2011), "Charity Watchdogs and the Limits of Information-Based Regulation," *Voluntas*, 22, pp.112-141.
- Szper, R.(2013), "Playing to the Test: Organizational Responses to Third Party Ratings," *Voluntas*, 24, pp.935-952.
- Tang, T.(2009), "Information asymmetry and firms' credit market access: Evidence from Moody's credit rating format refinement," *Journal of Financial Economics*, 93(2), pp.325-351.
- The Staff of Entrepreneur Media and Corbin Collins(2017), "*Start Your Own Nonprofit Organization: Your Step-By-Step Guide to Success*," Entrepreneur Press: 2 edition.
- Tinkelman, D.(1999), "Factors affecting the relation between donations to not-for-profit organizations and an efficiency ratio," *Research in Government and Nonprofit Accounting*, 10, pp.135-161.
- Yörük, B. K.(2012), "The effect of media on charitable giving and volunteering: Evidence from the "Give Five" campaign," *Journal of Policy Analysis and Management*, 31(4), pp.813-836.
- Yörük, B. K.(2016), "Charity ratings" *Journal of Economics & Management Strategy* 25 (1), pp.195-219.

## The Effect of Third-Party Ratings on the Change in the Charitable Giving to Nonprofit Sector

Keun-Hyo Yook\*

### Abstract

In this study, we investigate the role of third-party ratings (star ratings) given by watchdog organizations and empirically examines whether the zero- to three star ratings provided by GuideStar KOREA have additional information content for donors using the Regression Discontinuity Design (RDD) approach. To test the relationship between charity ratings and donations, we examine 315 non-profits rated by GuideStar in Korea for the period 2015-2018.

The results are as follows: First, the overall ratings have a statistically significant effect on the amount of contributions receives. Both star ratings and rating score(overall ratings) have a significant positive impact on the contributions, especially star ratings have shown a strong correlation with donations rather than rating score. That's because stars are relatively easy to interpret than overall ratings. Second, before doing RDD analysis, we also verified that assumptions for applying RDD are properly established in our model. The results show that there is no evidence of a significant change in control variables at the threshold values of the overall ratings, the forcing variable (a function of overall rating) is not manipulated, and McCrary's density test yields a nonrandom sorting of charities across cutoff ratings. This finding implies that the assumption of applying RDD is found to be satisfied. Third, for NPOs that are rated as 3 stars, a one point increase in overall ratings is associated with a 1.4 times increase in charitable contributions received, while for NPOs that are rated as 2 and 3 stars, the estimation results show that an increase in overall ratings does not have a significant impact on charitable contributions received. Fourth, this study shows that for large NPOs that are slightly below the 3 star threshold rating, a one star increase has a significant impact on

---

\* Professor, College of Business, Busan University of Foreign Studies, First Author

the amount of contributions that the NPO would receive. However, large NPOs that are slightly below the 2 star threshold rating and small NPOs that are below 2, 3 star threshold rating, a one star increase has no significant impact on the amount of contributions an organization receives. Fifth, we find that the responses to the ratings of individual donors and corporate donors are all significant positive, but contrary to expectations, the treatment effect of the ratings at 90 cut-points is more than twice as strong as that of corporate donations. Lastly, the study indicate that third-party rating information and Web site disclosure information on NPOs are relatively relevant to the contributions received compared to media coverage. It also show that NPOs received the first grade for the third consecutive year have higher contributions than those who do not.

Overall, we find that a change in GuideStar ratings is associated with a directionally consistent change in donations. Above all, the study suggests that star rating informations at the cut-point (assignment point) of the actual rating, plays an important role in the decision-making for charity giving. However, the effect of these star ratings is limited to changes in the relatively high star ratings, from two star ratings to three. Therefore, in order to address complaints from NPOs that do not receive three star ratings due to the difference of actual marginal rating, various measures should be worked out to supplement star ratings system, just as the latest credit rating system shifts from star rating to scoring system. The results of this study will be useful information to improve the rating system and activation of donations by NPO CEOs and the government.

Key words: Donation, Regression Discontinuity Design (RDD), NPO (non-profit organization), star ratings, rating score

---

• 저자 육근효는 현재 부산외국어대학교 경영학부 회계 전공 교수로 재직 중이다. 부산수산대학교와 부산대학교에서 경제학석사와 경영학석사, 동아대학교에서 경영학박사를 취득하였다. 일본 고베대학, 요코하마국립대학, 오사카부립대학, Akron대학에서 초빙교수 및 객원교수로 활동하였다. 주요 연구분야는 원가계획, 가치분석, CSR회계 등이다.