

The Decay in Contributions in a Public Goods Game: Learning Hypothesis, Strategy Hypothesis and Reciprocity Hypothesis Revisited*

Jung-Kyoo Choi[†] Junsok Huhh[‡]

Abstract Most studies on public goods game reported that in a finitely repeated public good game subjects' contributions begin with an average contribution of about 50% of their initial endowment and decay toward the free riding level as the game progresses. Unconditional free ridings were seldom observed. To explain what causes this behavioral patterns, three hypotheses have been suggested; learning hypothesis, strategy hypothesis and reciprocity hypothesis. This paper investigates these hypotheses, and our findings from public goods experiments suggest that subjects' reciprocity is one of the main causes that produce subjects' behavioral patterns.

Keywords Public Goods Game, Learning Hypothesis, Strategy Hypothesis, Reciprocity Hypothesis.

JEL Classification D03, C91

*This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2008-332-B00038).

[†]School of Economics and Trade, Kyungpook National University

[‡]Seoul National University

공공재 게임 실험에서 기여율의 하락: 학습 가설, 전략 가설, 상호적 맞대응 가설의 재평가*

최정규[†] 허준석[‡]

Abstract 공공재 게임을 유한한 회수로 반복할 때, 대부분의 실험 연구가 보여주는 바는 실험 참가자들이 초반 50% 대의 기여율에서 출발하여 게임이 반복됨에 따라 그 기여율이 점차 하락한다는 것이다. 실험 참가자들의 무조건적인 무임승차는 잘 관찰되지 않는다. 이로부터 실험 참가자들이 왜 이런 행동패턴을 보이는지, 그리고 이러한 행동패턴은 경제이론의 출발점이 되는 ‘경제인’이라는 가정으로부터 이탈하는 것인지 등을 둘러싸고 많은 논의가 이루어져왔다. 본 연구에서는 이러한 논의의 연장선상에서 이러한 행동을 설명하기 위해 제시된 주요한 몇몇 가설들을 재평가하고자 한다. 본 논문에서는 공공재 게임 실험을 통해서 얻어진 결과를 기초로, 실험 참가자들이 상호성을 갖고 행동한다는 것, 그리고 행위자들의 상호적 맞대응이 공공재 실험에서 보여주는 일반적인 패턴을 설명하는 데 유효한 요인이 된다는 것을 보일 것이다.

Keywords 공공재 게임, 학습 가설, 전략 가설, 상호적 맞대응 가설

JEL Classification D03, C91

*이 논문은 2008년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2008-332-B00038).

[†]경북대학교 경제통상학부

[‡]서울대학교 정보문화학 연합전공

1. 도입

공공재 게임 실험에서 유한한 회수로 게임을 반복할 때, 대부분의 연구가 보여주는 바는 이 게임에서 우월전략인 무임승차 행동이 이론의 예측만큼 관찰되지 않는다는 점이다. Journal of Public Economic에 마웰과 에임스의 선구적인 실험 결과가 보고된 이후(Marwell and Ames, 1981), 많은 연구자들이 공공재 게임을 실험한 결과를 보고해왔는데, 이들 연구들이 공통적으로 보여주는 결과는 다음과 같다(Marwell and Ames, 1981; Isaac, Walker and Thomas, 1984; Isaac, McCue and Plott, 1985, Isaac and Walker, 1988; 그리고 전반적인 리뷰로는 Keser, 2002; Ledyard, 1995; Andreoni and Croson, 1998).

- (1) 공공재 게임을 동일한 참가자들을 상대로 수차례 반복 시행할 때, 실험 참가자들은 평균적으로 50% 정도를 공공계정에 기여하면서 게임을 시작한다.
- (2) 이 기여율은 게임이 반복되면서 점점 하락한다.
- (3) 마지막 회에서 조차 기여율은 0으로 떨어지지 않으며, 완전한 무임승차는 잘 관찰되지 않는다.
- (4) 마지막 회에서의 기여율이 0 이상이지만, 마지막 회에는 이전 회에 비해 기여율이 크게 하락한다.

이로부터 실험 참가자들이 왜 이런 행동패턴을 보이는지, 그리고 이러한 행동패턴은 경제이론의 출발점이 되는 ‘경제인’(Homo economicus)이라는 가정으로부터 이탈하는 것인지 등을 둘러싸고 많은 논의가 이루어져왔다. 위의 결과를 둘러싸고 이를 설명하기 위해 몇 가지 가설들이 제시되었는데 그 중 주요한 몇 개의 가설들은 다음과 같다.

첫째, 공공재 게임에서 초기에 50%에 달하는 기여율이 나오는 것, 그리고 마지막 회에도 0 이상의 기여율이 관찰되는 것 등은 실험 참가자들이 게임의 구조를 정확히 이해하지 못해서 어떻게 행동하는 것이 자신들에게 가장 이익이 되는지를 깨닫지 못했기 때문에 얻어지는 결과일 수도 있다(이를 ‘학습가설’이라고 부른다). 이렇게 본다면 실험을 통해 관찰되는 패턴, 즉 반복에 따른 기여수준의 점차적 하락이라는 현상은 실험 참가자들이 게임을 반복적으로 수행하면서 점점 게임의 구조를 이해할 수 있게 되고 그 과정에서 어떻게 행동하는 것이 자신들에게 가장 유리한지를 깨닫는 과정일 뿐이게 된다. 이 가설에 따르면, 이러한 학습이 충분히 진행된 후라면 이

론이 예측하는 바대로 경기자들이 우월전략인 무임승차전략을 채택하는 모습을 관찰할 수 있을 것이다. 다른 말로 하자면 게임의 초반에 나타나는 양(+)의 기여율은 실험 참가자들이 게임을 제대로 이해하지 못한 상태에서 나타나는 일시적인 현상이며, 실험 참가자들이 게임의 구조를 정확히 이해하게 되면 이론에서 예측하는 0의 기여율이 관찰될 것이라는 것이다. 따라서 게임을 반복하면서 기여율이 점차 낮아지는 것은 학습이 진행되고 있음을 보여주는 증거가 된다는 것이다(Selten and Stoecker, 1986; Andreoni, 1995; Palfrey and Prisbrey, 1996; Houser and Kurzban, 2002).

둘째, 실험 참가자들이 게임의 구조를 정확히 이해하고 있고, 무임승차가 자신에게 가장 유리한 전략임을 알고 있더라도, 이 사실이 공통지식으로 성립하지 않는다면 그 우세전략에 따라 행동하지 않고 게임의 초반에 양(+)의 기여를 하는 것이 전략적인 선택일 수도 있다는 것이다(이를 ‘전략가설’이라고 부른다). 만일 자신은 게임을 잘 이해하고 있더라도 상대방 경기자들이 게임을 제대로 이해하고 있지 않다고 믿는다면, 첫 회부터 무임승차를 해서 상대방들로 하여금 무임승차가 유리한 전략임을 곧바로 깨닫게 하는 것보다는 게임의 초반부에는 어느 정도 높은 기여 수준을 유지함으로써 상대방으로부터 높은 기여를 유도하는 것이 전략적으로 유리할 수 있다는 것이다(Kreps et al, 1982). 하지만 경기가 종료시점에 다가가면서 이러한 전략적 행동의 유효성은 점차 떨어질 것이고 특히 마지막 회에서는 이후 게임이 없기 때문에 굳이 상대방의 기여를 유도하기 위해 자신도 높은 기여를 할 필요가 없다. 위 (1), (2), (4)는 참가자들이 이렇게 전략적으로 행동하기 때문에 나타나는 결과이다(Croson, 1996; Muller et al, 2008; Sonnemans et al., 1999).

셋째, 실험 참가자들은 게임의 구조를 정확히 이해하고 있고, 공공계정에의 기여가 자신에게 유리한 선택이 아님을 잘 알고 있지만, 실험 참가자들은 자신의 몫을 어느 정도는 희생하더라도 타인(혹은 자신이 속한 집단)의 이득을 높이는 방향으로 행동하려는 성향이 있기 때문에, 즉 실험 참가자들이 사회적 선호를 갖고 있기 때문에 양의 기여율이 관찰되는 것일 수 있다(Bowles, 1998; Sobel, 2005; Rabin, 1998; Levine, 1998). 하지만 타인의 이득을 긍정적으로 고려하려는 성향은 언제나 무조건적으로 긍정적인 방향으로 나타나지는 않는다. 상대방도 자신과 마찬가지로 타인의 이득을 긍정적으로 고려하려는 성향을 가졌다고 판단할 때는 그러한 방향으로 계속 행동하지만, 상대방이 그렇지 않아 자신의 이득만 챙기는(즉 무임승차하려는) 성향을 가졌다는 것이 판단될 때에는 언제든지 상대방에 대한 배려(즉 공공계정에 대한 양의 기여)를 철회함으로써 상대방의 무임승차에 보

복할 수 있다. 즉 실험 참가자들이 상호성을 갖고 행동한다면 상대방들의 높은 기여수준에 대해서는 자신도 기여 수준을 높여 대응할 것이지만, 상대방들이 낮은 기여수준을 보인다면 자신도 기여 수준을 낮추는 식으로 행동한다는 것이다. 따라서 위 (1), (2), (3)은 실험 참가자들이 한편으로는 타인의 이득에 대해 긍정적으로 고려하고 있음을 반영하며, 상대방의 무임승차에 대해서는 자신도 기여를 철회함으로써 보복 혹은 맞대응한다는 상호적 태도에서 비롯되는 것이다(이를 ‘상호적 맞대응 가설’ 혹은 ‘사회적 선호 가설’이라고 부른다). (Andreoni, 1995; Croson et al., 2005; Carpenter et al, 2009; Fehr and Gächter, 2000; Fehr and Fischbacher, 2002)

이상의 세 가지 가설의 타당성을 둘러싸고 많은 연구들이 진행되어 왔다. 본 연구에서는 이를 둘러싼 기존 연구들에 대한 검토를 수행하고(2절), 각 가설의 타당성을 검증하고자 우리가 진행한 실험을 소개한 후(3절), 우리의 실험 결과를 토대로 기존 가설에 대한 몇 가지 추가적인 평가를 내리도록 할 것이다(4절).

2. 기존 문헌 검토

학습 가설의 타당성을 검증하기 위해 여러 연구들이 진행되어 왔지만 그 중에서도 대표적인 것이 안드레오니(Andreoni, 1988)의 실험이다. 안드레오니는 실험 참가자들에게 공공재 게임을 몇 회 반복할 것인지를 미리 알려주고 게임을 진행한 후, 동일한 참가자들을 상대로 예고되지 않은 실험을 한 차례 더 진행했다. 만일 기여율 수준의 하락이 실험 참가자의 학습의 진행과정을 반영하는 것이라면, 이러한 갑작스런 실험 재개(sudden start) 후에 관찰되는 기여율 수준은 재개 이전(즉 이전 게임의 종료 시점)의 기여율과 연속성 상에 있어야 한다. 하지만 안드레오니는 최초 반복을 통해 점차 하락하던 기여수준은 갑작스런 실험 재개와 더불어 다시 50% 수준으로(즉 게임의 첫 회 때 보여준 기여율 수준으로) 상승한다는 것을 발견했다. 갑작스런 실험 재개와 동시에 기여율이 초반 수준으로 복귀한다는 것은 매 회 기여율 수준의 하락이 실험 참가자들의 학습을 반영하는 것은 아님을 보여주는 강력한 반증이 되었다. 다른 한편, 아이작, 워커 그리고 윌리엄스(Isaac, Walker and Williams, 1994)는 실험 참가자들에게 공공재 게임을 각각 10회, 40회, 60회로 실시했는데, 기여율이 하락 속도가 실험의 반복 횟수에 따라 다르게 나타남을 발견했다. 즉 10회, 40회, 60회 반복 게임들에서 보여주는 마지막 회의 평균 기여율 수준은 거의 유사했는데(약 10% 내외) 흥미롭게도 이 10%대의 기여율에 도달하는 시간이 각각 10

회, 40회, 60회가 걸렸다. 만일 기여율의 하락이 학습 과정을 반영하는 것이라면 10회 반복 게임에서 10회 때 10%의 기여율을 보였다면, 40회, 60회 반복 게임에서도 10회 쯤 10%의 기여율에 도달한 후 그 기여율이 나머지 회 동안 지속되어야 하는데 그렇지 않았다는 것이다. 학습과정이 반복회수에 의존해서 느리게 혹은 빠르게 일어날 리가 없으므로, 이러한 결과 역시 학습 가설을 반박해주는 강력한 반증으로 받아들여졌다.

전략가설을 테스트하기 위한 실험도 진행되었다. 안드레오니(Andreoni, 1988)는 유한 반복 공공재 게임을 (1) 집단의 구성원이 바뀌지 않은 채 게임을 유한한 회수로 반복하는 조건(이를 파트너 조건이라 부른다)과 (2) 매 회 집단 구성원을 바꾸면서 게임을 유한한 회수로 반복하는 조건(이를 스트레인저 조건이라 부른다)으로 나누어 실시했다. 만일 공공재 게임에서의 양(+)의 기여율이 실험 참가자들의 전략적인 고려에서 기인한 것이라면 매 회 다른 구성원과 게임을 벌이기 때문에 전략적 고려가 존재할 여지가 없는 스트레인저 조건에서는 0의 기여율을 나타내야 하지만, 안드레오니가 발견한 것은 스트레인저 조건에서의 기여율이 파트너 조건에서의 기여율보다 오히려 높다는 것이었다. 물론 이러한 결과가 이후 연구에서 재현되지 않았고 따라서 전략가설을 기각하기에는 충분치 않다는 반박도 있었다(Weimann, 1994; Keser and van Winden, 1996). 또한 소네만 등(Sonnemans et al., 1999)의 연구에서는 스트레인저 조건을 부분적으로 적용한 경우(즉 모든 사람이 집단을 바꾸는 것이 아니라 4명 중 한 사람씩만 차례로 집단을 바꾸게끔 세팅한 경우) 이 집단을 떠나게 된 사람들의 기여 수준은 떠나는 시점에 가까워갈수록 (남아 있는 사람들에 비해) 현저히 낮아진다는 사실을 발견했다. 또한 이들은 집단에 머무는 기간을 3, 6, 9, 12회로 다르게 통제해보았는데, 전체 길이와 상관없이 떠나는 시점이 얼마 남았는가에 따라 기여율 하락 속도가 거의 유사하게 됨을 발견할 수 있었다. 이 결과는 실험 참가자들의 의사결정에 전략적인 고려가 크게 작용한다는 것을 입증하고 있다.

만일 학습가설과 전략가설이 공공재 게임에서의 양의 기여율의 존재와 반복에 따른 기여율의 하락을 잘 설명해내지 않는다면 남아 있는 가능성은 사회적 선호에 기반한 상호성 가설¹이다. 이 가설에 따르면 실험 참

¹ 기존 문헌들에서 전략 가설을 이야기할 때 상호성을 포함하는 경우들도 많다. 상대방 경기자들의 이전 회의 행동에 현재 나의 행동을 조건짓는 것은 상호성의 정의이기도 하지만, 이 역시 전략의 일종으로도 해석할 수 있기 때문이다. 하지만 이 글에서는 양자를 엄밀하게 구분하고자 한다. 즉 Keser의 구분에 따라 전략가설이라 할 때에는 이후 상대방의 기여를 유도하기 위해 현재 양(+)의 기여를 하는 경우, 즉 행위자들이 이와 같이 미래지향적(forward-looking)으로 행동할 것이라고 보는 입장을 전략가설이라고 하고, 경제주체가

가자들의 기여수준은 같은 집단 구성원들이 과거에 자신보다 평균적으로 많이 기여했으면(혹은 상대방이 자신보다 많이 기여할 것이라고 예상한다면) 높아지는 경향을 갖고, 자신의 기여보다 평균적으로 적게 기여했으면(혹은 적게 기여할 것이라고 예상한다면) 낮아지는 경향을 갖는다고 본다(Figuieres et al., 2011; Fischbacher and Gächter, 2010). 경제주체는 어느 정도 상대방을 위해서 자신의 몫을 희생할 준비가 되어 있지만 오직 상대방도 같은 태도로 나올 때에만 그 성향을 유지할 것이고, 그렇지 않다면 자신의 기여를 철회하고자 한다는 것이 상호성 가설의 주요 내용이다. 상호성 가설에 따를 때, 다른 가설들의 예측과 달리 공공재 게임에서 반드시 기여율의 하락이 관찰되어야만 하는 것은 아니다. 하지만 만일 상호적 성향을 갖는 실험 참가자들 사이에 무임승차자들이 끼어 있다면, 이들의 무임승차 때문에 상호적 태도를 갖는 실험 참가자들의 기여율이 하락하는 경향을 가질 것이다(Fischbacher and Gächter, 2010). 혹은 모든 실험 참가자들이 상호적 성향을 갖더라도 자신의 이해를 조금 더 고려하려는 태도를 보인다면(즉 자기 이익에 편이된 상호성이 존재한다면) 마찬가지로 기여율의 하락을 설명할 수 있다(Figuieres et al., 2011). 하지만 하우스저와 커즈반(Houser and Kurzban, 2002)은 일부 실험 참가자들에 대해서 사람이 아니라 사전 프로그램된 컴퓨터와 같은 집단을 이루어 게임을 하게 하고(물론 이 사실을 실험 참가자에게 알려주고) 반복 공공재 게임을 실시해본 결과, 여전히 이들에게서 양(+)의 기여가 나타나며 이 기여율이 반복에 따라 하락함을 발견했다. 이 세팅에서는 컴퓨터가 사전 프로그램되어 있으므로 참가자들은 현재 자신의 기여수준에 따라 컴퓨터가 반응하지 않을 것임을 알기에 전략적인 고려가 들어설 여지가 없고, 더 나아가 상대가 컴퓨터라는 사실을 알고 있기 때문에 어떠한 이타성이나 상호성도 들어설 여지가 없으므로, 여기서 나타나는 모든 양의 기여율은 실험 참가자들이 게임의 구조를 잘 이해하지 못했기 때문인 것으로 해석할 수 있다. 이를 기반으로 하우스저와 커즈반은 통상적인 반복 공공재 게임에서 나타나는 상호성이 기인한다고 여겨지는 기여율 중 약 50% 가량이 실험 참가자들의 게임 구조에 대한 이해부족에서 기인하는 것이라고 주장하였다.

이처럼 기존 연구 결과는 어느 하나의 가설로 합의되고 있지 않다. 본 논문에서도 하나의 가설을 배타적으로 지지하는 결론을 유도하지는 않을 것이다. 기존 연구 결과도 그렇고 앞으로 볼 본 연구의 결과도 학습, 전략,

상대방의 과거 행동에 대해 현재 행동을 조건짓는 방향으로, 상대방의 호의에는 호의로 그리고 배신에는 배신으로 맞대응한다고 보는 입장을 상호성 가설로 구분하고자 한다(Keser, 2000).

상호성 세 가지 요인이 복합적으로 작용함으로써 공공재 게임에서의 패턴을 만들어낸다는 것을 보여주고 있다. 본 연구에서는 기존 연구에서 확인되지 않았던 요인들 몇 가지를 추가적으로 확인함으로써 공공재 게임에서의 결과를 둘러싼 기존 논의를 한 걸음 진척시키고자 한다.

3. 실험의 설계

3.1. 공공재 게임

실험은 다음과 같은 보수 구조를 갖는 공공재 게임으로 진행되었다. 참가자들은 4인 1조로 구성된 조의 일원이 되었고, 매 회 20개의 토큰을 부존자원으로 받았다. 매 회 참가자들에게 지급받은 20개의 토큰 중 몇 개를 공공계정에 기여할 것인지를 의사결정하도록 했다. 20개의 토큰 중 공공계정에 기여한 토큰 수를 빼 나머지는 개인계정에 보유되는데 개인계정에 보유한 토큰 하나 당 1점씩이, 그리고 공공계정에 보인 토큰 하나당 0.5점씩이 부여되었다. 요약하자면 집단 G 에 속한 개인 i 의 공공계정의 기여토큰 수를 x_i 라고 할 때, 개인 i 가 매 회 얻게 되는 보수는 다음과 같이 결정된다.

$$\pi_i = 20 - x_i + 0.5 \sum_{j \in G} x_j$$

공공재 게임은 다음과 같은 특징을 갖는다. (1) 공공계정에 기여를 하지 않는 것이 개인적으로 가장 유리하다: 토큰 하나를 공공계정에 기여할 때 자신에게 돌아오는 몫은 0.5인 반면 개인계정에 보유하면 1을 얻게 되기 때문이다. $d\pi_i/dx_i < 0$ 이므로, 이 게임에서는 상대방의 기여수준에 관계없이 항상 0의 기여를 하는 것, 즉 무임승차하는 것이 우월전략이다. (2) 토큰을 공공계정에 기여하면 집단 전체의 몫이 커진다. 토큰을 개인계정에 보유하면 1점이 생기는 반면, 공공계정에 보유하면 4명의 조원 모두에게 0.5점씩, 총 2점이 생긴다. 모든 실험 참가자들이 가진 토큰 전체를 공공계정에 모두 기여할 경우 전체적으로 총 120점이 생기고, 각 개인은 30점씩을 얻게 되지만, 모든 사람들이 자신의 토큰을 모두 개인계정에 보유하면 총 80점이 생기고 각 개인은 20점씩을 얻는다.

즉 공공재 게임에서 개인들의 의사결정을 관찰하면, 개인들이 어느 정도로 자신의 이익을 극대화하는 방향으로 행동하는지, 또 어느 정도로 자신의 이익을 희생하면서 타인의 이익을 증대시키는 방향으로 행동하고자 하는지를 살펴볼 수 있다.

3.2. 실험 진행

본 실험은 2010년 5월과 6월 사이에 경북대학교에서 실시되었다. 학교 홈페이지 알림 게시판을 통해 실험 참가자들을 모집하였고, 총 232명의 학생들이 참가하였다. 실험 참가자들에게는 공지 당시 약속한 5,000원의 참가 사례비에 덧붙여 게임에서 얻은 수익을 실험이 끝남과 동시에 현금으로 지불받았다. 참가 학생들에게 지불된 금액은 참가 사례비를 포함하여 평균 13,470원(최대 18,350원, 최소 10,430원)이었다.

참가자들은 실험 시작 10분전까지 약속된 장소(컴퓨터 40대가 설치된 전산실)에 모이도록 했고, 도착과 동시에 미리 프로그램이 설치된 컴퓨터 앞에 자리를 배정했다. 실험 참가자들에게는 무작위적으로 컴퓨터 자리를 배정했으며, 각자의 자리는 다른 사람의 컴퓨터 화면이 보이지 않을 정도로 충분히 거리를 두고 배치하였다. 실험이 시작되면 약 20분에 걸쳐 실험 안내를 받았다. 실험 안내는 3 페이지짜리 실험 안내문을 나누어 준 후, 실험 안내문의 내용을 큰 목소리로 읽어 내려가는 방식으로 진행되었다(실험 안내문은 요청이 있으면 제공할 수 있음). 안내가 진행되는 도중에 질문이 있으면 받았고, 안내문을 모두 읽은 후, 그리고 게임의 구조를 정확히 이해했는지를 확인하기 위해 한 페이지 가량의 연습 문제를 풀고, 답을 함께 확인하면서 게임의 내용을 정확히 숙지할 수 있도록 했다.

실험은 약 20분에서 30분가량 소요되었다. 실험은 실험 참가자들이 컴퓨터 앞에 앉아 컴퓨터 화면에 나오는 내용에 따라 의사결정 내용을 입력하고 그 결과를 확인하는 식으로 이루어졌으며, 실험 프로그램은 z-tree 라는 개발도구를 이용하여 필자들이 작성한 프로그램을 이용하였다(Fischbacher, 2007).

실험은 3.1절에서 소개된 공공재 게임으로 진행되었고, 매 회는 (1) 공공계정에 얼마의 토큰을 기여할 것인지 의사결정을 묻는 화면과 (2) 본인의 공공계정에의 기여 토큰 수, 본인이 속한 조의 공공계정에의 총 기여 토큰 수, 그리고 본인의 개인계정으로부터의 수익과 공공계정으로부터의 수익, 그리고 총 수익 등이 나타나는 결과 화면으로 구성되었다(다른 구성원들 개별적인 기여수준과 그들의 수익은 공개하지 않았다. 하지만 다른 구성원들의 평균적인 기여수준은 쉽게 계산될 수 있는데, 실험 안내를 통해서도 이것이 가능함을 확인했다).

총 10개의 세션이 진행되었고, 이 중 5개의 세션은 다음에서 설명할 S 조건으로, 그리고 나머지 5개의 세션은 P 조건으로 진행되었다. 모든 세션에서 각 참가자들은 4인 1조로 구성된 조에서 각각 공공재 게임을 진행하

였고, 총 232명 중 S 조건으로 진행된 실험에 참여한 실험 참가자 수는 116명(29개 조) 그리고 P 조건으로 진행된 실험에 참여한 실험 참가자 수는 116명(29개 조)이었다.

- S 조건: 10회를 매회 조구성을 바꾸면서 진행함 (즉 스트레인지어 조건).
- P 조건: 10회를 매회 동일한 조 편성을 유지한 채 진행함(즉 파트너 조건).

3.3. 가설 테스트

t 회째 기여수준을 C_t 라고 하자. t 회째 기여 수준 중 학습의 미비로, 즉 게임의 구조를 제대로 이해하지 못해 나타나는 부분을 x_t^L , 전략적 고려에서 기인하는 부분을 x_t^S , 그리고 마지막으로 실험 참가자들의 사회적 선호에서 기인하는 부분을 x_t^R 이라고 하면, t 회째의 기여 수준을 다음과 같이 구분할 수 있을 것이다.

$$C_t = x_t^L + x_t^S + x_t^R.$$

x_t^L : 게임이 매회 반복되면서 게임 구조에 대한 학습이 진행될 것이고, 따라서 x_t^L 의 크기는 t 가 늘어나면서 점차 감소해 종료시점에서는 0으로 수렴할 것이다.

x_t^S : P 조건에서는 게임이 종료시점에 다가오면서 전략적으로 기여를 할 유인이 점차 없어질 것이므로 전반부에서는 기여수준이 0보다 크게 나타나겠지만 종료시점에 가까워 오면서 x_t^S 는 급격히 하락하여 마지막 회에는 0이 될 것이다. 그리고 S 조건에서는 전략적 고려를 할 여지가 없으므로 x_t^S 은 0일 것이다.

x_t^R : 공공재 게임에서 어느 정도를 기여해야 하는가를 둘러싼 사회적 규범이 존재한다면, 거기서는 각 개인들이 자신의 행동을 둘러싼 의사결정에서 얼마만큼이나 타인에 대해 고려하는가와 타인들은 자신과 비슷한 상황에서 어떻게 행동하는가에 대한 각 개인들의 예측이 중요할 것이다. 즉 사회적 규범이란 집단 구성원들 간에 공유된 행동 원칙이므로, 개인들은 게임의 출발과 더불어 나름대로 사회적 규범에 대한 상을 갖고 상호작용을 시작하지만 게임이 진행됨에 따라 개인들은 타인들의 행동을 관찰하면서 자신의 사회적 규범에 대한 상을 조정해 나가는 일종의 사회적 조정 과정을 거친다. 공공재 게임에서 얼마만큼을 공공계정에 기여할 것인가의 문제는 한편으로는 각 개인들이 타인의 이득을 위해 자신의 몫을 희생하고자 하는 성향을 반영할 것이고, 다른 한편으로는 얼마만큼 기여하는 것이 사회적으로 공유된 규범인지에 대한 예측을 반영할 것이다. 특히 후자의 경우 개인들은 상대방들의 기여수준이 자신보다 높으면 자신의 기여를

높이는 방향으로 그리고 상대방들의 기여수준이 자신보다 낮으면 자신의 기여를 낮추는 방향으로 기여수준을 둘러싼 사회적 규범의 상을 조정해 나간다. 사회적 조정과정은 S 조건과 P 조건 모두에게서 진행될 것이지만, P 조건에서는 방금 언급한 기여수준을 둘러싼 규범의 사회적 조정 과정에 덧붙여 ‘상호적 맞대응’이 추가될 수 있다. 즉 상대방이 높은 기여 수준을 보일 때 이에 대해서 자신의 기여를 높이면서 호의로 답하고 상대방이 낮은 기여수준을 보일 때 이에 대해서 자신의 기여를 낮추면서 적대적으로 대응한다는 것이 상호적 맞대응의 의미이다. 처벌 혹은 포상을 위한 별도의 절차가 마련되어 있지 않은 단순한 공공재 게임의 경우 상대방의 높은 기여에 대해서 자신의 기여 수준을 높이는 것이 유일한 보답이며, 상대방의 무임승차에 대해서 보복을 할 수 있는 유일한 가능성은 자신도 기여를 철회하는 것뿐이기 때문이다. 이러한 상호적 맞대응은 매 회마다 구성원이 바뀌는 S 조건에서는 나타날 여지가 없고, 동일한 구성원과 반복적으로 게임을 벌이는 P 조건에서만 나타날 수 있다. 상호성 가설 그 자체로만 볼 때 기여율의 하락이 필연적이라는 예측은 유도되지 않는다. 다만 몇몇 연구는 상호성 가설에 따라 기여율이 하락할 것이라고 예측하기도 한다.²

첫째, S 조건에서는 전략적 고려가 작용하지 않을 것이므로(즉 $x_i^S = 0$), 회 차의 진행에 따른 기여율의 하락에 대해 다음과 같은 가설이 세워질 수 있다.

- 학습가설이 맞다면, S 조건에서 실험 초반부에는 0 이상의 기여율이 나타날 수 있지만, 게임이 반복되면서 기여율은 점차 하락하여 0으로 수렴할 것이다. 만일 S 조건에서 게임의 종료시점에도 0 이상의 기여율이 나타난다면, 그것은 게임의 구조에 대한 학습이 충분이 이루어진 이후에 나타나는 기여율이라고 볼 수 있으며, 따라서 이는 참가자들의 사회적 선호를 반영하는 것이다.

²이 때 사회적 선호의 조정에 따른 기여수준은 상승할 수도 있고 하락할 수도 있지만, 대부분의 연구에 따르면 개인들이 갖고 있는 사회적 선호를 반영하는 기여수준은 회 차를 거듭하면서 점차 하락한다. 한 집단 구성원들이 갖고 있는 사회적 선호가 서로 다르다면, 그래서 집단에 사회적 선호를 전혀 갖고 있지 않은 경기자가 존재한다면, 그의 존재는 회 차를 거듭함에 따라 사회적 선호를 갖고 있는 다른 구성원들의 기여수준을 하락시키는 요인이 된다(Fischbacher and Gächter, 2010; Fischbacher, Gächter and Fehr, 2001). 다른 한편 모든 경기자들이 동일한 사회적 선호를 갖더라도 그것이 개인들이 갖고 있는 자신의 이득 추구하고 결할되면서 도덕적 규범이 재조정되어 기여율을 하락시키는 요인이 되기도 한다(Figueres, Masclat and Willinger, 2011).

따라서 S 조건에서 평균 기여수준의 변화를 분석할 때, 반복에 따른 기여율의 하락이 나타난다면 그것은 학습이 진행되고 있음을 반영하고 있다고 볼 수 있으며, 게임의 종료시점에 0 이상의 기여율이 나타난다면 그것은 개인들이 가지고 있는 이타성을 반영하는 것이라고 판단할 수 있다.

둘째, P 조건하에서 실행된 공공재 게임에 대해서는 다음과 같은 가설이 세워질 수 있다.

- 전략 가설이 맞는다면 초반부에는 높은 기여율이 나타나다가 후반부로 갈수록 기여율이 하락할 것이다. 마지막 회에서는 전략적 고려가 작용할 여지가 없으므로 $x_{10}^S=0$ 일 것이다.

셋째, 마지막 회에서는 더 이상 전략적 고려가 작동하지 않으므로($x_{10}^S=0$), 그리고 10회 동안의 기간이 충분한 학습이 가능한 기간이라면($x_{10}^L=0$), 마지막 회에서 0 이상의 기여율이 나타난다면 그 (+) 수준의 기여율은 사회적 선호를 반영하는 것이다($x_{10}^R > 0$). 10회가 지난 후에도 여전히 학습이 불충분해서 $x_{10}^L > 0$ 이더라도, 10회라는 기간 동안 일어나는 학습의 크기가 S 조건에서나 P 조건에서나 다르지 않다면 P 조건에서 10회 째 나타나는 평균 기여 수준은 S 조건하에서 10회 째 나타나는 평균기여율과 동일할 것이다.

- 전략가설이 맞다면, 10회째 평균 기여수준은 P 조건에서나 S 조건에서 동일할 것이다.

4. 실험 결과

4.1. 평균 분석 결과

결과 1: S조건과 P조건 모두에서 1회 때 공공계정의 토큰 기여율은 평균 약 50% 정도였고, 점차 하락하여 종료시점 평균 기여수준은 두 조건 모두에서 초반 평균 기여수준보다 낮았다.

그림 1은 10 회를 반복하는 동안 P 조건과 S 조건 각각에서 평균기여율이 어떻게 변화했는지를 보여주고 있다. P 조건에서는 1회 때 평균 기여 토큰수가 9.66 개였고 10회에 이르면 평균 기여 토큰 수는 5.38개로 줄어들었다(실험에 참가한 30개 집단 평균들을 대상으로 한 Wilcoxon rank-sum test에 따라 1회와 10회에 차이가 없다는 귀무가설이 기각될 확률 p 값은

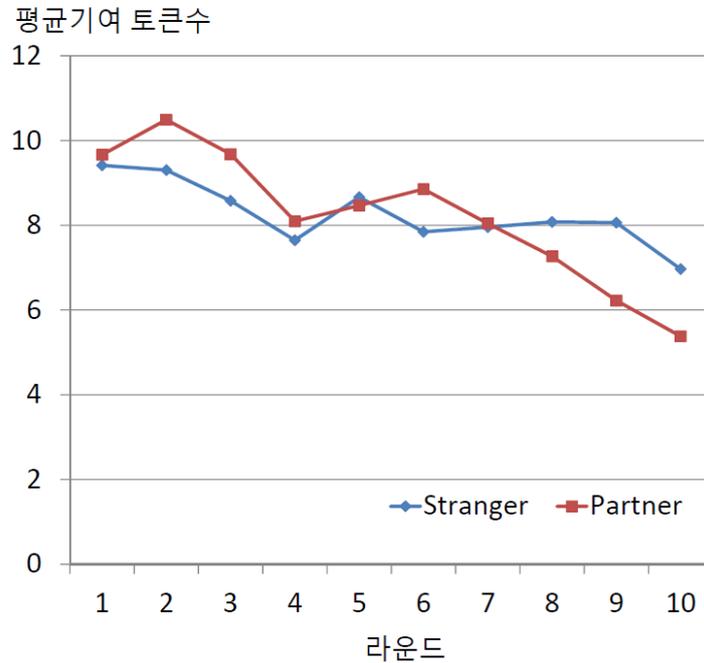


그림 1: 회 차에 따른 각 조건에서의 평균기여율의 변화

0.002). S 조건에서는 1회 때 평균 기여 토큰수가 9.41 개였고 10회에 이르면 평균 기여 토큰수는 6.97 개로 감소했다(Wilcoxon rank-sum test에 따른 p 값은 0.011).

S 조건에서는 전략적 고려가 작동하지 않을 것이므로, 1회 때의 기여율에 비해 10회 때의 기여율이 낮다는 것은 1회 때 기여수준 중 실험 참가자들의 게임구조에 대한 무지에서 비롯되는 부분이 존재한다고 해석할 수 있다. 여기서 주목할 것은 S 조건에서 10회에도 실험 참가자들은 총 20개의 토큰 중 평균 6.97개의 토큰을 공공계정에 기여했다는 것이다. 10회 동안의 반복이 게임의 구조 및 자신에게 유리한 전략이 무엇인지를 찾아내는 데 충분한 학습이 이루어졌다고 한다면, 마지막 회에서 나타나는 평균 6.97개의 토큰 기여 수준은 실험 참가자들의 타인 혹은 집단의 이익에 대한 고려를 반영하는 것이라고 볼 수 있다.

이 결과를 요약하자면, S조건에서 기여수준의 변화를 통해 우리는 실험 참가자들의 학습의 효과(회 차에 따른 기여수준의 감소) 및 사회적 선호의 존재(10회 때 0보다 큰 기여수준) 모두를 확인할 수 있다.

결과 2: 종료시점에 가까이 가면서 평균 기여 수준은 S 조건에서보다 P 조건에서 더 낮게 나타났다.

P 조건에서 9회와 10회 째의 평균 기여 토큰 수는 각각 6.22, 5.38 개였고, S 조건에서는 9회와 10회째 각각 8.06과 6.97개로 이 두 회에서 P조건에서의 기여수준이 S조건에서의 기여율보다 낮다(Wilcoxon rank-sum test에 따른 p 값은 9회에는 0.04, 10회에는 0.03).

S 조건에서는 전략적 고려가 작용할 여지가 없고, P 조건에서도 10회 째에는 전략적 고려가 작용하지 않는다. 따라서 두 조건에서 10회의 반복을 통해 학습을 거친 후에도 여전히 양의 기여수준을 보인다면, 그것은 실험 참가자들의 타인에 대한 고려, 즉 사회적 선호를 반영하는 것이라고 해석할 수 있다(결과 1 해석 참조). 실험 참가자들이 두 조건에 무작위로 배치 되었으므로 두 조건에 참가한 실험 참가자들 사이에 사회적 선호의 정도 차이는 없어야 한다. 두 조건 사이에 학습이 일어나는 속도가 다르지 않다고 한다면, 3.3절에서의 논의에 비추어 이야기하자면, P 조건에서 10회 째 나타나는 기여수준이 S 조건에서 10회째 나타나는 기여수준보다 작다는 사실은 P 조건하에서 기여율의 하락이 (1) 종료시점에 다가오면서 전략적으로 기여를 할 유인이 감소한다는 것에서 기인하는 측면과 (2) 10회 동안 반복을 통해 진행되는 학습에서 기인하는 측면 이외에 다른 요인에 추가로 작용하고 있다는 것을 암시한다. 다시 말해 이 결과는 P 조건에서의 기여율의 하락에는 학습과 전략이외에도 실험 참가자들의 상호적 맞대응, 즉 상대방의 기여수준이 낮다는 것이 확인되었을 때 이에 대한 보복의 의미로 자신의 기여수준을 낮추는 방향으로 대응한 것이 원인이 된다는 암시해주고 있다.

4.2. 회귀분석 결과

앞서 공공재 게임 집단의 평균적인 기여수준의 변화를 보다 엄밀히 분석하기 위해 우리는 전체 232명의 실험 참가자들의 기여수준에 대해 회귀 분석을 실시했다. 이를 통해 각 개별 실험 참가자 수준에서 기여율에 영향을 미치고 있는 요인들 중 학습 혹은 전략적인 고려를 통제하더라도 상호성이 여전히 유의미한 결정 요소임을 확인할 수 있다.

회귀분석에서는 i 번째 실험 참가자의 t 회 때의 기여토큰 수를 종속변수로 하고, (1) 회 차(round), (2) $t-1$ 회 때의 i 번째 실험 참가자 자신의 기여토큰 수($Cont_{i,t-1}$), (3) $t-1$ 회 때 자신을 제외한 다른 조원들의 기여토큰 수의 평균($Cont_{-i,t-1}$), (4) i 번째 실험 참가자가 참가한 실험이 P 조건 실험인지의

여부를 나타내는 더미(D_i^P : P 조건에 참여한 경우 1, S 조건에 참여한 경우 0), 그리고 마지막으로 (5) P 조건인지, S 조건인지의 여부와 위 (2), (3) 변수에 대한 교차 효과를 보여주는 항을 설명변수로 했다. 우선 최소자승법에 기반을 두고 모든 자료를 취합한 합동 회귀 모형(pooled regression)과 각 회의 고유 요소를 도입한 고정효과 모형 혹은 더미변수 최소자승 모형(least squares dummy variable model), 그리고 확률효과 모형에 따라 해당 계수들을 추정해보았다. 아울러 해당 모형이 시차 변수들을 설명변수로 지니기 때문에 최소자승법에 기반한 방법은 일치 추정량을 얻는 데 실패할 수 있다(Anderson and Hsiao 1981; Arellano and Bond 1991). 따라서 우리는 동태적 패널 모형에 따라서 점근적으로 일치 추정량을 확보하는 차분 GMM과 시스템 GMM에 따른 추정 결과와 함께 제시할 것이다(Arellano and Bond 1991; Blundell and Bond. 1998).

결과 3: 이전 회의에서의 타인들의 기여수준의 평균이 현재 자신이 기여수준에 미치는 영향은 P조건일 때 더 크다.

표 1은 앞서 말한 설명변수들을 가지고 회귀분석한 결과를 보여주고 있다. round 변수의 계수가 (-)의 값을 갖고 유의미하다는 것은 회 차가 진행됨에 따라 학습이 진행되고 있음을 의미한다. 표 1 따르면 회 차에 따른 학습을 통제했을 때, S 조건하에서의 $Cont_{-i,t-1}$ 의 계수의 부호는 양의 값을 갖는 것으로 나타났다(1% 수준에서 유의미함). 즉 S 조건에서도 상대방(들)이 자신보다 이전 회에 더 많이 기여했다면 그에 따라 이번 회 자신의 기여수준을 높이고, 반대로 상대방(들)이 자신보다 이전 회에 더 적게 기여했다면 그에 따라 이번 회 자신의 기여수준을 낮추는 것으로 나타났다. S 조건에서는 상호적 맞대응도 전략적 고려도 있을 수 없으므로, 이 계수 값은 상대방 행동을 관찰하면서 경기자들의 사회적 선호가 사회적으로 조정되고 있다는 것을 반영하는 것으로 해석할 수 있다. 즉 한편으로는 반복에 따라 게임의 구조를 점점 깨달으면서 내쉬균형 전략을 찾아나가는 학습이 진행되고(이것이 round의 계수로 통제되는 부분이다), 다른 한편으로는 상대방 경기자들의 행동을 관찰하고 이에 입각하여 일종의 사회적 모방을 통한 기여수준의 조정이 진행되었음을 보여준다.

마지막으로 $D^P \times Cont_{-i,t-1}$ 의 계수는 P 조건에서 학습 및 사회적 규범을 둘러싼 사회적 조정을 제외하고 난 후에 상대방의 행동에 대한 맞대응으로 남아 있는 부분, 즉 상호적 맞대응의 추가적 효과를 반영하는 것이라고 볼 수 있다. 표 1에 따르면 합동 회귀분석의 결과에서나 고정효과 모형 모두에서 이 값은 양의 값을 갖고 1% 수준에서 유의미했다. 우리의 결과에 따르면

표 1: P 조건과 S 조건에 대한 추정 결과

	합동회귀 모형 pooled regression	고정효과 모형 fixed effect regression	확률효과 모형 random effect regression
Intercept	3.488 (0.560)***	- (0.571)***	4.456
D_i^P	-1.787 (0.580)***	-1.796 (0.580)***	-2.178 (0.625)***
round	-0.129 (0.051)**	-	-0.149 (0.042)***
$Cont_{i,t-1}$	0.497 (0.026)***	0.498 (0.026)***	0.394 (0.027)***
$D_i^P \times Cont_{i,t-1}$	0.010 (0.037)	0.011 (0.037)	0.034 (0.039)
$Cont_{-i,t-1}$	0.148 (0.043)***	0.151 (0.043)***	0.149 (0.042)***
$D_i^P \times Cont_{-i,t-1}$	0.181 (0.057)***	0.182 (0.057)***	0.205 (0.059)***

주1: 괄호안의 숫자는 로버스트 표준오차. ***표시는 1% 수준에서, **은 5% 수준에서 그리고 *는 10% 수준에서 유의함을 나타냄.

주2: 232개의 개체와 9개의 시점을 지니는 2,088개의 관찰을 사용함.

주3: 고정효과 모형은 각 회차를 별도의 더미로 고려하여 추정한 결과

P 조건에서의 이전 기의 상대방의 기여 수준에 대한 경기자들의 반응은 S 조건에서와 비교했을 때 두 배 이상이 된다.

표 1에서 설명 변수에 시차변수를 둘 때 발생하는 추정량의 불일치성의 문제를 완화하기 위해서 동태적 패널 모형을 통해 차분 GMM과 시스템 GMM을 이용한 2단계 추정을 시도해보았다.

표 2는 앞서 최소자승법에 기반을 둔 추정치들이 일치성에는 문제가 있을지 모르나 그 부호가 가지는 의미는 여전히 타당함을 보여준다.³ 표 1

³도구변수가 제대로 설정되었음을 검증하는 Sargan test에서 차분GMM의 경우에는 문제가 없으나 시스템GMM의 경우에는 과잉식별의 가능성이 존재함을 알 수 있다. 한편, AR(2)에서는 양쪽 모두 계열 상관성이 사라졌음을 확인할 수 있다.

표 2: P 조건과 S 조건에 대한 GMM 추정 결과

계수	차분 GMM	시스템 GMM
$(1 - D_i^P) \times Cont_{i,t-1}^S$	-0.070 (0.174)	0.203 (0.174)
$D_i^P \times Cont_{i,t-1}^P$	0.508 (0.193)***	0.242 (0.181)
$(1 - D_i^P) \times Cont_{-i,t-1}^S$	-0.065 (0.153)	0.187 (0.089)**
$D_i^P \times Cont_{-i,t-1}^P$	0.972 (0.462)**	0.456 (0.145)***
Intercept	-	4.387 (0.768)***
회차 더미		
D^3	-1.118 (0.548)**	-0.604 (0.510)
D^4	-1.894 (0.548)**	-1.445 (0.462)***
D^5	-0.449 (0.640)	-0.096 (0.562)
D^6	-0.496 (0.624)	-0.668 (0.582)
D^7	-1.234 (0.562)	-1.086 (0.521)**
D^8	-1.293 (0.709)**	-1.514 (0.609)**
D^9	-1.004 (0.842)*	-1.495 (0.641)**
D^{10}	-1.195 (0.964)	-1.994 (0.627)***
Sargan Test	44.489 (p=0.12)	69.761 (p=0.0045)
AR(1)	-7.935 (p=0)	-6.488 (p=0)
AR(2)	-0.117 (p=0.45)	0.661 (p=0.25)

주1: 괄호안의 숫자는 로버스트 표준오차. ***표시는 1% 수준에서, **은 5% 수준에서 그리고 *는 10% 수준에서 유의함을 나타냄.

주2: 232개의 개체와 10개의 시점을 지니는 2,320개의 관찰 중 차분GMM에서 사용된 관찰수는 1,856개이고 시스템 GMM은 3,944개의 관찰을 사용함. 모두 2회부터 10회까지의 시차 종속변수를 도구 변수로 사용.

에서 확인한 $Cont_{-i,t-1}$ 계수의 유의성이 차분 GMM에서는 사라져있음을 제외하면, 차분 GMM과 시스템 GMM 모두 상대방의 기여 수준에 대한 반응에 있어서 P 조건에서의 $Cont_{-i,t-1}$ 의 계수값이 S 조건에서의 $Cont_{-i,t-1}$ 의 계수값보다 훨씬 크다는 점을 잘 보여준다. 회차 더미에 대한 추정결과는 뒤쪽으로 갈수록 기여율이 낮아지는 경향이 나타나는데(이 결과는 시스템 GMM에서 뚜렷이 나타난다), 이는 회차의 진행에 따른 학습의 진행(즉 OLS회귀결과에서의 round가 음의 계수값을 갖는다는 것)을 반영한다고 해

석할 수 있을 것이다.

결과 4. P 조건에서 상대방의 기여수준에 대한 경기자들의 반응은 회에 따라 크게 변하지 않는다.

마지막으로 우리는 P 조건에서 나타나는 상대방의 기여수준에 대한 경기자들의 반응이 전략적 고려에서 유래하는지를 확인하고자 했다. 상대방의 기여수준이 낮다면 상호적 맞대응이 아니더라도, 즉 전략적으로 고려하더라도 기여수준을 낮출 유인이 있을 수 있기 때문이다. 이 가능성을 점검하기 위해 우리는 P 조건에서 상대방의 기여수준에 대한 계수 값이 회 차에 따라 변하는지의 여부를 확인해보았다. 만일 상대방의 기여수준에 대한 경기자들의 대응을 나타내는 계수 값(즉 $Cont_{-i,t-1}$ 의 계수 값)이 전략적 고려를 반영한 것이라면 그 크기는 매 회 다르게 나타날 것이고, 특히 게임의 종료시점에서의 그 계수 값과 다른 시점에서의 계수 값의 차이가 나타날 것이기 때문이다.

이를 검증하기 위해서 우리는 비교적 단순한 접근법을 사용했다. 우리가 관심을 두는 바는 P 조건에서 매 회 마다 실험 참가자들의 기여 수준이 직전 회에서 같은 집단에 속한 다른 참가자들의 기여 수준에 영향을 받는 정도가 달라지는가의 문제다. 이는 사실상 각각 116개의 관찰로 구성된 매 회를 독립된 것으로 보아 해당 계수들을 추정한 후 다른 집단 구성원의 평균 기여 수준에 대한 반응이 회 차 별로 같은지 다른지에 대한 검정을 실시하는 방식으로 접근해볼 수 있다. 이러한 접근은 사실상 앞서 본 합동회귀 모형에서 매 회를 더미 변수로 설정한 뒤 이 더미 변수들과 다른 참가자들의 기여 사이의 교차효과를 살펴보는 것과 동일하다.⁴ 표 3의 추정 결과에 따르면,⁵ 각 회 차를 더미 변수로 설정하고 같은 팀에 속한 다른 상대들의 기여율에 반응하는 정도는 결과가 보여주듯이 전체에 대한 추정치와 유의미하게 다르지 않았다. 아래 더미와 곱해진 추정치들이 모두 0이라는 귀무가설도 F-검정을 통해 기각되지 않았다($p=0.760$).

즉 상대방들의 기여수준에 대한 대응 정도는 회 차에 따라 유의미하게 변하지 않고 안정적인 크기를 유지하고 있다는 것이다. 이는 P 조건에서 상대방의 기여 수준이 높을 때 자신의 기여 수준을 높이고, 상대방의 기여 수준이 낮을 때 자신의 기여 수준을 낮추는 식으로 행동하는 것이 전략적으

⁴패널 자료로 구조화하여 해당 계수 값을 보다 효율적이고 정확하게 추정하는 문제와 달리, 매 회의 독립적인 반응을 고찰하는 하는 데 있어서는 앞서와 같이 계열 상관의 문제를 고려할 필요가 없다.

⁵추정 계수가 많은 관계로 주장과 관련된 추정치만을 요약했다.

표 3: P 조건에서 상대의 기여에 대한 효과의 회 차별 추정 결과

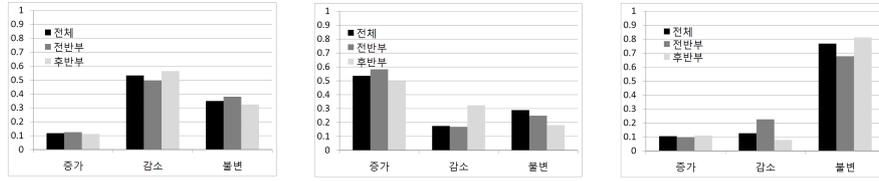
계수	추정치
(Intercept)	2.025 (1.528)
(중략)	
$Cont_{i,t-1}^P$	0.588 (0.077)***
$Cont_{-i,t-1}^P$	0.288 (0.130)**
$D^3 \times Cont_{-i,t-1}^P$	0.088 (0.179)
$D^4 \times Cont_{-i,t-1}^P$	0.075 (0.170)
$D^5 \times Cont_{-i,t-1}^P$	0.249 (0.172)
$D^6 \times Cont_{-i,t-1}^P$	-0.068 (0.169)
$D^7 \times Cont_{-i,t-1}^P$	0.005 (0.167)
$D^8 \times Cont_{-i,t-1}^P$	0.030 (0.169)
$D^9 \times Cont_{-i,t-1}^P$	0.048 (0.170)
$D^{10} \times Cont_{-i,t-1}^P$	0.010 (0.173)

주1: 괄호안의 숫자는 로버스트 표준오차. ***표시는 1% 수준에서, **은 5% 수준에서 그리고 *는 10% 수준에서 유의함을 나타냄.

주2: 116개의 개체와 9개의 시점을 지니는 1,044개의 관찰을 사용함.

로 이루어진 다기 보다는 상호적 맞대응에서 이루어진다는 우리의 해석을 뒷받침해준다고 볼 수 있다.

아래 그림 2는 P 조건에서 이전 기 상대방들의 기여수준이 자신의 기여수준보다 클 때, 작을 때 그리고 같을 때의 각각의 경우에 대해 이번 기 자신의 기여를 높인 사람, 줄인 사람, 불변인 사람의 비중을 나타내고 있다. 그리고 그 비중을 2-10회 (전체), 2-5회 (전반부), 6-10회 (후반부)로 나누어 나타내고 있다. 그림에서 보듯이 상대방의 이전 기 기여수준이 자신보다 높을 때 이번 기 기여수준을 높이고, 또 반대로 상대방의 이전 기 기여수준이 자신보다 낮을 때 이번 기 기여수준을 낮추는 식으로 행동하는 사람들이 과반수를 차지하고 있다. 그리고 이러한 상호적 맞대응을 하는 사람들의 비중은 전반기와 후반기를 비교했을 때 크게 변하지 않은 채 안정적으로 유지되고 있음을 확인할 수 있다. 다만 바로 전기 상대방의 기여수준이 자신보다 높았을 때에도 여전히 자신의 기여를 낮추는 사람들의 비중이 후반부에서 증가하는 것으로 보아 전략적 고려에 따라 행동하는 실험 참가



(a) $Cont_{i,t-1} < Cont_{i,t-1}$ (b) $Cont_{i,t-1} > Cont_{i,t-1}$ (c) $Cont_{i,t-1} = Cont_{i,t-1}$

그림 2: $t - 1$ 기 타인의 기여수준과 자신의 기여수준의 차이에 대한 t 기 자신의 기여수준의 증감여부

자들이 존재한다는 사실 자체는 부정할 수 없는 듯하다.

5. 결론

우리는 대부분의 기존 공공재 게임 연구들이 보고하고 있는 회 차의 진행에 따른 기여율의 하락이라는 현상을 설명하는 각 가설들의 설명력을 판단해보고자 했다. 우리의 실험 결과 역시 기존의 연구들과 유사한 기여율의 패턴을 보여주었는데, 우리는 이 결과를 토대로 다음과 같은 결론을 유도할 수 있었다. 첫째, 회 차의 진행에 따른 기여율의 하락은 개별 경기자들 차원에서 진행되는 학습(게임의 구조를 이해하고 무임승차가 자신에게 가장 유리한 전략임을 알아나가는 과정)을 반영한다. 둘째, 전략적 고려가 들어설 여지가 없는 최종 회에서의 기여율이 P 조건에서 더 낮다는 사실은 P 조건에서 낮은 기여율을 보이는 구성원에 대한 보복적 맞대응이 이루어지고 있음을 반영한다. 셋째, 상대방의 기여 수준에 대한 상호적 대응은 S 조건에서도 나타나지만 P 조건에서 더 크게 나타난다. S 조건에서 나타나는 상호적 대응은 공공계정의 기여를 둘러싼 사회적 규범이 타인의 행동을 관찰하면서 형성되고 있음을 의미하며, P 조건에서 나타나는 상호적 대응은 이러한 규범의 사회적 조정과정 뿐 아니라 상대방의 낮은 수준의 기여에 대한 보복, 그리고 상대방의 높은 수준에 대한 포상의 의미를 갖는 상호적 맞대응도 포함하고 있다.

본 연구의 결과는 학습, 전략, 상호성 세 가지 요인이 복합적으로 작용함으로써 공공재 게임에서의 패턴을 만들어낸다는 것을 보여주고 있다. 하지만 본 논문에서 강조하고자 했던 것은 학습 및 전략적 고려로 환원되지 않는 사회적 선호에 기반한 상호적 맞대응이 존재하며 이것이 기여율의 패턴에 중요한 요인으로 작용한다는 점이다. 이러한 점에서 본 연구는 공

공재 게임에서의 결과를 둘러싼 기존 논의를 한 걸음 진척시키는 데 기여할 수 있을 것이다.

참고문헌

- Anderson, T. W. and C. Hsiao (1981). Estimation of dynamic models with error components, *Journal of the American Statistical Association* 76, 589–606.
- Andreoni, J. (1988). Why free ride? Strategies and learning in public goods experiments, *Journal of Public Economics* 37, 291–304.
- Andreoni, J. (1995). Cooperation in public-goods experiments: Kindness or confusion? *American Economic Review* 85, 891–904.
- Andreoni, J., and R. Croson (1998). Partners versus strangers: Random rematching in public goods experiments, in Charles R. Plott and Vernon L. Smith (eds), *Handbook of Experimental Economics Results*.
- Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte-carlo evidence and an application to employment equations, *Review of Economic Studies* 58, 277–297.
- Blundell, R., and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel-data models, *Journal of Econometrics* 87, 115–143.
- Bowles, S. (1998). Endogenous preferences: The cultural consequences of markets and other economic institutions, *Journal of Economic Literature* 36, 75–111.
- Carpenter, J., Bowles, S., Gintis, H., and S-H. Hwang (2009). Strong reciprocity and team production: Theory and evidence, *Journal of Economic Behavior and Organization* 71, 221–232.
- Croson, R. (1996). Partners and strangers revisited, *Economics Letters* 53, 25–32.
- Croson, R., Fatas, E., and T. Neugebauer (2005). Reciprocity, matching and conditional cooperation in two public goods games, *Economics Letters* 87, 95–101.

- Croson, R. (2007). Theories of Commitment, Altruism & Reciprocity: Evidence from linear public goods games, *Economic Inquiry* 45, 199–216.
- Fehr, E., and U. Fischbacher (2002). Why social preferences matter-the impact of non-selfish motives on competition, cooperation, and incentives, *Economic Journal* 112, C1–C33.
- Fehr, E., and S. Gächter (2000). Cooperation and punishment in public goods experiments, *American Economic Review* 90, 980–994.
- Figuières, C., Masclet, D., and M. Willinger (2011). Weak moral motivation leads to the decline of voluntary contributions, CIRANO working paper.
- Fischbacher, U. (2007). Z-tree: Zurich Toolbox for readymade economic experiments, *Experimental Economics* 10, 171–178.
- Fischbacher, U., Gächter, S., and E. Fehr (2001). Are people conditionally cooperative? Evidence from a public goods experiment, *Economics Letters* 71, 397–404.
- Fischbacher, U., and S. Gächter (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments, *American Economic Review* 100 (1), 541–556.
- Houser, D., and R. Kurzban (2002). Revisiting kindness and confusion in public goods experiments, *American Economic Review* 92 (4), 1062–1069.
- Isaac, R. M., Walker, J. M., and S. H. Thomas (1984). Divergent evidence on free riding: an experimental examination of possible explanation, *Public Choice* 43 (2), 113–149.
- Isaac, R. M., Walker, J. M., and A. W. Williams (1994). Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups, *Journal of Public Economics* 54, 1–36.
- Keser, C. (2000). Cooperation in public goods experiments, CIRANO working paper.
- Keser, C., and F. van Winden (2000). Conditional cooperation and voluntary contributions to public goods, *Scandinavian Journal of Economics* 102 (1), 23–39.

- Kim, Oliver and J. M. Walker (1984). The free rider problem: experimental evidence, *Public Choice* 4, 3–24.
- Kreps, D., Milgram, P., Roberts, J., and R. Wilson (1982). Rational cooperation in finitely repeated prisoner's dilemma, *Journal of Economic Theory* 27, 245–252.
- Ledyard, J. O. (1995). Public Goods: A survey of experimental research, in Roth and Kagel (eds), *Handbook of Experimental Economics*.
- Levine, D. K. (1998). Modeling altruism and spitefulness in experiments, *Review of Economic Dynamics* 1, 593–622.
- Marwell, G., and R. E. Ames (1981). Economists free ride, does anyone else?, *Journal of Public Economics* 15, 295–310.
- Muller, L., Sefton, M., Steinberg, R., and L. Vesterlund (2008). Strategic behavior and learning in repeated voluntary contribution experiments, *Journal of Economic Behavior and Organization* 67, 782–793.
- Palfrey, T. R., and J. E. Prisbrey (1996). Altruism, reputation and noise in linear public goods experiments, *Journal of Public Economics* 61, 409–427.
- Selten, R., and R. Stoecker (1986). End behavior in sequences of finite prisoner's dilemma supergames: A learning theory approach, *Journal of Economic Behavior and Organization* 7, 47–70.
- Sobel, J. (2005). Interdependent preferences and reciprocity, *Journal of Economic Literature* 43, 392–436.
- Sonnermans, J., Schram, A., and T. Offerman (1999). Strategic behavior in public goods games: When partners drift apart, *Economics Letters* 62, 35–41.
- Weimann, J. (1994). Individual Behavior in a free riding experiment, *Journal of Public Economics* 54, 185–220.