

協力關係 構築에 있어서의 信賴의 役割：
南北韓 關係를 중심으로

安 錫 煥(淑明女大 教授)

林 源 赫(本院 專門研究員)

◇ 要 約 ◇

이 논문에서는 남북한 관계를 중심으로 분쟁해소와 협력관계 구축에 있어서 信賴가 할 수 있는 役割에 대해 검토해 보았다. 이 논문의 모델은 Kreps-Wilson(1982)의 신규진입 게임을 협력관계 구축이라는 새로운 상황에 맞게 변환시킨 순차적 게임모델로서 有限反復 게임구도와 非對稱的 情報所有를 그 근간으로 한다.

북한이 남한정부가 진보적이라고 믿는 확률을 남한에 대한 '信賴'라고 정의할 때, 신뢰는 크게 나누어 다음과 같은 역할을 한다. 북한이 남한에 대해 작은 규모의 신뢰감만 가지고 있어도 被疑者의 딜레마(Prisoner's Dilemma) 상황을 피할 수 있으며 상당한 기간동안 南北協力關係가 지속될 수 있다. 왜냐하면 보수적 성향의 남한정부도 북한이 개혁정책을 추진할 때 對北 압박정책 대신 包容政策을 통해 얻을 수 있는 수익이 적지 않은 점을 고려하여 상당기간동안 포용정책을 선택할 동기가 있기 때문이다. 하지만 이와 같이 자신의 성향을 은폐하는 위장행동의 가능성이 높은 동안 信賴는 쉽게 증진되지 않는다. 다만 남북한이 협력을 통해 실질적인 수익을 얻을 수 있을 뿐이다.

남북협력관계를 구축하기 위해 남한이 할 수 있는 일은 정책대응에 소요되는 시간을 줄이는 한편, 개혁정책을 통해 북한이 얻을 수 있는 期待收益을 늘리고 남한에 배신당할 경우 입는 피해를 줄여주는 것이다. 특히 장기적인 관점에서 북한의 기대수익을 점진적으로 늘려 '信賴도가 넘어야 할 문턱'을 낮추는 것이 효과적인 정책이다. 따라서 남북관계를 진전시키기 위해서는 북한의 지속적인 개혁을 조건으로 북한의 기대수익이 점진적으로 증가하는 KEDO식의 장기적 經協事業을 고안하여 추진해야 한다. 이에 반해 인도주의적 명분하에 일과성의 식량지원을 북한에 제공하는 것은 북한의 構造的 經濟難 해소나 남북관계의 지속적 진전에 별 도움이 안되는 정책이다.

북한은 북한 나름대로 남한을 한번 믿어볼 필요가 있다. 왜냐하면 상대방에 대한 不信은 결국 협력관계가 유지될 수 있는 기간을 감축시켜 비효율적인 결과를 초래하기 때문이다.

I. 序 論

1. 信賴에 관한 통상적 見解

사람들은 흔히 말하기를 남북한 사이의 뿌리깊은 불신 때문에 남북관계가 진전되지 못하고 있다고 한다. 어떤 이들은 남북한이 불신의 벽을 넘어 상대방을 “무조건 한번 믿어보면” 쌍방에 이로운 결과를 얻을 수 있을 것이라고까지 한다. 하지만 상호신뢰는 그렇게 쉽게 구축될 수 있는 것이 아니다. 신뢰는 자기가 가지고 있는 상대방에 대한 믿음을 상대방이 행동을 통해 확인시켜 줄 때 다져진다. 오랜 친구 사이에도 단 한번의 실수로 신뢰가 파괴될 수 있는데 전쟁까지 치른 남북한이 상대방을 믿기란 어려운 일일 것이다. 따라서 남북관계 진전의 실마리를 “무조건 한번 믿어보는 것”에서 찾는다든 것은 현실성이 없다고 할 수 있다.

하지만 상호신뢰가 전혀 없는 상태에서 관계가 진전되기도 어려울 것이다. 상호불신의 골이 깊은 경우, 이쪽에서 양보하면 저쪽도 성의를 보일 것이라는 희망 아래 온건노선을 채택하기보다는 강경노선으로 밀고 나가는 편이 배신당하는 데서 오는 피해를 방지할 수 있기 때문에 양쪽은 대결구도를 유지하게 될 가능성이 높고, 결국 어느 쪽도 불신을 해소하려는 엄두를 내지 못하기 때문에 긴장이 증폭되게 되는 것이다.

설령 한쪽에서 한번 속을 썩치고 온건노선을 채택해도 어느 정도의 신뢰가 존재하지 않는 한 상대방이 이에 화답할 가능성은 높지 않다. 상대방에 대한 신뢰 여부는 상대방의 행동을 해석하고 자기 자신의 대응책을 수립하는 데 지대한 영향을 미치기 때문이다. 예

를 들어, 남한의 對北지원은 북한을 남한에 의존하게 함으로써 흡수통일을 유도하려는 정책인가, 아니면 북한의 점진적 개혁을 지원하려는 것인가? 또 남한의 북방정책은 북한의 외교적 고립을 유도하여 체제붕괴를 촉진하려는 정책이었던가, 아니면 북한을 우회하여 북한의 우방국과 외교관계를 맺음으로써 북한을 남북대화의 장으로 끌어내고 교차승인의 원칙하에 북-미 및 북-일 수교노력을 지원, 남북 화해협력에 필요한 대외적 여건을 조성하려는 시도였는가? 반대로, 북한의 경제적 개혁·개방정책은 궁극적으로 북한체제를 공고히 하여 對南적화역량을 강화하려는 시도인가, 아니면 본격적 체제변혁의 출발점인가? 같은 행동에 대해서라도 그 행동의 의도나 궁극적 목표에 대한 의견이 다를 수 있는데, 상호신뢰가 존재하지 않는 상황에서는 상대방의 행동이 나쁜 방향으로 해석되고 대응책 또한 비협조적인 방향으로 설정될 가능성이 높게 되므로 이쪽에서 아무리 선의의 의도를 가지고 그동안의 상호 불신을 해소하여 관계진전을 모색해 보려고 해도 저쪽에서 호응해 주지 않기 때문에 관계개선은 요원하게 되는 것이다.¹⁾

이와 같이 기존의 적대관계를 해소하고 서로 협력하면 양쪽 다 이익을 볼 수 있다는 사실을 알면서도 협력하지 못하는 상황을 타개하는 데 있어 신뢰는 어떤 역할을 할 수 있는가? 믿을 만한 행동을 해야 신뢰가 생기는 것인가, 아니면 신뢰가 있어야 믿을 만한 행동을 할 수 있는 것인가? 마치 “닭이 먼저냐, 달걀이 먼저냐”는 논쟁을 연상케 하는 이 상황을 어떻게 해소할 수 있는가?

1) 분쟁을 해소하고 관계진전을 이루기 위해서는 상대방에 대한 기존감정을 협상내용에서 분리하여 상대방의 ‘궁극적인’ 의도나 동기에 초점을 맞추기보다는, 우선 객관적 원칙을 설정하고 협상을 통해 얻을 수 있는 수익과 위험부담을 냉철하게 계산한 후에 따라 행동하는 것이 바람직하다. 이와 같은 협상 원칙에 의거한 원칙협상(principled negotiation)에 대해서는 Roger Fischer and William Ury, *Getting to YES: Negotiating Agreement Without Giving In* (New York: Houghton Mifflin, 1981)를 참조할 것.

2. 信賴에 관한 게임이론적 分析

이 글에서는 분쟁해소와 협력관계 구축에 있어서의 신뢰의 역할에 대해 게임이론적인 관점에서 분석해 보고 그 교훈을 남북한 관계에 적용해 보고자 한다. 게임이론적인 관점에서 볼 때 남북한 관계는 피의자의 딜레마(Prisoner's Dilemma) 게임을 연상케 한다.²⁾ 즉 긴장해소와 경제협력을 통해 서로 얻을 것이 많은 남북한이 자기 자신과 상대방의 행동을 구속하는, 믿을 만한 협조 장치(credible commitment to cooperation)를 찾지 못하기 때문에 (배신, 배신)의 결과를 유발하고 있는 것으로 볼 수 있기 때문이다. 주지하는 바와 같이 상대방의 수익구조(payoff structure)가 완전히 노출되어 있는 상황에서 행해지는 피의자의 딜레마 게임은, 아무리 반복되어도 그 횟수가 한정되어 있는 한 (배신, 배신)의 결과를 낳으며 이 과정에서 신뢰가 할 수 있는 역할은 없다고 할 수 있다.³⁾

하지만 상대방의 수익구조가 확실하지 않고 지배적 전략(dominant strategy)이 존재하지 않을 경우, 상대방의 성향에 대한 주관적인 판단(prior belief)은 게임의 결과에 상당한 영향을 미칠 수 있다. 또 상대방은 이와 같은 불확실성을 역이용하여 자신의 성향을 상당 기간동안 위장함으로써 더 높은 수익을 올릴 수도 있다. 예를 들어 신규진입 게임의 경우, 신규진입기업이 독점기업의 성향에

2) Prisoner's Dilemma는 보통 '囚人の 딜레마'로 번역되지만, Albert Tucker가 소개한 게임에서의 prisoner는 죄수나 그 동의어인 수인처럼 '죄를 지어 교도소에 갇힌 사람'이 아니라 법정에서 형이 확정될 때까지 갇혀 있는 刑事被疑者 또는 未決囚를 뜻한다. 따라서 이 논문에서는 Prisoner's Dilemma를 피의자의 딜레마로 번역하였다. Prisoner's Dilemma 및 협력관계 구축과 연관된 흥미있는 실험에 대해서는 Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984)를 참조할 것. 게임이론을 통해 남북한 관계를 분석한 저서로는 김재한, 『게임이론과 남북한 관계: 갈등과 협상 및 예측』(서울: 한울, 1995)을 참조할 것.

3) 이 결과에 대해서는 Reinhard Selten, "The Chan-Store Paradox," *Theory and Decision* 9(2), 1978, pp. 127~159를 참조할 것.

대한 확실한 정보를 가지고 있지 못할 때에는 온건한 독점기업도 게임의 초기 단계에서는 신규진입기업에 대해 강경하게 대처함으로써 이익을 늘릴 수 있는 것이다.⁴⁾ 이 논문에서는 이와 같은 점에 착안하여 Kreps-Wilson(1982) 모형을 응용, 상당히 작은 규모의 신뢰도 분쟁해소와 협력관계 구축에 큰 역할을 할 수 있다는 것을 순차적 게임모형을 통해 입증하고자 한다. 여기에서 ‘신뢰’(trust)란 상대방이 협조적 성향을 가지고 있다고 믿는 확률이며, 이 신뢰도는 각 期마다 Bayes Rule에 따라 수정된다.

3. 論文의 基本構圖

이 논문에서 제시될 남북한 모델의 기본틀은 다음과 같다. 북한은 경제난과 전반적인 체제위기를 타개하기 위해 개혁정책과 체제고수정책 사이에서 선택을 해야 하는 상황에 처해 있고, 남한정부는 북한의 움직임에 따라 포용 또는 압박정책으로 대응할 수 있다. 북한은 남한정부가 진보적(협조적)인지 보수적(비협조적)인지 알지 못하며, 남한정부의 성향에 대한 주관적인 판단에 따라 정책선택을 해야 한다. 이 모델에서 진보파는 긴장해소와 경제협력에 큰 의의를 부여하는 반면, 보수파는 북한에 대해 정치군사적 우위를 확보하는 것을 선호한다. 따라서 진보파는 북한이 정치경제적 개혁 및 軍備統制를 시도하는 한 이에 상응하는 포용정책을 추진하지만, 보수파는 북한이 개혁지향적 노선을 표방하더라도 그 체제를 더욱 약화시키기 위해 경제제재와 군비증강 노력을 경주할 수도 있다.

4) 수익구조의 불확실성과 연관된 ‘명성 효과’(reputation effect)에 대해서는 David Kreps and Robert Wilson, “Reputation and Imperfect Information,” *Journal of Economic Theory* 27(2), 1982, pp. 253~279와 David Kreps, Paul Milgrom, John Roberts, and Robert Wilson, “Rational Cooperation in the Finitely Repeated Prisoners’ Dilemma,” *Ibid.*, pp. 245~252를 참조할 것.

그러나 보수파도 전쟁위험까지 감수하며 북한을 고립시키고 남북 긴장을 지속적으로 증폭시키기에는 위험부담이 너무 크기 때문에 북한체제붕괴와 흡수통일에 대한 차선책으로서 남북협력을 선택할 수도 있다. 즉 보수적 성향을 가진 남한정부라도 처음부터 강경책으로 밀어붙여 전쟁위험을 고조시키기보다는 북한의 신뢰를 역이용하여 최소한 초기단계에서는 진보파처럼 행동함으로써 더 높은 수익을 올릴 수 있는 것이다. 하지만 게임의 종료시기가 다가오게 되면 보수파가 진보파처럼 행동함으로써 얻을 수 있는 수익은 상대적으로 줄어들게 되므로 북한에 대한 강경노선으로 전환할 가능성이 높아지게 된다.

이 모델에서 신뢰는 협력관계를 구축하는 데 중요한 역할을 한다. 만약 북한이 남한정부에 대해 전혀 신뢰감을 가지고 있지 않다면 북한은 강경노선으로 일관할 수밖에 없고 이에 대해 남한정부도 강경노선으로 대응할 수밖에 없겠지만, 신뢰감이 조금이라도 있다면 상호협력관계를 구축할 수 있는 길이 열리게 되는 것이다. 특히 남한정부는 정책대응 모색기간을 줄이는 한편, 북한이 상호협력을 통해 얻을 수 있는 기대수익을 점진적으로 늘리고 남한에게 배신당하는 데서 오는 피해를 줄여줌으로써 남북관계 진전에 필요한 신뢰도를 낮출 수 있다.

또 이 모델에서의 신뢰는, 보수파가 진보파로 위장할 수 있는 한 쉽게 증진될 수 없다는 특성을 가진다. 예를 들어 보수파도 진보파처럼 행동할 가능성이 높은 게임 초기단계에서는 경제협력 및 군비통제와 같은 포용정책도 북한이 남한정부의 진정한 성향을 파악하는데 도움을 주지 못하므로 남북한간의 신뢰를 증진하는 데는 별 구실을 못하게 된다. 즉 신뢰는 상대방이 비협조적 또는 기회주의적인 행동을 할 가능성이 높음에도 불구하고 협조적인 행동을 할 때 비로소 증진되는 것이다. 하지만 비록 남북간의 신뢰가 증진되지는 않더라도

도 쌍방은 경제협력과 군비통제 등을 통해 실질적인 수익을 얻을 수 있으므로 보수파가 진보파처럼 행동하는 것은 바람직한 일이다.

북한의 경우, 일정 한도내에서는 남한에 대한 초기 신뢰도가 높으면 높을수록 파레토 효율적인 결과를 얻는다. 왜냐하면 북한의 초기 신뢰도가 높으면 높을수록 남한의 보수파가 진보파처럼 행동할 가능성도 높아지고 따라서 남북협력을 통해 쌍방이 누릴 수 있는 혜택도 늘어나기 때문이다.

II. 모 델

1. 南北韓의 基本戰略과 有限反復게임의 構圖

이 모형에서는 각 時期마다 북한이 먼저 정책선택을 하고 이에 대해 남한이 대응하는 것으로 간주한다. 북한정부는 남한정부가 대북정책에 있어서 진보적인지 보수적인지 사전에 알지 못하며, 남한 정부의 정책대응을 관찰함으로써 남한정부의 성향에 대한 자신의 생각을 고쳐나갈 뿐이다. 이 게임은 한쪽의 정권이 교체되어 상대방의 성향에 대한 평가가 새로 이뤄져야 할 때까지 계속된다. 현 한반도 정치구도하에서는 남한의 대통령이 5년에 한번씩 새로 선출되므로 게임은 한시적일 수밖에 없고, 이 기간내에는 남북간의 정책교환 횟수도 제한될 수밖에 없으므로 게임은 무한정 반복되지 않는다. 여기에서 한 時期는 원칙적으로 1개월이나 1년처럼 고정된 일정 기간이 아니라 남북한 사이에 정책교환이 이뤄지는 기간을 의미한다. 따라서 남북한 사이에 정책선택과 대응이 빠르게 이뤄질수록 한 時期의 실제시간은 줄어들게 되며 게임이 반복되는 횟수는 늘어

나게 된다.⁵⁾

북한은 크게 나뉘 개혁과 대결이라는 두 개의 정책 사이에서 선택을 할 수 있다. 북한은 구조적인 경제난을 극복하기 위해 개혁전략을 채택하여 외부세계에 대한 개방조치와 시장지향적 체제개편조치를 취하고 군사비를 축소할 수 있다. 반대로, 북한은 현 체제 고수를 목표로 외교적 고립노선과 명령경제체제를 고집하고 주민에 대한 통제를 더욱 강화할 수도 있다.⁶⁾

남한은 크게 나뉘 포용과 압박이라는 두 개의 전략으로 정책대응을 할 수 있다. 체제개혁을 향한 북한의 조심스런 움직임에 화답하고 북한내 온건파에 힘을 실어주기 위해 남한은 포용전략의 기초하에 인도주의적 차원의 지원을 제공하고, 경제협력을 확대하며, 남북간 신뢰구축 및 군비감축정책을 추진할 수 있다. 이와는 반대로, 북한을 압박하기 위해 방어용 이상으로 군비를 증강하고, 북한에 대한 외교적 봉쇄조치를 취하며, 경제협력 및 지원을 줄이는 정책도 취할 수 있다.

-
- 5) 예를 들어, 남북한이 관망 및 탐색 기간을 줄여 정책선택과 대응이 평균 2개월마다 한번씩 이뤄지게 한다면 5년 동안에는 총 30회의 정책교환이 있게 되고 이 게임은 30期 게임이 되는 것이다.
- 6) 북한은 이 두 전략을 절충하여 對外的으로 남한에 대해 대결정책을 유지하면서도 對內的으로는 경제난 타개를 위해 체제개혁을 추진할 수 있다. 실제로 북한은 조심스럽게 개혁정책을 추진하면서 1994년 10월의 북-미 제네바 합의 이후, 미국과 일본 등에는 유화적 태도를 보이면서도 남한은 배제하는 듯한 태도를 보여왔다. 하지만 북한은 남한전체가 아니라 단지 남한정부에 대해 강경한 태도를 취하고 있을 뿐이다. 서독기업의 직접투자를 허가하지 않았던 옛 동독정부와는 달리 북한은 남한기업에 우호적인 태도를 보이고 있다. 현실적으로 볼 때 남한을 제외하고 북한이 현재 경제개발에 필요한 막대한 자금을 지원받을 수 있는 곳은 일본뿐이며, 남한이 1965년 대일 청구권 명목으로 받은 8억달러가 이자와 환율 등을 고려한다 해도 현 시가로 100억달러 정도밖에 되지 않는다는 점을 감안하면, 북한이 남한기업에 대해서까지 대결정책을 추진할 가능성은 낮다. 따라서 북한의 對南 대결-개혁 절충전략은, 남한정부가 중국과 대만처럼 先경제적 실리추구-後정치적 타협 모색이라는 소위 ‘政經分離’ 원칙에 따라 행동한다면 개혁정책으로 귀결될 수 있는 여지가 있는 것이다. 이와 같은 점을 고려하여 본 논문에서는 북한의 정책을 크게 대결-체제고수와 유화-체제개혁 전략으로 나뉘 2×2 구도를 유지하려 한다.

북한의 개혁과 남한의 포용전략을 통해 쌍방이 얻을 수 있는 수익은 다음과 같이 크게 안보와 경제적 측면으로 나뉘 생각할 수 있다. 우선, 남북한 사이의 화해협력과 군비통제는 한반도 내의 긴장 상태를 해소하고 동족상잔의 비극이 다시 한번 연출되는 것을 방지할 수 있을 것이다. 또 남북경협은 북한의 노동력과 남한의 자본 및 경제개발 경험을 결합함으로써 남북 양쪽에 경제적 이득을 제공할 뿐 아니라 장기적으로는 남북간의 소득격차를 해소하여 통일비용을 줄이는 역할도 할 수 있을 것이다.⁷⁾

2. 南韓政府의 性向：進歩와 保守의 論理

앞에서 언급한 바와 같이 남한정부는 대북정책을 추진함에 있어 진보적 또는 보수적일 수 있다.⁸⁾ 진보적 성향의 남한정부는 북한에

7) 남북한의 軍備統制와 함께 북한의 경제개발과 체제전환에 이은 남북한의 경제 통합은 남북통일을 위해 해결되어야 할 가장 중요한 과제 중의 하나이다. 현재 남한은 지속적인 고도의 경제성장을 통해 매년 북한경제 전체의 규모보다 더 큰 경제를 창출해 내고 있으며 남북한의 일인당 소득격차는 10 대 1 이상 벌어진 것으로 추정되고 있다. 만약 남북한의 소득격차가 어느 정도 해소되지 않고 통일이 이뤄진다면 대규모의 '무작정 월남' 현상으로 인해 남북한 노동시장의 통합이 심각한 후유증을 유발하게 될 수도 있다. 이와 같은 사태를 미연에 방지하기 위해서는 남북한 경제의 보완성을 충분히 활용하여 통일이 이뤄지기 전에 북한의 주민소득을 최소한 남한의 50~60% 수준으로 끌어올리는 경제협력 조치가 필요하다. 물론 이와 같은 조치는 북한이 대남적화전력을 포기하고 진정으로 개혁을 추구할 경우에만 실현될 수 있을 것이다. 남북통일의 경제적 과제에 대해서는 전홍택, 『북한의 체제전환과 남북한 경제통합의 주요 과제』, KDI 정책연구시리즈 96-01, 1996을 참조할 것.

8) 진보적이든 보수적이든 남한정부는 안보 및 경제적 측면을 고려하여 합리적으로 수익을 계산하고 단지 대북정책의 기본원칙 및 가정에서만 차이를 보인다. 즉 이 게임의 결과는 비합리적인 플레이어의 존재 여부에 달려 있는 것이 아니다. 물론 신진기업을 못살게 구는 데서 즐거움을 얻는 독점기업의 경우처럼 보수적 성향의 남한정부도 북한을 곤궁에 빠지게 하는 데서 심리적 만족감을 얻는다고 가정해도 되겠지만, 이 게임에서는 안보 및 경제적 고려에 기초한 합리적 계산만으로도 게임의 결과를 충분히 얻을 수 있다는 점을 부각시키기 위해 '비합리적' 요소의 개입을 배제한다. 만약 국내정치, 특히 선거前的 대북강경정책을 통한 안보정국 조성 등이 보수파의 계산에 포함된다면 진보파와 보수파 사이의 차이는 더욱 분명해질 것이다.

대해 군사적 우위를 확보하려고 하는 것은 군비경쟁을 격화시킬 뿐, 평화공존을 위해 불필요하다고 생각하고 ‘합리적 충분성’ 원칙에 입각한 군사적 균형 확립이 더 효과적인 안보정책이라고 믿는다. 진보파는 경험확대를 통해 물질적 수익뿐 아니라 한반도 내의 긴장을 해소하는 효과도 얻을 수 있으리라고 생각한다. 따라서 진보적 성향의 남한정부는, 북한정부가 개혁조치를 취했을 때 반드시 포용정책으로 화답한다. 진보파에게 있어 남북간 협력에서 이탈하는 행위는 교역을 통해 경제적 이득을 취할 기회를 사전에 차단하므로, 압박정책은 일회게임에서도 지배적 전략이 되지 못하고 단지 북한이 먼저 강경노선을 채택했을 때 사용하는 방어적 조치의 성격을 가진다.

이에 반해 보수파는 ‘힘에 의한 평화’ 원칙에 따라 행동하고 북한에 대한 정치군사적 우위를 확보하는 것을 기본목표로 하며 경제적 이득은 부수적 차원에서 고려한다. 만약 군비경쟁, 경제적 제재, 그리고 외교적 봉쇄에 의한 소위 북한 압살정책이 제2의 한국전쟁을 촉발시키지 않고 북한의 붕괴나 항복을 유도할 수 있다면 좋겠지만, 이와 같은 강성 압박정책은 위험부담이 클 뿐 아니라 주변 강대국의 지지를 받을 확률도 적으므로 보수적 성향의 남한정부도 이런 정책을 무조건 밀어붙이지는 못한다.⁹⁾ 또 보수파로서도 체제 불안정으로 예측 불가능한 북한정권과 상대하기보다는, 비록 북한이 남한의 뜻대로 순순히 움직여주지는 않더라도 개혁을 지향하는 편이

9) 실제로 남북한은 1994년 6월, 북한의 핵개발 의혹에 대한 한-미 정부의 비포괄적 접근방법이 북한의 양보를 얻어내지 못하게 되자 전쟁 일보직전 단계까지 나아갔다. 결국 카터 前 미국 대통령의 訪北으로 극적으로 해소된 이 전쟁 위기는, 남북한의 강경 맞대응이 엄청난 결과를 초래할 수 있다는 것뿐만 아니라 ‘통일비용’을 논함에 있어 남북한의 잠재적 ‘분단비용’도 고려되어야 한다는 것을 보여준다. 북한 핵문제를 외교적 ‘주고 받기’ 대신 죄를 지은 자에 대해 벌을 내린다는 관점에서 접근했던 미국 정책의 한계에 대해서는 Leon V. Sigal, “American Nuclear Diplomacy with North Korea,” paper presented at the 8th Hokkaido Conference for North Pacific Issues, Oct. 28-29, 1996을 참조할 것.

나으므로 포용정책은 차선책이 될 수 있다. 북한 압살전략이 가져올 불확실성을 감수하지 못할 바에야 남북경협과 군비통제를 통해 얻을 수 있는 한반도의 긴장완화와 경제적 이득을 취하는 편이 낫기 때문이다. 이런 상황하에서는 보수적 성향의 남한정부가 일정기간동안은 진보파처럼 행동하는 것이 더 이로울 수 있다.

3. 南北韓의 政策選擇과 對應

이 모형의 일회게임(stage game)에서 발생할 수 있는 결과들은 다음과 같이 요약할 수 있다.

(개혁, 포용)

남한은 북한의 개혁정책이 남북간의 우호적인 분위기 속에서 진전될 수 있도록 포용정책으로 화답하며, 궁극적인 통일을 목표로 협력관계를 발전시켜 나간다. 이 결과는 피의자 딜레마 게임의 (협력, 협력)과 유사하며, 화해협력-남북연합-통일로 이어지는 ‘한민족 공동체’ 통일방안은 바로 이런 구도를 상정한 통일방안이다.

(개혁, 압박)

북한의 개혁 움직임에도 불구하고 남한은 강압적인 조치를 취한다. 예를 들어, 남한정부는 북한에 대한 외교적 봉쇄정책을 취하고 대북투자 등 본격적인 경제교류를 지연시키면서 북한에 좀더 근본적인 체제개혁조치와 군비감축을 취할 것을 요구한다. 외부자본의 유치가 이뤄지지 않는 상황에서 북한의 경제자유화 조치는 실질적인 경제적 혜택을 낳지 못하고 중앙의 통제능력을 상실하게 할 뿐이며, 북한이 평화배당금을 창출하기 위해 개혁정책의 일환으로 실시한 군비통제는 북한의 군사적 위치를 상대적으로 약화시킨다. 결

국, 북한은 남한에 이용당했다고 생각하게 된다.¹⁰⁾ 이 상황은 피의자 딜레마 게임의 (협력, 배신)과 유사하다.

이 상황은 북한에 있어서는 최악의 시나리오이지만, 북한체제의 약화를 노리는 보수적 성향의 남한정부에게는 최선이다. 반면, 북한체제의 점진적 개혁을 바라는 진보적 성향의 남한정부에게 이 결과는 (개혁, 포용)에 비해 못한 결과이다.

(대결, 포용)

북한은 개혁조치를 전혀 취하지 않으면서 대남적화전책을 고수하며 긴장을 증폭시키려 한다. 하지만 남한은 경제적 지원을 제공할 뿐 아니라 일방적으로 군비축소를 실시한다. 마치 1936년에 비무장지대인 라인란트에 독일군을 배치하고 1938년에는 오스트리아를 합병한 히틀러를 제재하기는커녕 체코슬로바키아가 수데텐란트를 독일에 양보하도록 종용한 영국처럼, 평화를 위협하는 북한을 남한이 일방적 유화정책을 통해 보상하는 상황이 전개된다.

이 결과는 북한에 있어서 최선이다. 남한의 일방적 군비축소로 북한의 군사적 위치가 향상될 뿐 아니라 남한이 제공하는 경제적 지원을 통해 본격적인 체제개혁 없이도 경제난을 어느 정도 해소할 수 있게 되기 때문이다. 그러나 진보적이든 보수적이든 이 결과는 남한정부에 있어서 최악의 시나리오이므로 순차적 게임에서 북한의 강경정책에 대해 포용정책으로 대응하는 경우는 없다.¹¹⁾

10) 북한내에서 상대적으로 개혁지향적 인사로 분류되는 김정우 대외경제협력추진위 위원장은, 1996년 8월 『주간조선』과의 인터뷰에서 통신분야를 대북투자 지연의 예로 들면서, 1995년 2월 삼성과 계약을 추진했지만 한국정부가 허가를 해주지 않아 결국 그해 9월 태국의 Loxley 그룹과 계약을 체결했다고 밝히며, 한국이 수차례 공식표방한 것과는 달리 남북한간의 본격적인 경제협력과 교류를 추진하지 않는 저의를 추궁했다.

11) 여기에서 한가지 유념할 점은 북한의 강경책에 맞서는 남한의 대응강도 조절일 것이다. 비록 이 모델과 같은 2 × 2 구도에서는 나타낼 수 없지만 압박정책에는 다양한 유형이 있을 수 있고 그 수位도 어느 정도 조절할 수 있을 것

(대결, 압박)

남북한은 상대방에 대한 불신감을 극복하지 못하고 기존의 대결 상태를 유지한다. 남북간의 군비증강에 의해 한반도 내의 긴장감은 고조된다.¹²⁾ 이 상황은 피의자 딜레마 게임의 (배신, 배신)과 유사하다.

북한의 정책선택과 남한의 대응에 따른 (북한, 남한)의 수익을 표로 정리하면 다음과 같다. 대문자로 표기된 수익은 북한의 개혁을 전제로 한다.

	(북한, 진보적 남한정부)	(북한, 보수적 남한정부)
(개혁, 포용)	(E, F)	(E, F*)
(개혁, 압박)	(G, H)	(G, H*)
(대결, 포용)	(e, f)	(e, f*)
(대결, 압박)	(g, h)	(g, h*)

남북한의 수익 순위는 다음과 같다.

북한	진보적 남한정부	보수적 남한정부
$e > E > g > G$	$F > H > h > f$	$H^* > F^* > h^* > f^*$

이다. 예를 들어, 북한의 대결정책에 남한이 초강경책으로 맞대응하는 것은 북한내 강경파의 입지만 강화해 주는 부작용을 낳을 수도 있으므로 북한이 강경노선에서 온건노선으로 전환하도록 유도하기 위해서는 압박정책의 강도를 적절한 수준으로 설정해야 할 것이다.

12) 물론 현 상황에서 북한보다 잃을 것이 많은 남한은 압박전략을 추진하면서도 '부자 몸조심'하는 심정으로 북한체제가 견잡을 수 없는 혼란에 빠지거나 궁지에 몰린 북한정권이 무모한 도발을 감행하지 않도록 하기 위해 한반도 안정유지에 필요한 최소한의 대북지원은 할 수 있다.

Ⅲ. 게임의 結果

1. 順次均衡의 構圖: 均衡戰略과 信賴度

이 순차적 게임에서 남아 있는 時期를 n 이라고 하자. 즉 T 期 게임에서 $n=1$ 은 이 게임의 마지막 期를 뜻하며 $n=T$ 는 첫 期를 뜻한다. 각 期마다 북한은 개혁과 대결정책 사이에서 선택을 해야 하고, 남한은 포용 또는 압박정책으로 대응한다. n 期가 남았을 때 북한이 남한을 진보적이라고 믿는 확률을 P_n 이라고 하자. 그렇다면 P_T 는 게임이 시작하기 전에 북한이 남한정부가 진보적이라고 믿는 확률, 즉 ‘초기 신뢰도’라고 할 수 있을 것이다. 어느 期에서든 북한이 대결정책을 채택하면 남한은 성향에 관계없이 압박정책으로 대응하므로 이 경우 남한의 성향에 대한 북한의 판단은 수정될 수 없다($P_n = P_{n+1}$). 또 북한이 개혁정책을 취했을 때 보수적 남한정부가 진보적 성향을 가진 듯 행동한다면 이 경우에도 남한의 성향에 대한 북한의 생각은 수정될 수 없다. 결국, 북한이 남한의 성향을 완전히 파악할 수 있는 경우는, 자신의 개혁정책에 대해 남한이 압박정책으로 대응함으로써 남한정부가 보수파임이 드러났을 때뿐이다. 그 외의 경우에 북한은 보수적 남한정부가 진보적으로 행동할 확률을 고려하여 남한정부의 성향에 대한 판단을 수정해 나갈 수밖에 없다. 여기에서 유념할 것은 보수적 남한정부가 북한의 개혁정책에 압박정책으로 대응하여 이득을 취할 수 있는 것은 한 期에 국한된다는 점이다. 북한은 남한이 보수적이라는 것을 파악하게 되면 그 이후에는 대결정책을 채택하게 되기 때문이다. Selten의 연쇄점 패러독스가 시사하는 것처럼 수익구조가 완전히 밝혀지면 일회게임

이 계속 반복되게 되는 것이다.¹³⁾

남한정부의 성향에 대한 북한의 판단은 다음과 같이 요약될 수 있다.

- (1) 게임이 시작될 때($n = T$) 북한이 남한에 대해 가지는 신뢰도는 P_T 이다.
- (2) 북한이 $n+1$ 시점에서 대결정책을 선택한다면, 남한은 압박정책으로 대응하게 되고 $P_n = P_{n+1}$.
- (3) 북한이 $n+1$ 시점에서 개혁을 선택할 때 남한이 포용정책으로 대응하고 $P_{n+1} > 0$ 이라면 $P_n = \max[L^n, P_{n+1}]$. 여기서 $L = (g - G) / (E - G)$ 로서 일종의 분기점(threshold) 역할을 한다.
- (4) 북한이 $n+1$ 시점에서 개혁을 선택할 때 남한이 압박정책으로 대응하거나 $P_{n+1} = 0$ 이라면 $P_n = 0$.

위의 결과에서 L 은, 북한이 대결정책을 통해 얻을 수 있는 수익 g 와 남북협력을 통해 얻을 수 있는 수익 E 를, 북한이 남한에 배신당할 경우의 수익 G 를 기준으로 하여 비교한 수치이다. L 이 낮으면 낮을수록 북한의 남한에 대한 신뢰도가 낮아도 남북협력이 이뤄질 수 있다.

순차균형(sequential equilibrium) 개념을 사용할 때 이 순차적 게임의 균형전략은 다음과 같은 특성을 갖는다. (증명은 부록으로 넘긴다)

북한의 전략:

- (1) $P_n > L^n$ 일 경우 북한은 1의 확률로 개혁정책을 선택한다.

13) 즉 상대방의 배신에 대한 응징을 일회로 한정하고 제2의 기회를 주며 받은 만큼 갚는 tit-for-tat 전략 대신 상대방이 한번 배신하면 끝까지 용서하지 않겠다는 이른바 grim 전략이 이 모델의 균형전략의 기초를 이루고 있는 것이다.

- (2) $P_n < L^n$ 일 경우 북한은 1의 확률로 대결정책을 선택한다.
- (3) $n = 1$ 에서 $P_1 = L^1$ 일 경우 북한은 r 의 확률로 개혁과 대결정책을 혼합한다.

보수적 남한정부의 전략:

- (1) $n = 1$ 에서 보수적 남한정부는 1의 확률로 압박정책을 선택한다.
- (2) $n > 1$ 에서 $P_n \geq L^{n-1}$ 일 경우 보수적 남한정부는 북한의 개혁정책에 대해 1의 확률로 포용정책을 선택한다.
- (3) $n > 1$ 에서 $P_n < L^{n-1}$ 일 경우 보수적 남한정부는 북한의 개혁정책에 대해 a_n 과 $1-a_n$ 의 확률로 포용과 압박정책을 혼합한다.

진보적 남한정부의 전략:

- (1) 북한이 개혁정책을 취한다면 진보적 남한정부는 1의 확률로 포용정책을 선택한다.
- (2) 북한이 대결정책을 취한다면 진보적 남한정부는 1의 확률로 압박정책을 선택한다.

여기서 $r = (H^* - F^*) / (H^* - h^*)$ 로서 북한이 혼합전략을 구사할 때 개혁정책에 두는 비중이고, $a_n = (1 - L^{n-1})P_n / (1 - P_n)L^{n-1}$ 로써 $n > 1$ 의 期에 $P_n < L^{n-1}$ 이러므로 보수적 남한정부가 혼합전략을 구사할 때 포용정책에 두는 비중이다.

2. 順次均衡의 例

이 순차균형의 성격을 이해하기 위해 구체적인 예를 하나 분석해 보자. 우선, 게임 시작 전에 북한이 남한정부가 진보적이라고 믿는

확률을 0.05라고 하자. 또 남북협력관계를 구축할 경우 북한은 $E = 3$ 의 수익을 기대할 수 있고, 남북대치상태를 유지할 경우에는 $g = 2$, 남한에 배신당할 경우에는 $G = 1$ 을 얻는다고 하자. 그러면 $L = 0.5$ 이며, $0.05 < 0.5^n$ 의 조건은 $n = 5$ 부터 성립된다.

T 가 5보다 크다고 가정하면, T 期 게임의 시발점에서 북한이 남한에 대해 가지고 있는 초기 신뢰도 $P_T = 0.05$ 는 L^T 보다 크므로, 북한은 개혁정책을 추진하게 되고 남한정부는 성향에 관계없이 포용정책으로 화답한다. 그렇게 되면 북한은 남한정부의 성향을 파악할 수 없으므로 $P_{T+1} = P_T$ 가 되고, 남북협력은 $n = 6$ 시점에 이를 때까지 지속된다. 이 시점부터 보수적 성향의 남한정부는 진보파처럼 행동해서 얻을 수 있는 수익이 상대적으로 감소하였으므로 ($0.05 < 0.5^{n-1}$, $n = 6$) 혼합전략을 구사하게 된다. 만약 남한정부가 $n = 6$ 에서 압박정책을 추진한다면, 북한은 남한정부가 보수적이라는 것을 알아차리게 되므로 $P_5 = 0$ 이 되고, 반대로 남한정부가 포용정책을 추진한다면 $P_5 = 0.5^5$ 이 된다. 남한정부가 압박정책을 추진할 가능성에 대응해 $n = 5$ 부터는 북한도 혼합전략을 구사하게 된다. 이제 보수파가 진보적으로 행동할 확률은 1이 아니므로, 북한의 개혁정책에 대해 남한정부가 포용정책으로 화답하면 $P_4 = 0.5^4$, $P_3 = 0.5^3 \dots$ 식으로 신뢰는 증진된다. 하지만 북한은 $n = 1$ 시점에서 $P_1 = 0.5$ 이상으로 남한정부의 성향에 대해 확신을 하지 못한다. 비록 초기 신뢰도인 0.05에 비해서는 신뢰도가 10배로 증진되었으나 북한은 게임이 끝나고서야 비로소 남한정부의 성향을 확실히 파악하게 되는 것이다.

3. 信賴의 役割

T 期 게임의 마지막 期인 $n = 1$ 을 좀더 자세히 분석해 보자. 예

를 들어, $n=1$ 은 민주선거를 통해 선출된 남한정부의 임기가 거의 종료된 시점이라고 봐도 될 것이다. 이 경우 남한정부는 진보적이라는 명성을 쌓아서 수익을 올릴 수 있는 기간이 더 이상 남아 있지 않으므로 일회게임 상황에서처럼 행동하게 된다. 따라서 보수적 성향의 남한정부는 북한의 정책선택에 구애받지 않고 압박정책을 추진하게 된다. 이와는 대조적으로 진보적 성향의 남한정부는 $n=1$ 에서도 다른 期와 마찬가지로 북한이 개혁정책을 취한다면 포용정책으로, 대결정책을 선택한다면 압박정책으로 대응하게 된다.

북한의 정책선택은 남한에 대한 신뢰도의 크기에 따라 달라진다. 마지막 期가 시작되기 전까지 남한이 북한의 개혁정책에 대해 일관적으로 포용정책으로 대응해 왔을 때 북한이 남한을 진보적이라고 믿는 확률을 P_1 이라고 하자. 그럴 경우 북한이 마지막 期에 개혁정책을 선택함으로써 취할 수 있는 수익은 $p_1E + (1-p_1)G$ 이고, 대결정책을 추진할 때 얻는 수익은 g 이다. 따라서 $L = (g-G)/(E-G)$ 은 북한의 정책선택을 좌우하는 일종의 분기점 역할을 한다. 즉 북한은 $p_1 > L$ 일 경우 개혁정책을, $p_1 < L$ 일 경우에는 대결정책을, 그리고 $p_1 = L$ 일 경우에는 개혁과 대결정책을 혼합하게 되는 것이다.

결국 북한이 남한에 대해 일정한 크기의 신뢰도를 가지고 있을 경우, 남북협력을 통해 얻을 수 있는 상대적 수익($E-G$)이 크면 클수록, 또 남한의 배신에 의해 입는 상대적 손해($g-G$)가 작으면 작을수록 북한이 개혁정책을 추진할 확률이 높아지게 되는 것이다. 북한과 협력관계를 구축하기 위해서 남한이 할 수 있는 일은, 북한이 개혁정책을 취할 경우 기대할 수 있는 혜택을 늘리고 남한이 압박정책으로 대응할 때 북한이 입는 피해를 줄여 북한이 거의 '밋져야 본전'이라는 식으로 개혁정책을 펼 수 있는 대외적 여건을 조성해 주는 것이다.

북한의 경우, 북한이 남한에 대해 가지고 있는 초기 신뢰도가 어떻게 형성되는지는 알 수 없지만, 그 정도가 높으면 높을수록 남북

협력관계가 구축되고 유지될 확률이 높아지므로, 속는 셈치고 남한을 믿어보는 것이 바람직하다.

IV. 討議 및 政策的 示唆點

이 논문에서는 순차균형 개념을 사용하여 상당히 작은 규모의 신뢰도만 있으면 유한반복되는 순차적 게임에서도 피의자의 딜레마에서 벗어날 수 있음을 입증했다.

1. 상대방 행동에 대한 解釋과 分類

이 모델의 가장 본질적인 문제는 상대방의 궁극적 의도에 대한 불신에서 파생되는 상대방 행동에 대한 해석과 분류의 문제이다. 현재의 남북관계하에서 남한은 북한이 체제전환을 통해 국력을 축적한 뒤 무력도발을 자행하지 않을까 하는 의심을 떨쳐버릴 수 없을 것이고, 북한은 북한대로 남한이 대북지원을 통해 북한을 남한에 의존하게 한 후 집어삼키지 않을까 하는 불안을 떨쳐버릴 수 없을 것이다. 그렇다면, 예를 들어 북한의 나진-선봉지구 개발은 개혁정책의 일환으로 분류해야 하는가, 아니면 궁극적인 대결정책의 시발점으로 해석해야 하는가?

의도와 결과가 별개의 것이 될 수도 있다는 사실은, 상대방에 대한 신뢰 그 자체와도 관련된 이 문제를 더욱 복잡하게 한다.¹⁴⁾ 어떤 기준에 따라 상대방의 행동을 해석하고 분류할 것인가? 이 모델에

14) 예를 들어, 고르바초프는 사회주의를 강화하기 위해 현실 사회주의에서 일탈하는 정책을 폈지만, 변화를 통제하지는 못했고 결국은 소련의 붕괴와 체제변혁을 초래했다.

서는 각 행동의 불확실한 궁극적 의도보다는 단기적 효과에 초점을 맞춰 분석했다.

2. 政策交換 횃수와 期待收益의 조절

이 게임의 결과는 정책선택을 좌우하는 분기점인 L^T 의 상대적 크기에 따라 달라진다. 초기 신뢰도가 이 수준에 미치지 못할 경우 분쟁 당사자들은 피의자의 딜레마 상황을 극복하지 못하고 협력관계를 전혀 구축할 수 없게 된다. 하지만 $L = (E-G)/(g-G)$ 나 T 는 고정된 수치가 아니므로 분쟁 당사자들은 협력관계를 구축하기 위해 여러가지 방안을 강구해볼 수 있다.

우선, T 는 한정된 기간내에 정책선택과 대응이 오가는 횃수이지 고정된 숫자가 아니므로 분쟁 당사자들은 정책교환에 소요되는 평균시간을 줄임으로써 협력관계 구축의 가능성을 높일 수 있다. 예를 들어, 북한이 어떤 제안을 해올 경우 남한정부는 되도록 빨리 이에 대응하여 T 를 높임으로써 남북협력관계 구축을 위해 북한의 '초기 신뢰도가 넘어야 할 문턱' L^T 을 낮추는 효과를 가져올 수 있는 것이다.

또 분쟁 당사자들은 협력을 통해 얻을 수 있는 수익을 늘리고 배신을 당하는 데에서 입는 손해를 줄여 L 을 낮춤으로써 협력관계 구축의 가능성을 높일 수 있다. 남북한의 경우 남한은 북한이 개혁정책을 추진할 경우 경협 등을 통해 얻을 수 있는 실질적 수익을 늘려주기 위해 초기의 시험적 공동개발 프로젝트나 임가공무역이 성공하면 대북투자를 적극 장려하겠다고 밝힐 필요가 있다. 또 나진-선봉 자유무역지대와 같은 제한된 개방실험도 과소평가하지 말고 북한이 외부로부터 배신당할 경우 입는 피해를 줄이는 조치의 일환이라고 발상의 전환을 할 필요가 있다.

그리고 비록 이 모델에서는 E 와 G 를 고정되어 있는 수치인 것처럼 간주했지만 실제 상황에서는 순차적 게임이 진행됨에 따라 그 규모를 E_n , G_n 식으로 바꿀 수 있는 정책변수인 점을 감안한다면 남한정부의 운신의 폭은 더욱더 넓어지게 된다. 예를 들어, 남북협력관계를 유지하기 위해서는 $P_n > L^n$ 이 북한이 개혁을 추진하는 데 필요한 기본조건이라는 점에 착안하여, 점진적으로 북한의 기대수익(E_n)을 높이고 기대피해(G_n)를 낮추어 북한의 '신뢰도가 넘어야 할 문턱' L^n 을 낮춰주는 정책이 효과적일 것이다.

이에 반해, 장기적 사업과 연계하지 않고 식량 등을 일과성으로 지원하는 것은 남북협력관계 유지에 그리 도움이 되지 않는 비효율적인 방법이다. 인도주의적 차원에서 북한에 쌀을 제공하더라도 수십만톤을 일시에 무상지원하는 것보다는, 이것을 수만톤 규모로 나눠 북한의 대응을 주시하면서 단계적으로 늘려 지급하고 공동 자원 개발 또는 경작사업 등과 연계하여 지원에 대한 대가를 어느 정도 확보하는 것이 북한의 경제난을 북한 스스로, 구조적으로 풀어가도록 돕는 방법이다.¹⁵⁾

또 게임 초기에 남북협력이 이뤄지는 데 필요한 초기 신뢰도 P_7 는 그리 크지 않아도 된다는 점을 감안한다면, 초기의 대규모 대북 지원은 '과잉포용'이 될 수도 있다. 그보다는 오히려 임가공무역의 확대나 시험적 프로젝트의 개시 등 그리 수익이 높지 않은 경협사업을 초기에 추진하면서, 북한이 지속적으로 개혁개방을 확대할 경우에는 남한 및 외국자본의 직접투자 유치 등 기대수익이 큰 사업

15) 이런 관점에서 볼 때 한반도 에너지 개발기구(KEDO)는 상당히 효과적인 남북협력 형태이다. 우선, 남한은 미국 등 우방국과 함께 북한의 핵개발 프로그램 철폐를 경수로 및 중유 제공과 장기적으로 연계하여 남북한 쌍방이 실질적 수익을 거둘 수 있도록 했으며, 북한도 북한 나름대로 KEDO를 국제기구화함으로써 남한이 일방적으로 사업을 도중에 중단할 가능성을 낮췄다. 이에 반해 1995년에 이뤄진 15만톤의 대북 쌀 지원은, 북한의 지속적인 개혁을 촉진할 아무 장치도 마련하지 않은 비효율적인 경협 형태라고 할 수 있을 것이다.

도 실현될 수 있을 것임을 약속하고 또 실제로 행동을 통해 보여주는 것이 더 효과적인 방법일 것이다.

북한의 관점에서 볼 때에도 일단 실질적인 투자사업이 진척되어 남한기업이 북한에서 공장 등을 운영하게 되면 북한이 남한에 배신당할 가능성도 줄어든 것이다. 그렇게 되면 남북간의 신뢰가 증진되지 않더라도 북한의 신뢰도가 넘어야 할 문턱이 상대적으로 낮아지므로 남북협력관계가 오래 지속될 수 있게 된다.

3. 初期 信賴度: 형성과 전파과정에 관한 의문

북한이 개혁을 추진하는 데 필요한 기본조건이 $P_n > L^n$ 이라면 P_n 을 높여주는 방법은 없는가? 이 게임의 순차균형 개념을 유지할 경우 북한의 남한에 대한 신뢰도 P_n 은 Bayes Rule에 따라 개정되어야 하며 게임 도중에 북한이 임의로 P_n 을 설정할 수는 없다. 하지만 북한의 초기 신뢰도 P_T 의 경우에는 Bayes Rule에 따라 결정되는 것이 아니므로, 남한의 L^T 처럼 정책변수로 운용할 여지가 있어 보인다. 과연 북한은 남한에 대한 초기 신뢰도를 높게 설정함으로써 남북협력관계 구축에 기여할 수 있을까?¹⁶⁾

초기의 주관적 판단을 내생적 변수로 간주할 경우 이 모델은 예상치 못한 결과를 낳는다. 북한이 남한에 배신당해 입는 피해는 한기에 국한되므로, 만약 게임이 충분히 반복된다면 북한은 남한에 대한 초기 신뢰도를 높게 설정하는 것이 더 유리하게 되는 것이다. 위에 제시하였던 순차균형의 예에서 만약 $P_T > 0.5$, 또는 극단적으로 $P_T = 1$ 이라면 북한정부는 마지막까지 개혁정책을 추진하고, 남

16) 게임이론에서 초기의 주관적 판단(prior belief)은 보통 외생적 변수로 간주된다. Kreps-Wilson(1982) 모델에서도 초기의 주관적 판단은 '주어진 것'으로 간주되며 이것이 신규진입 게임의 전체적 수익구조와 연관되어 내생적으로 결정될 수 있다는 언급은 없다.

한정부는 성향에 관계없이 게임의 마지막 期가 오기 전까지는 포용 정책으로 화답하게 되므로 $P_T = 0.05$ 보다 더 파레토 효율적인 결과를 낳게 된다.¹⁷⁾

이 모델에 따르면 한 방향으로 불확실성(one-sided uncertainty)이 존재할 경우 ‘무조건 한번 믿어보는 것’이 분쟁해소와 협력관계 구축에 효과적인 정책이라고 할 수 있을 것이다.¹⁸⁾ 하지만 남북한간의 신뢰도는 극히 낮은 것이 현실이다. 높은 수준의 초기 신뢰도가 남북관계 진전에 크게 기여할 수 있다는 것을 북한이 안다면 왜 남한을 신뢰하지 않는 것인가? 남한에 새 정부가 들어설 경우, 북한은 미래지향적인 관점에서 차후 5년간의 기대수익을 계산해 보지 않고 단지 과거의 남북관계에 근거하여 초기 신뢰도를 낮게 설정하는 것일까?

초기 신뢰도의 형성에 관한 이와 같은 질문들은 기존의 게임이론에서 그리 크게 다루지 않고 있는 것이 사실이며, 경제학보다는 오히려 심리학적 관점에서 접근하는 것이 문제를 해결하는 데 더 적절한 방법일 수도 있다. 하지만 이 논문의 모델 구도는 게임이론적 관점에서도 초기 신뢰도의 형성 문제에 대해 어느 정도 접근할 수 있음을 보여준다. 우선, 과연 남북관계라는 게임이 5년을 단위로 남한의 정부가 바뀔 때마다 無의 상태에서 다시 시작되는 것인지 검토해볼 필요가 있다. 물론 5년마다 바뀌는 남한정부의 연계성을 지

17) 만약 남한정부가 그 이전에 압박정책을 추진할 경우 북한은 대결정책으로 전환해야 될 것이지만, 남한정부의 이와 같은 행동은 순차균형에서 이탈한 것이다. 이 경우 북한은 보수적 남한정부와 상대하고 있음을 알아차리고 그 다음 시기부터 대결정책으로 전환하게 되는데, (대결, 압박)이 반복되는 이 구도하에서 보수적 남한정부가 얻을 수 있는 수익은 진보적으로 행동할 때의 수익보다 낮으므로 수익극대화의 원칙에 위배된다.

18) 모델의 구도를 바꾸어 남한이 북한의 성향을 모른다고 해도 이 모델의 결론들은 유효하다. 서로 상대방의 성향을 모를 경우(two-sided uncertainty)에는 각자 optimal stopping rule에 따라 행동하게 된다. 이에 대해서는 Kreps-Wilson(1982)를 참조할 것.

나치게 강조하여, 기존의 여당이 정권 재창출에 성공할 경우 대북 정책이 그대로 유지된다고 가정하는 것은 이 모델의 유한반복 구도에도 배치되는 논리의 비약이다. 오히려 여기서 주목하고자 하는 것은 대통령 후보들이 선거유세 등을 통해 자신의 對北觀을 집권 전에 이미 어느 정도 노출할 수밖에 없다는 사실이다. 따라서 북한은 새 남한정부의 성향에 대해 어느 정도 주관적인 판단을 할 수 있을 것이다.¹⁹⁾ 만약 무슨 이유에서든지 초기 신뢰도 형성에 있어 상대방의 과거 언행에 기초한 정보가 미래 기대수익에 대한 고려보다 훨씬 더 중요한 역할을 한다면, 초기 신뢰도의 유래에 대한 의문도 어느 정도는 해소될 수 있을 것이다.

초기 신뢰도에 대한 또 하나의 의문은 그 전과과정에 대한 것이다. 북한의 초기 신뢰도가 아무리 높다고 해도, 보수적 남한정부가 낮은 것으로 착각해서 게임의 종료시기가 다가오기 훨씬 전부터 압박정책을 구사하기 시작하면 게임은 (대결, 압박) 구도로 전환되고 만다. 이런 상황이 벌어지지 않도록 남북한 사이에 의사소통을 원활히 하여 초기 신뢰도를 일종의 공유지식(common knowledge)으로 만드는 방법이 있는가?

사실 북한이 남한에 대해 높은 초기 신뢰도를 가지고 있다고 선언한 후 개혁정책을 추진할 때 남한정부에게 있어 최선의 방책은 북한을 한번 믿어보는 것이다. 만약 북한의 실제 초기 신뢰도가 낮

19) 일단 게임이 시작된 후 북한의 개혁정책에 대해 남한이 압박정책으로 대응한다면 북한은 남한이 보수적 성향을 지녔음을 알게 되고 대결정책으로 전환하게 된다. 그 이후 남한이 이제부터는 포용정책을 추진하겠다고 선언해도 북한은 믿지 않을 것이다. 만약 게임이 시작되기 전에, 예를 들어 대통령 선거유세 기간중에 북한에 대해 강경발언을 한 후보가 당선된 후 진보적 노선을 표방한다면 북한은 새 남한정부의 성향을 어떻게 파악해야 하는가? 이 모델구도 하에서 성향이 100% 노출되는 것은 북한의 개혁정책에 대하여 남한이 압박정책으로 대응할 때뿐이다. 그 외의 경우에는 불확실성(uncertainty 또는 noise)이 있게 되어 있으므로 이와 같이 선거 前과 後의 행동이 다른 새 남한정부에 대한 북한의 초기 신뢰도도 0은 아닌 것으로 간주해야 할 것 같다.

아서 남한이 예상했던 것보다 일찍 대결정책으로 전환한다고 해도 남한은 언제나 적절히 대응할 수 있다. 왜냐하면 이 모델구도하에서는 북한이 언제나 먼저 움직이기 때문이다. 흥미롭게도 초기 신뢰도의 형성과정과 마찬가지로 전과과정에서도 ‘한번 믿어보는 것’이 일리있는 정책일 수도 있는 것이다.

4. 信賴의 增進: 末期的 現象?

이 모델에 따르면 신뢰증진은 게임의 말기에서나 가능한데 과연 현실적으로 신뢰증진이 ‘말기적 현상’인가? 사람들은 흔히 신뢰는 인연을 처음 맺게 된 이후부터 점진적으로 늘려갈 수 있는 것으로 믿는데 이것은 잘못된 생각인가?

이 모델에서 신뢰는 상대방이 기회주의적인 행동을 할 수 있음에도 불구하고 그와 같은 행동을 하지 않을 때 비로소 증진되는 것으로 정형화되어 있다. 이에 비해 상대방이 기회주의적인 행동을 할 여지가 있었는지 생각해 보지도 않고 그가 단순히 협조적인 행동을 계속해 왔다고 해서 그 사람을 믿는 것은 ‘신뢰’보다는 ‘맹신’에 가깝다고 할 수 있다.

V. 結 論

이 논문에서는 남북협력관계 구축에서 신뢰가 할 수 있는 역할에 대해 분석함에 있어, 남한이 양보하면 북한도 ‘감지덕지’해서 보답을 하고 순화될 것이라든지, 남북협력을 통해 상호신뢰가 증진될 것이라든지 하는 희망적인 가정을 도입하지 않고 게임이론의 논리

에 따라 북한의 정책선택과 남한의 정책대응을 검토해 보았다. 이 논문의 모델은 Kreps-Wilson(1982)의 신규진입 게임을 협력관계 구축이라는 새로운 상황에 맞게 변환시킨 순차적 게임모델로서 유한반복 게임구도와 비대칭적 정보소유를 그 근간으로 한다.

이 논문의 모델에서처럼 북한이 남한정부가 진보적이라고 믿는 확률을 남한에 대한 '신뢰'라고 정의할 때, 신뢰는 크게 나누어 다음과 같은 역할을 한다. 북한이 남한을 전혀 신뢰하지 않는다면, 남북간의 대치상태는 결코 해소될 수 없다. 하지만 남북관계가 진전되는 데 상당히 높은 수준의 신뢰가 요구되는 것은 아니다. 本文의 예를 통해 본 것처럼 북한의 남한에 대한 초기 신뢰도가 0.05만 되어도 남북협력관계는 상당기간동안 지속될 수 있다. 보수적 성향의 남한정부도 대북포용정책을 통해 얻을 수 있는 수익이 적지 않다는 점을 감안하여 상당기간동안 진보적으로 행동할 동기가 있기 때문이다. 하지만 이와 같이 흉내내기 가능성이 다분한 동안에는 신뢰가 쉽게 증진되지 않으며, 다만 남북한이 협력을 통해 실질적인 수익을 얻을 수 있을 뿐이다.

남북협력관계를 구축하기 위해 남한이 할 수 있는 일은 정책대응에 소요되는 시간을 줄이는 한편, 개혁정책을 통해 북한이 얻을 수 있는 기대수익을 늘리고 남한에 배신당할 경우 입는 피해를 줄여주는 것이다. 특히 장기적인 관점에서 북한의 기대수익을 점진적으로 늘려 '신뢰도가 넘어야 할 문턱'을 낮추는 것이 효과적인 정책이다. 남북관계를 진전시키기 위해서는 북한의 지속적인 개혁을 조건으로 북한의 기대수익이 점진적으로 증가하는 KEDO식의 장기적 경험 사업을 고안하여 추진해야 한다. 이에 반해 인도주의적 명분하에 일과성의 식량지원을 북한에 제공하는 것은 북한의 구조적 경제난 해소나 남북관계의 지속적 진전에 별 도움이 안되는 정책이다.

북한은 북한 나름대로 남한을 한번 믿어볼 필요가 있다. 왜냐하

면 상대방에 대한 불신은 협력관계가 유지될 수 있는 기간을 감축시키기 때문이다.

이상과 같이 게임이론을 통해 남북협력관계 구축문제를 분석해보아도 해소되지 않는 가장 큰 문제는 상대방의 궁극적 의도에 대한 불신이다. 즉 남한은 북한이 대남적화전략을 포기하지 않은 채 개혁정책을 통해 국력을 축적한 뒤 무력도발을 자행하지 않을까 하는 의심을 떨쳐버릴 수 없을 것이고,²⁰⁾ 북한은 북한대로 남한이 포용정책을 통해 북한을 남한에 의존하게 한 후 흡수통일하지 않을까 하는 불안을 떨쳐버릴 수 없을 것이다.

하지만 상대방의 모든 정책을 음모적인 시각에서 해석하는 것은 비생산적인 결과를 초래할 뿐이다. 만약 북한이 개혁정책을 통해 군사력을 강화한다고 해도 남한은 이에 대응하여 군비경쟁을 벌일 국력을 충분히 가지고 있다. 남한정부의 경우, 북한 체제붕괴와 그 이후의 북한 경제재건에 필요한 준비가 충분히 갖춰져 있지 않은 상태에서 고도로 무장된 상대를 궁지로 몰아넣는 것보다는, 비록 북한의 궁극적 의도에 대해 확신할 수는 없어도 북한의 개혁정책을 지원하는 것이 더 현명한 정책일 것이다.²¹⁾ 왜냐하면 북한의 시장지향적 개혁은, 북한의 남한에 대한 적개심을 고조시키기보다는 기

20) 북한의 先개혁-後대결정책 추진 가능성에 대한 이런 우려는 중국이 개혁정책을 통해 국력을 키운 후 미국과 대결하지 않을까 염려하는 미국내 일부의 시각과 일맥상통하는 측면이 있다.

21) 제Ⅱ장에서 밝힌 바와 같이 북한이 대결정책을 고수할 때에는 압박정책이 효과적이며 보수파든 진보파든 일방적인 유화정책을 펴서는 안된다. 단, 냉전기간 중 미국이 소련에 대해 사용한 적극적 군비경쟁과 ‘힘에 의한 평화’ 정책을 한 반도에 적용하는 것이 과연 바람직한가에 대해서는 재고해 볼 필요가 있다. 왜냐하면 북한은 남한과 대치하고 있는 위협적 상대일 뿐 아니라 봉쇄정책으로 가중될 경제난 등에 의해 그 체제가 붕괴될 때는 남한이 떠맡아야 할 가능성이 높은 (잠재적) 관리의 대상이기 때문이다. 소련이 해체된 후 경제적 지원은 별로 제공하지 않고 안보차원에서 러시아의 움직임을 관찰하고 있는 미국과, 북한 체제붕괴시 당장 난민 걱정부터 해야 하는 남한은 결코 같은 처지에 있지 않다. 대북정책의 기초는 결국 일방적인 유화정책과 일방적인 압박정책이라는 양극단을 피하면서 북한의 변화를 유도해 나가는 방향으로 설정되어야 한다.

존의 ‘우리식 사회주의 체제’를 약화시켜 남북한간의 이념적 갈등을 해소하고 남북협력을 본격화하는 기반을 제공할 가능성이 더 높기 때문이다. 북한의 경우에도, 만약 남한에 대한 과다의존이 우려된다면 북한개혁에 있어서의 남한정부 및 기업의 역할을 일정선에서 제한하면 될 것이다. 이에 반해 지금처럼 남한정부를 배제하는 것은 북한이 원하는 남한기업의 대북진출을 차단하고 북-미, 북-일 관계개선과 외국자본 유치를 어렵게 만드는 비효율적인 정책이다.

결국 남북한은 원칙협상(principled negotiation) 이론에서도 제시된 것처럼 의심을 하려면 한이 없는 상대방의 의도에 초점을 맞추기보다는 정책선택과 대응이 가져올 실질적인 수익과 위험부담을 냉철하게 계산하고 이에 따라 행동해야 할 것이다.

〈附 錄〉

첫기와 마지막기로 구성되어 있는 2期模型을 상정하기로 한다 (T기 모형으로의 확장은 쉽게 가능하므로 본 연구에서는 2期の 경우를 증명할 것임). $n=2$ 는 마지막期까지는 2기가 남아 있음을, 즉 이 2期게임의 첫기를 의미한다. 북한의 남한에 대한 주관적 믿음(belief)은 Bayes Rule에 따라 改新된다고 하자.

증명에 앞서 편의상 남북한의 전략과 성향을 다음과 같이 영문약자로 표기하기로 하자.

R: Reform	(북한의 개혁정책)
NR: No Reform	(북한의 비개혁 또는 대결정책)
A: Accommodation	(남한의 포용정책)
NA: No Accommodation	(남한의 비포용 또는 압박정책)
PS: Progressive South	(진보적 남한정부)
CS: Conservative South	(보수적 남한정부)

$n=2$ 에서 北이 R을 택한 것에 대하여 南이 A로 반응하였다고 할 때 $n=1$ 에서 北이 南은 진보적(PS)이라고 믿을 확률을 p_1 으로 나타내기로 한다. 그러면,

$$\begin{aligned} p_1 &= \Pr[PS|A] \\ &= \Pr[A|PS]\Pr[PS] / \Pr[A] \end{aligned}$$

단, $\Pr[A] = \Pr[A|PS]\Pr[PS] + \Pr[A|CS]\Pr[CS]$ 이다. p_2 를 이 2期게임이 시작되는 시점에서 北이 南이 진보적이라고 믿고 있는 ‘초기

신뢰도'라고 하기로 한다. 즉 $p_2 = \Pr[PS]$ 이고 $1-p_2 = \Pr[CS]$ 이다. 南이 진보적이라면 北이 R을 선택한 것에 대해서 A로 대응할 것이므로 $\Pr[A|PS] = 1$ 이고 $\Pr[A|CS]$ 에 관해서 정리하면, 다음의 식(1)을 구하게 된다.

$$\Pr[A|CS] = p_2(1-p_1) / p_1(1-p_2) \quad (1)$$

$\Pr[A|CS] \leq 1$ 이므로 $p_2 \leq p_1$ 이어야만 한다. 즉 北의 R에 대해서 南이 A로 대응할 경우, $n = 1$ 에서 北의 南에 대한 신뢰(trust)의 정도가 $n = 2$ 에서보다 작을 수는 없다.

아래의 보조정리 1과 보조정리 2는 $n = 1$ 에서 본문에 기술되어 있는 순차균형(sequential equilibrium)에서 성립하는 사실에 대한 설명이다. 이에 대한 증명은 본문의 수익구조표를 보면 쉽게 얻을 수 있으므로 여기선 생략한다.

보조정리 1 : 게임의 $n = 1$ 에서 보수적인 남한정부(CS)는 제시된 순차균형에서 NA를 확률 1로 실행한다.

보조정리 2 : 게임의 $n = 1$ 에서 제시된 순차균형에서 북한은 다음과 같이 행동한다(단, $L = (g-G) / (E-G)$ 이라고 정의한다).

- (1) $p_1 > L$ 인 경우, 북한은 확률 1로 R를 선택한다.
- (2) $p_1 < L$ 인 경우, 북한은 확률 1로 NR을 선택한다.
- (3) $p_1 = L$ 인 경우, 북한은 R과 NR을 혼합한다.

이제 $n = 2$ 에서 본문에 기술되어 있는 순차균형에서 성립하는 사실에 대하여 설명하기로 한다.

보조정리 3 : 게임의 $n = 2$ 에서 $p_2 > L$ 인 경우, 제시된 순차균형에서 $\Pr[A|CS] = 1$ 이고 $p_1 = p_2$ 이다.

증명 : 이것이 사실이 아니라고 가정해 보자. 즉 $\Pr[A|CS] < 1$ 이라고 하자. 식(1)에 의하면 $\Pr[A|CS] < 1$ 은 $p_2 < p_1$ 임을 의미한다. $p_2 > L$ 의 경우를 상정하고 있으므로 이미 얻어진 결과에 의하여 $p_1 > L$ 이어야 한다. 北의 신뢰도가 충분히 크므로 보조정리2에 의하여 北은 $n=1$ 에서 R을 실행한다. 보수적인 南이 $n=2$ 에서 A를, 그리고 $n=1$ 에서 NA를 실행함으로써 얻게 되는 총수익(total payoff)은 $F^* + H^*$ 가 된다. 반면, 보수적인 南이 $n=2$ 에서 NA를 실행하여 자신의 실체(true identity)를 드러낼 경우 얻게 되는 총수익은 $H^* + h^*$ 가 된다. $F^* > h^*$ 이므로 보수적인 南이 $n=2$ 에서 A를 실행하는 것은 NA를 실행하는 것보다 우월한(strictly better) 전략이며 이는 이 경우 $\Pr[A|CS] = 1$ 임을 의미한다. 이는 모순이다. (증명 끝)

이 경우(즉 $p_2 > L$ 의 경우)는 北의 초기 신뢰도가 높은 상황을 나타내고 있다. 첫期에서 진보적인 유형을 모방함으로써 보수적인 南은 자신의 실체를 숨기고 초기의 높은 신뢰도가 마지막期까지 유지되도록 하려 한다. 결국, 南은 마지막期에 가서야 비로소 두 정부간의 협력을 파기하게 된다. 마지막期가 오기 전에 선불리 협력을 파기하는 일은 보수적인 南으로서는 최적일 수 없다. 그 이유는 그런 행위는 北으로 하여금 자신의 南에 대한 주관적 믿음(belief)을 改新(update)하게 하여 마지막期에서 대결적인 자세로 회귀하게 할 것이기 때문이다.

보조정리 4 : 게임의 $n=2$ 에서 $p_2 < L$ 인 경우, 제시된 순차균형에서 $\Pr[A|CS] \in (0,1)$ 이고 $p_1 = p_2$ 이다.

증명 : $\Pr[A|CS] = 0$ 이라고 가정해 보자. 그러면 식(1)로부터 $p_1 = 1$ 이 된다. $p_1 = 1 > L$ 이 되므로 보조정리2에 의해서 北은 $n=1$ 에서 확률 1로 R을 실행한다. 보수적인 南이 $n=2$ 에서 A를 선

택함으로써 $F^* + H^*$ 의 총수익을 얻게 된다. 반면, 보수적인 南은 NA를 실행할 경우 $H^* + h^*$ 의 총수익을 얻게 된다. $F^* > h^*$ 이므로 $n = 2$ 에서 NA를 실행하는 것은 보수적인 南에게 최적일 수 없으며 따라서 $\Pr[A|CS] = 0$ 은 사실일 수 없다.

이번에는 $\Pr[A|CS] = 1$ 이라고 가정해 보자. 이 경우에는 식(1)에 의해서 $p_1 = p_2$ 가 성립한다. $p_2 < L$ 이므로 $p_1 < L$ 이 사실이며, 보조정리2에 의하여 $n = 1$ 에서 北은 확률 1로 NR을 실행한다. 이럴 경우, 보수적인 南은 $F^* + h^*$ 의 총수익을 얻게 되는데, 이것은 $n = 2$ 에서 NA를 실행함으로써 얻을 수 있는 총수익 $H^* + h^*$ 보다 낮다. 말하자면, 이것은 최적행위일 수 없다. 그러므로 보수적인 南은 두가지의 순수전략(pure strategy)을 혼합하게 되며 이는 곧 $\Pr[A|CS] \in (0,1)$ 과 $p_1 \geq L$ 을 의미한다.

이제 $p_2 < L$ 이면 $p_1 = 1$ 이어야 함을 증명하기로 한다. 그렇지 않다고 가정해 보자. 즉 $p_1 > L$ 이라고 가정해 보자. 보조정리2에 의하여 北은 $n = 1$ 에서 확률 1로 R을 선택한다. 어떤 유형이든 간에 南은 $n = 2$ 에서 A를 확률 1로 선택한다. 이것은 모순이다(南이 유형에 관계없이 $n = 2$ 에서 A를 실행하는 경우, 식(1)에 의하여 $p_1 = p_2$ 가 성립한다. 그러면 $p_1 > L$ 은 $p_2 < L$ 에 모순이 된다. 그러므로 $p_1 = L$ 이어야만 한다). 그러므로 식(1)에 의해 $p_1 = p_2$ 이어야만 한다. (증명 끝)

식 (1)에 의하여 $\Pr[A|CS] = p_2(1-p_1) / p_1(1-p_2)$ 이고 $p_1 = L$ 이므로 보수적인 南은 A에 $a = \Pr[A|CS] = p_2(1-L) / L(1-p_2)$ 의 확률을 부여하게 된다.

이것은 北의 초기 신뢰도가 낮은 경우($p_2 < L$)이다. $n = 2$ 에서 두 유형이 구별되지 않도록 행동하는 것은 이처럼 낮은 수준의 초기 신뢰도가 마지막期까지 유지되도록 하는 것일 뿐이기 때문에, $n = 2$

에서 南이 포용적인 행동(A)을 하여도 北에게 긍정적인 영향을 줄 수 없으며 따라서 마지막期에서의 개혁(R)을 유도해낼 수도 없다. 달리 표현하면, 보수적인 南이 진보적인 듯이 행동하는 일(즉 $n=2$ 에서 $\Pr[A|CS]=1$ 인 경우)은 신뢰를 증진시키지 못한다. 보수적인 南이 그 실체를 드러내는 일(즉 $n=2$ 에서 $\Pr[A|CS]=0$ 인 경우) 또한 최적이지 않다. 최적전략은 두 순수전략의 확률적 혼합(randomization)을 통해 불확실성의 여지를 남겨두는 것이 된다.

보조정리 5 : $p_2 < L$ 라 하자. 제시된 순차균형(sequential equilibrium)에서 북은 $n=1$ 에서 R에 r 의 확률을, 그리고 NR에 $1-r$ 의 확률을 부여한다(단, $r = (H^* - F^*) / (H^* - h^*)$ 이다).

증명 : 보조정리4에 의하면 $p_2 < L$ 인 경우 $p_1 = L$ 이다. 보조정리2에 따르면, 이럴 경우 북은 $n=1$ 에서 R과 NR에 관해 차별을 두지 않는다. r 을, $n=2$ 에서 北이 R을 실행했을 때 南이 A로 대응한 경우, $n=1$ 에서 北이 R에 부여하는 확률이라고 하자. 보조정리4에 의하여 $n=2$ 에서 $\Pr[A|CS] \in (0,1)$ 이므로 南은 $n=2$ 에서 A와 NA를 혼합할 것이며 따라서 p_1 은 南이 어떤 순수전략을 실행하더라도 南에게 동일한 수익을 발생하게 한다. 즉 다음과 같은 관계가 성립한다.

$$F^* + rH^* + (1-r)h^* = H^* + h^*$$

이것을 r 에 관해서 정리하면 北이 R에 부여하는 확률이 유일하게 결정된다. (증명 끝)

보조정리 6 : 제시된 순차균형에서는 게임의 $n=2$ 에 대하여 다음이 사실이다.

- (1) $p_2 > L^2$ 인 경우, 북은 확률 1로 R을 실행한다.

- (2) $p_2 < L^2$ 인 경우, 복은 확률 1로 NR을 실행한다.
 (3) $p_2 = L^2$ 인 경우, 복은 r 과 $1-r$ 의 확률로 R과 NR를 혼합한다.

증명 : a 를 北이 $n=2$ 에서 R을 선택한 경우 南이 A를 실행할 확률이라고 하자. 그러면 a 는 다음과 같이 다시 쓸 수 있다:

$$\begin{aligned} a &= \Pr[A|PS]\Pr[PS] + \Pr[A|CS]\Pr[CS] \\ &= p_2 + (1-p_2) \Pr[A|CS] \end{aligned}$$

식(1)을 이용하면, $a = \frac{p_2}{p_1}$

$n=2$ 에서 확률 1로 R을 실행함으로써 北은

$$a[E + p_1E + (1-p_1)G] + (1-a)(G + g) \quad (2)$$

의 수익을 얻게 된다.

$p_2 > L$ 라고 가정하자. 그러면 보조정리3에 의하여 $\Pr[A|CS] = 1$ 이고 이것은 $a = 1$ 및 $p_1 = p_2 > L$ 임을 의미한다. 정의에 의하여 $p_2 > L$ 와 $p_2E + (1-p_2)G > g$ 는 동일하므로, $n=2$ 에서 R을 실행하는 것은 北에게는 NR을 실행하는 것보다 우월한(strictly better) 전략이 된다. 따라서 $p_2 > L$ 인 경우 北이 $n=2$ 에서 확률 1로 R을 실행하게 된다.

이제 $p_2 < L$ 을 가정해 보자. 보조정리4에 의하여 $p_1 = L$ 이며 따라서 $a = p_2/L$ 이 된다. 이 경우 北은 $n=1$ 에서 R과 NR에 관하여 무차별하고 마지막期에 g 의 수익을 얻게 된다. 식(2)에 따라 $a = p_2/L$ 일 때 $n=2$ 에서 R을 실행함으로써 北은

$$(p_2/L)[E + p_1E + (1-p_1)G + [1 - (p_2/L)](G + g)]$$

의 수익을 얻게 된다. 그러므로 北이 R을 실행하게 하는 p_2 의 가치는 다음의 부등식에 의해서 결정된다:

$$(p_2/L)[E + p_1E + (1-p_1)G] + [1-(p_2/L)](G + g) > g + g.$$

이것을 p_2 에 관해서 정리하면 다음의 조건을 구하게 된다.

$$p_2 > L^2$$

$p_2 < L^2$ 인 경우라면 $n=2$ 에서 R을 실행함으로써 北은 NR을 실행할 경우보다 더 큰 수익을 얻게 되며 따라서 이 경우 北은 NR을 확률 1로 실행한다.

이 보조정리의 증명을 마무리하기 위하여 마지막으로 입증해야 할 것은 $p_2 = L^2$ 인 경우 두 순수전략의 확률적 혼합(randomization)이 있다는 것이다. $p_2 = L^2$ 인 경우에는 반드시 $p_2 < L$ 이며 보조정리4에 의하면 이는 $p_1 = L$ 임을 의미하므로 北은 $n=1$ 에서 확률적 혼합을 한다. $n=2$ 의 경우 r_2 를 北이 R을 실행할 확률이라고 하고 $1-r_2$ 를 NR을 실행할 확률이라고 하자.

北의 혼합확률을 구하기 위해서는 南의 수익을 살펴봐야 하는데, 우선 게임의 구도상 北이 혼합전략을 구사하기 전에 南이 포용과 압박정책을 혼합한다는 점을 주목하자. 2期 게임의 경우 北이 $n=2$ 부터 혼합전략을 구사한다는 것은 게임이 시작하기 전인 $n=3$ 에 이미 南이 포용과 압박정책을 혼합한 것으로 상정할 수 있다.

$n=3$ 에서 南이 혼합전략을 구사한다는 것은 南의 기대수익이 포용정책을 추진하든 압박정책을 추진하든 같다는 것을 의미하므로 다음과 같은 공식이 성립된다.

$$F^* + r_2[a_2(F^* + r_1H^* + (1-r_1)h^*) + (1-a_2)(h^* + H^*)] + (1-r_2)(h^* + H^*) = H^* + 2h^*$$

南은 $n=3$ 에 이어 $n=2$ 에서도 혼합전략을 구사하므로 다음과 같은 관계도 성립된다.

$$F^* + r_1 H^* + (1 - r_1) h^* = H^* + h^*$$

아랫공식을 윗공식에 대입하면 다음과 같은 결과를 얻는다.

$$F^* + r_2(H^* + h^*) + (1 - r_2)(h^* + h^*) = H^* + 2h^*$$

이것을 정리하면 $n = 2$ 와 $n = 3$ 에서의 조건이 같음을 알 수 있다.

$$F^* + r_2 H^* + (1 - r_2) h^* = H^* + h^*$$

따라서 $r_2 = r_1 = r = (H^* - F^*) / (H^* - h^*)$

즉 복한이 혼합전략을 구사할 경우 그 혼합확률은 期에 관계없이 동일하다. (증명 끝)

論 評

牟 鍾 璘

(延世大學校)

이 논문은 남북관계에서의 상호 신뢰의 역할에 대해 게임이론을 적용하여 체계적으로 분석함으로써 북한의 정책선택을 이해하고 우리의 대응정책을 세우는 데 크게 기여할 수 있는 연구라 사료된다.

논문의 기본 가정은 첫째, 남북간의 협력은 남한정부의 성향에 따라 결정된다는 것이다. 즉 진보적인 남한정부와 북한 사이에는 (개혁, 포용)이라는 협력적인 정책 교환이 이루어질 수 있는 반면, 보수적인 남한정부와 북한은 서로 협력하지 못하고 (대결, 압박)이라는 긴장관계를 유지한다는 가정이다. 둘째, 북한은 남한정부의 성향을 정확히 알지 못한다고 가정하고 있다. 이와 같은 가정하에서의 북한정책은 남한정부에 대한 신뢰도, 즉 남한정부가 진보적일 확률에 따라 결정된다. 이 논문은 남북한간에 게임이 반복되었을 때 북한의 대남 신뢰도가 어떻게 변화하는가를 살펴봄으로써 남북관계에서의 신뢰의 역할을 설명하고 있다.

이 논문의 주요 결론은 북한이 남한에 대해 어느 정도의 신뢰감만 갖고 있어도 상당한 기간 남북협력관계가 지속될 수 있다는 사실이다. 왜냐하면, 보수적인 남한정부도 게임의 초기에는 북한과의 협력관계에서 얻는 수익을 확보하기 위해 포용정책을 선택할 동기가 있기 때문이다.

그러나 이 논문이 제시한 모델은 현재 남한정부가 직면한 딜레마를 정확하게 표현하기에는 미흡하다고 여겨진다. 특히 남한내 강경파와 온건파간의 이견이 존재하는 이유에 관한 저자들의 설명에 이

의를 제기하고자 한다. 저자들의 모델에서는 남한 강경파와 온건파 간의 차이가 (개혁, 압박), (개혁, 포용) 중 어느 쪽을 선호하느냐에 따라 결정된다고 가정했지만, 현실적으로 남한에서 북한의 개혁에 압박정책으로 대응하는 (개혁, 압박)을 포용정책으로 대응하는 (개혁, 포용)보다 더 선호하는 그룹이 정책결정그룹내에 존재한다고 보기 어려울 뿐만 아니라 설사 존재한다 해도 그들이 남한 강경파들의 입장을 대표한다고 볼 수 없다.

따라서 현재 남북내 강경파와 온건파의 갈등을 다른 각도에서 살펴보아야 한다고 생각한다. 그러기 위해서는 저자들의 가정을 일부 수정할 필요가 있다. 물론 저자들이 제시한 대로 남한의 정책선택은 압박과 포용정책으로 압축할 수 있다. 쌀 지원의 예를 들면, 포용정책은 쌀을 지원하는 것이며 압박정책은 쌀 지원을 거부하는 것이다. 그러나 북한의 정책선택이 개혁과 대결의 두 정책으로 한정되어 있다는 가정은 수정될 필요가 있다. 남한이 현재 인식하고 있는 북한의 선택은 세가지로 요약할 수 있다. 첫째는 ‘개혁’정책이다. 한마디로 북한이 대남 강경자세를 자제하고 남한이 제의한 4자 회담 등의 긴장완화정책을 수용하는 것이다. 둘째, 북한은 쌀 지원을 수용하면서도 4자회담 등 정치·군사적 협력조치를 취하지 않는 일종의 ‘대결’정책을 선택할 수 있다. 마지막으로 북한은 4자회담을 거부하고 오히려 쌀 지원, 남북경협 등으로부터 얻어지는 경제적 수익을 군비확충 등 대남적화통일전략에 사용하는 ‘배반’정책을 선택할 수 있다.

현재 남북한 상황으로 볼 때 남한이 먼저 정책을 결정해야 한다고 가정하는 것이 가장 현실적이라고 생각한다. 남한이 먼저 압박정책을 선택하면 그 결과는 현상유지(SQ)이다. 이 경우 북한은 SQ를 인정할 수밖에 없다. 반면, 남한이 포용정책을 선택하면 북한의 선택폭은 넓어지는데, 이는 다시 말해 개혁, 대결, 또는 배반으

로 대응할 수 있다고 가정하는 것이다. 북한의 대응에 따라 나타나는 게임의 결과는 앞으로 각각 LK(linkage), AD(aid), BT(betrayal)로 표기한다.

이들 네가지 결과 중 남한정부는 남한이 쌀 지원을 하고 북한이 4자회담을 수락하는 LK를 가장 선호하며, 남한의 쌀 지원을 북한이 군수용으로 전용하는 BT를 최악의 결과로 간주한다. 남한이 쌀 지원을 거부하는 SQ와 남한이 쌀 지원을 하고 북한이 4자회담을 거부하는 AD 중 어느 쪽을 선호하는지는 확실치 않지만 강경파들은 이 두가지 가운데 SQ를 선호하고 온건파들은 AD를 선호한다고 볼 수 있다. 즉 강경파의 수익순위는 $LK > SQ > AD > BT$ 인 것과는 달리, 온건파의 수익순위는 $LK > AD > SQ > BT$ 가 된다.

여기서 강경파와 온건파의 이견은 그들이 북한의 현재 상황을 다르게 평가하는 데서 기인한다. 온건파가 SQ를 AD보다 선호하는 이유는, 북한이 쌀 지원을 받고 4자회담 참가라는 정치적인 양보를 하지 않아도 그로 인해 북한정권이 안정되면 우발적인 도발을 하지 않을 것이라고 믿기 때문이다. 강경파의 경우는 SQ가 지속되어도 전쟁의 위험이 크게 늘지 않는다고 믿고 있고, 그렇다면 북한이 상응하는 정치적 양보를 하지 않는데 굳이 쌀을 지원할 필요가 없다고 생각하는 것이다. 현재 한국정부의 입장은 강경파에 가까우며 미국정부의 대북정책은 온건파의 논리에 기초한다고 할 수 있다. 여기서 주목해야 할 점은 강경파나 온건파 모두 LK를 제일 선호하며 BT를 가장 기피한다는 사실이다.

강경파와 온건파로 나누어지는 남한정부의 성향은 이같은 수익순위의 차이뿐만 아니라 북한정부의 성향에 대한 의견의 차이에서 기인할 수도 있다. 남한정부가 일치된 수익순위를 갖고 있어도 북한의 성향에 따라 정책선택이 달라질 수 있기 때문이다. 예를 들어 남한정부가 AD보다는 SQ를 선호하는 강경파의 수익순위를 갖고 있

다고 가정해 보자. 북한이 LK를 가장 선호한다면 남한정부는 쌀 지원을 지지하고 선택할 것이며, 만약 북한이 LK보다는 AD나 BT를 가장 선호한다면 남한정부는 쌀 지원을 거부할 것이다. 이 경우 북한이 전자의 성향을 갖고 있다고 믿는 남한의 그룹은 온건파가 되는 것이고 후자의 경우는 강경파로 인식이 된다. 남한정부의 정책이 북한정부의 성향에 따라 달라지는 것은 남한정부가 온건파 수익순위를 갖고 있어도 마찬가지이다. 북한이 가장 선호하는 결과가 BT라면 온건파 성향의 남한정부도 쌀 지원을 포기하게 된다. 결론적으로 말해, 남한내에서 대북 강경파와 온건파가 공존하는 이유는 남한의 정책결정그룹들이 북한정부의 성향과 현재 북한상황을 서로 다르게 평가하기 때문이다.

마지막으로 남북관계에 피의자의 딜레마 게임과 같은 협력이론을 적용하는 것이 적합하지 않다고 지적하고 싶다. 남북관계는 궁극적으로 양측의 협상력에 따라 결정되는데, 협력이론은 이같은 국력 또는 협상력의 차이점이 협상결과에 미치는 영향을 설명할 수 없다. 협력이론은 양측 모두가 합의된 협정을 위반할 동기를 갖고 있을 경우에 적합한 모델이다. 하지만 현재 남북은 협력 자체에 합의하지 못하는 상황이다. 물론 그 이유는 상대방의 준수의도를 불신해서가 아니라 양측 모두 서로에게 유리한 협상결과를 얻기 위해 줄다리기를 하고 있기 때문이다. 따라서 이 상황에서 적절한 모델은 협상모델이라고 생각한다.

朴 珍 佑

(國民大學校)

이 논문은 남북한 관계에서 협력을 구축함에 있어 신뢰가 갖는 중요성을 짚임새 있게 다루고 있다. 특히 분쟁해소와 협력관계 구

측에서 신뢰의 역할에 대한 게임이론을 사용하여 이론적인 접근을 시도하고 있다는 점을 높이 평가할 수 있다. 이 논문이 기여하는 바를 좀더 구체적으로 열거하면, 첫째 남북관계에서 신뢰가 남·북한 공동의 이익을 증진시키는 방향으로 작용될 수 있음을 지적한 점, 둘째 Kreps & Wilson(1982)의 게임이론적 모델을 저자들이 주장하고자 하는 바와 적절히 연결시켰다는 점, 셋째 IV장 토의 부분에서 다루고 있듯이 신뢰를 회복하기 위해 개선되어야 할 사항들을 제시하고 있다는 점을 들 수 있을 것이다.

이 논문이 기여하고 있는 바가 여러 부문에 걸쳐 상당히 많지만 평자의 생각에 그중 가장 중요하다고 판단되는 부분은 IV장 토의이다. 즉 논문에서 가장 큰 기여가 될 수 있는 부분은 믿을 만한 협조 메커니즘을 구축할 수 있는 방안에 대한 연구일 것이다. 논문의 IV장에서 일부 다루고 있으나 이 부분을 보다 구체적으로 자세히 다룰 필요가 있다. 모델에서 시사하는 바가 신뢰감이 조금만 있어도 협력이 유도될 수 있다는 점인 반면에, 현실적으로 남북관계가 협력적인 관계를 유지하지 못하고 있다는 사실(검토자의 주관적인 판단일 수 있음)은 모델이 현실의 여러 변수들을 다 고려하지 못하고 있거나, 혹은 모델이 시사하는 바와 같이 신뢰가 중요하다면 논문의 주된 초점이 이를 회복하기 위한 방안들에 있을 필요가 있다. 저자들도 정책선택 및 대응에 소요되는 평균시간을 줄이는 방안, 수익구조를 조정하여 협력을 통해 얻을 수 있는 수익을 늘리고 배신을 당하는 데에서 입는 손해를 줄이는 방향으로 조절하는 방안 등을 논문의 모델내에서 이미 제시하고 있지만, 다른 방안들이 있는가 좀더 생각해 볼 필요가 있으며, IV장의 내용을 보다 강조·확장하여 이를 제시할 수 있다면 보다 유익한 결과를 낳을 것으로 생각된다.

게임이론적 시각에서 볼 때 이 논문의 모델은 저자들도 지적하고

있듯이 Kreps & Wilson 모형을 응용했기 때문에 독창적이지는 못하지만, 반면에 새로운 모형이 가질 수 있는 오류는 발견하기 어렵다는 점에서 긍정적인 측면이 있다고 판단된다. 유한반복게임의 경우에 불완전정보가 있음으로 해서 협력의 결과를 낳을 수 있다는 결과를 보여주고 있는 Kreps & Wilson의 모형은 남북한의 관계를 유한한 관계라고 이해하는 한 나름대로 남북한 관계에 시사하는 바가 많다고 생각된다. 즉 남북한 관계에서 협력을 유도하여 상호의 이익을 증진시키기 위해 신뢰가 갖는 중요성을 지적하기에는 충분하다고 생각한다.

평자는 게임이론으로 사회현상을 설명하는 데 있어서 장점과 단점이 있다고 생각한다. 게임이론을 사용함으로써 모델의 내적 적합성을 상당히 향상시킬 수 있는 반면에, 게임이론이 다소 복잡하기 때문에 이를 적용함에 있어 복잡한 사회문제를 상당히 단순화하는 과정에서 현실과 다소 괴리가 있을 수 있다고 생각한다. 이러한 일반적인 한계가 이 논문에서도 일부 나타나고 있다고 생각된다. 즉 복잡한 남북한 문제를 이 논문의 게임모형이 만족스럽게 반영하고 있는가에 대해서는 다소 의문이 있다.

모델 측면에서 보면 논문에서 다루고 있는 남북관계의 복잡한 여러 요소들을 고려하는 방향으로 모델이 개발될 여지가 있어 보인다. 첫째, 정보의 비대칭성이 일방적인 형태를 취하고 있으나 남·북한 쌍방에 불완전한 정보상태를 생각해 볼 수 있을 것이다. 둘째, 남한과 북한의 수직구조가 외생적으로 주어져 있으나 현실적으로 볼 때 남북한의 수직구조는 각 체제 안에서의 게임의 축약된 수익구조(reduced payoff)일 가능성이 높다. 예컨대, 남한정부가 진보와 보수로 나뉠 수 있듯이 북한도 강경파와 온건파로 나뉠 수 있을 것이며, 이들은 정권획득 내지는 유지를 고려하면서 남북한 관계에서 정책적 선택을 할 것으로 생각된다. 만약 그렇다면 현재 북한에

서 강경파가 집권한 상태에서 게임의 수익구조는 강경파의 수익구조를 반영하고 있을 것이며, 집권하고 있는 강경파도 남북한 관계에서 온건한 정책선택은 온건파의 정권획득을 용이하게 해줄 가능성이 있다는 점을 고려하여 정책선택을 할 수 있다는 것이다. 즉 논문에서 제시되고 있는 수익구조는 이와 같은 복잡한 현실적 상황을 축약한 수익구조를 나타내고 있을 수 있다. 셋째, 남한정부가 임기 5년이기 때문에 남북관계를 유한게임으로 모형화하고 있지만, 5년 임기 후에 새로운 정부와의 연결관계에 따라 사실상 무한관계일 가능성이 있다.

마지막으로 모델 측면에서 지적하고 싶은 점은 초기의 주관적 판단에 관한 부분이다. 평자가 이해하는 바로는 이 논문이 전달하고자 하는 여러 정책적 시사점 중의 하나가 남북관계에서 신뢰가 회복되면 협력이 증진되어 상호 이익이 증대할 것이라는 것이다. 모델 측면에서 평자가 소화하기 다소 힘든 부분은 각주 16)에서 저자들도 지적하고 있듯이 Kreps & Wilson의 모델에서 초기의 주관적 판단은 외생적으로 주어진 것으로 보는 것이 일반적인 접근방법인데, 초기의 주관적 판단이 주어진 것으로 볼 경우 이와 같이 주어진 초기판단을 게임의 참여자가 자의적으로 조정하면 협력의 결과를 낳을 수 있다는 논문의 주장은 한계가 있다고 생각된다. 초기의 주관적 판단에 대한 논의는 현실문제를 다루고자 하는 저자들의 의도와는 달리 필요 이상으로 논문을 기술적인 측면에서 너무 복잡하게 하기 때문에 자세한 논의를 피하고 싶지만 문제점이 존재한다는 사실은 지적하고 싶다. 예컨대, 초기 신뢰도 P_T 는 어디에서 유래하는 것인가에 대해 깊이 생각해 볼 필요가 있다. 게임의 참여자가 자의적으로 선택할 수 있는 신뢰도라면 논문에서 제시하고 있는 T 기 모형의 어느 중간에 베이저안 조정(Bayesian updating)을 통해 신뢰도를 조정할 필요 없이 중간시점에서 남한이 보다 진보적이라고 신

되를 하면 보다 협력적인 결과가 나오지 않는가 하는 반문을 제기해 볼 수 있을 것이다.

이와 같은 모델과 현실현상과의 괴리는 앞서 평자가 게임이론의 한계라고 지적한 바 있다. 그러나 이러한 괴리를 게임이론 자체의 한계라고 쉽게 이야기할 수도 있지만, 평자를 포함하여 게임이론을 응용하는 연구자들이 보다 현실에 가까운 모형을 개발하고자 노력할 필요성이 있음을 지적하면서 논평을 마치고자 한다.