

分配的 正義의 두 概念과 獨立性的의 公理

金 完 鎭⁽¹⁾

功利主義를 정당화하는 하사니의 定理에서 하사니가 택하고 있는 가정중에서 사회적 選好가 독립성의 公理를 만족해야 한다는 조건에 관해서 많은 논란이 있어 왔다. 그러나 지금까지의 연구에서는 구체적 例와 反例를 통해서 독립성공리의 타당성을 논할 뿐 체계적인 분석은 이루어지지 않았다. 본연구에서는 獨立性的의 公理의 타당성을 결정하는 기준으로서 分配的 正義에 관한 도덕적 관점을 제시하고 어떠한 관련을 갖는가를 분석하였다. 도덕적 관점은 기본적으로 초인적 주체의 관점과 사회구성원의 관점으로 나눌 수 있으며 전자는 공리주의의 도덕원리의 기초가 되고 후자의 관점은 契約論의 입장과 관련지을 수가 있다.

1. 序 論

하사니(J. Harsanyi) 교수는 그의 1955년 논문에서 功利主義를 정당화하는 대단히 흥미로운 수학적 定理를 증명하였다. 이 정리는 그 당시에는 경제학계의 지배적인 입장이 序數論(ordinalism)이었기 때문에 큰 주목을 받지 못했으나 그 이후 점차 후생판단에 있어서 基數的 效用의 可測性과 개인간 비교가능성의 도입이 불가피함이 밝혀짐에 따라 새롭게 평가되어 서수론의 대표지중의 하나인 사뮈엘슨이 하사니定理를 후생경제학에서의 “획기적인 업적중의 하나(one of the few quantum jumps)” [Samuelson(1974, p.1255)]로 간주하기에 이르렀다. 그러나 이 정리의 의의에 관해서 아직도 충분히 해명되지 못한 채로 남아 있는 문제가 있으며 그 중의 하나에 관해 분석함으로써 이 중요한 定理가 함축하는 의미를 이해하는 데 약간의 공헌을 하고자 하는 것이 本研究의 목적이다.

본연구에서 검토하고자 하는 문제를 구체적으로 제시하기 위하여 먼저 하사니定理의 내용을 살펴보기로 한다. 이 정리는 다음과 같은 세 公理를 가정한다.

(1) 사회를 구성하는 개인들의 選好(preference)는 기대효용이론의 諸公理를 만족한다. 따라서 개인들의 선호는 본 노이만-모르겐슈테른效用函數(이하 略해서 vNM효용함수)에 의해 대표될 수 있고, 우리는 개인들의 불확실한 상황하에서의 選擇行動을 관찰함으로써

(1) 이 研究는 서울大學校 大學發展基金의 지원으로 이루어졌음을 밝힌다.

vNM효용을 측정할 수 있다. ⁽²⁾

(2) 사회적 선택의 기초가 되는 社會的 選好도 개인의 선호와 마찬가지로 기대효용이론의 諸公理를 만족한다. 이 때 측정될 수 있는 vNM효용을 社會厚生이라 부른다.

(3) 임의의 두 사회적 代案에 대해서, 모든 개인들의 선호가 두 대안을 무차별하다고 판단한다면 그 때의 사회적 선호도 두 代案을 무차별하다고 판단해야 한다. 이 가정은 사회적 선호가 가져야 할 최소한의 바람직한 성질로서 간주되며 만장일치의 公理라 부르기로 한다.

하사니定理에 의하면, 위의 세 公理가 충족된다면 社會厚生은 개인들의 vNM효용의 線形 結合이 된다는 것이다. 즉 社會的 代案 x 의 社會厚生을 $W(x)$ 라 하고 개인들의 vNM효용 함수를 $u_i(x)$ 로 표시하면, 적당한 λ_i 들을 선택해서

$$W(x) = \lambda_1 u_1(x) + \dots + \lambda_n u_n(x)$$

가 성립한다는 것이다. ⁽³⁾ 여기에서 개인들의 효용 u_i 의 단위를 적절히 조정해서 기중치 λ_i 가 모두 1이 되도록 할 수 있으므로 위의 社會厚生函數는

$$W(x) = u_1(x) + \dots + u_n(x)$$

로 다시 쓸 수가 있다. 즉, 社會厚生은 개인의 효용의 합과 같다는 功利主義의 결론에 도달하게 된다.

이 정리가 우리의 관심을 끄는 것은, 일견 매우 보편적이고 윤리적으로 중립적인 것으로 보이는 假定들이 사실은 매우 강력한 윤리적 주장을 내포하고 있다는 사실 때문이다. 이 점에서 하사니定理은 애로우의 不可能性定理과 흥미있는 대조를 이룬다. Arrow(1953)는 마찬가지로 윤리적인 가치판단이 거의 배제된, 그렇지만 합리적이고 민주적인 의사결정을 보장하기 위한 最小限의 條件으로 간주될 수 있는 가정들을 모두 만족하는 社會厚生函數는 존재하지 않는다는 사실을 증명하였다.

그런데 많은 경제학자들에게 功利主義는 받아들여질 수 없는 것이다. 공리주의적 후생 판단에 의하면 예를 들어 A의 效用的 증가량이 B의 效用的 감소량을 상쇄하고도 남을 만큼 클 때 社會厚生이 증가하였다고 할 수 있는 것이다. 이러한 판단이 실제로 가능하기 위해서는 개인의 基數的 效용을 측정할 수 있을 뿐 아니라, 서로 다른 개인간의 效용의 증감량의 비교가 가능해야 한다. 그러나 개인들의 시장행동의 관찰을 통해 얻어지는 것은 序數的 效용뿐이라는 사실이 이미 알려지 있을 뿐 아니라 個人間 效用比較라는 것은 과학의 영역 밖에 있는 형이상학적인 문제로 간주될 수 있다. 따라서 경제학자들은 실제의 여러 사회

(2) 期待效用理論의 공리는 선호가 이완성, 원전성, 연속성, 독립성을 만족하는 것이다. 구체적 定義에 관해서는 Harsanyi(1977b)를 참조한 것.

(3) 증명은 Harsanyi(1955)를 참조한 것.

적 대안들을 평가할 때 功利主義的 基準을 적용하기 위한 충분한 자료를 입수할 수 없으므로 공리주의적 가치판단은 불가능하게 되는 것이다. 사실 厚生經濟學의 根源은 영국의 공리주의사상이라고 할 수 있으나 이러한 효용의 가측성 및 개인간 비교가능성의 문제가 후생경제학자들을 괴롭혀 왔으며 功利主義를 대신할 수 있는 대안을 찾으려는 노력이 現代厚生經濟學의 基本的 효용이 되었다. 이러한 후생경제학의 발전에 있어서 획기적인 전환점을 마련한 것이 바로 애로우의 不可能性定理이다. 애로우의 정리의 의의는 근본적으로 효용의 측정과 개인간의 비교를 배제하고서는 사회적 후생판단이 불가능함을 증명하였다는 사실에 있다. 이제 功利主義를 새롭게 평가해보아야 할 이유가 여기에 있는 것이다.

하사니의 定理은 우리가 功利主義를 평가할 때에 효용의 가측성과 개인간 비교가능성의 문제와는 또 다른 차원의 문제를 고려하지 않으면 안된다는 사실을 보여주고 있다. 즉 우리가 공리주의를 비판한다면 하사니가 제시하는 세 公理중 구체적으로 어느 것이 타당하지 않은가를 명시하여야 하는 부담을 안게 되는 것이다. 하사니의 공리중 個人的 選好가 期待效用理論의 公理를 만족하여야 한다는 가정과 만장일치의 공리에 대해서는 경제학자들 사이에 대체적으로 그 타당성이 인정되고 있다. 물론 합리적 선택이론에 있어서 기대효용이론의 諸公理중 특히 獨立性的의 公理의 실증적 타당성에 관해서는 많은 논란이 있는 것이 사실이다. 그러나 합리적 행동의 규범적 조건으로서의 獨立性的의 公理의 타당성은 거의 모든 학자들에 의해 받아들여지고 있다. 獨立性的의 公理에 대해 간단히 설명하면 다음과 같다. 확실한 결과를 가져오는 두 개의 代案 x 와 y 가 있다고 하자. 그리고 가능한 상황은 A 와 B 두 경우 밖에 없고 또한 두 경우는 상호배타적이라고 하자. 즉 실제로 實現되는 것은 A, B 둘중의 하나일 뿐이다. 그런데 A, B 어느 것이 실현되든지 대안 x 를 택했을 때의 결과가 대안 y 를 택했을 때의 결과보다 언제나 더 선호된다고 하면 이러한 선호는 獨立性的의 公理를 만족한다고 말한다.⁽⁴⁾ 이러한 정의에 더욱 적절한 표현으로서 保證原理(surething principle)라 부르기도 한다. 만장일치의 공리는 파레토原理로도 불리우는 것으로서 서수적 효용이론에 기초한 후생경제학의 기본을 이루는 가정이다. 그 가정은 사회적 선택은 個人的 選好에 바탕을 두어야 한다는 개인주의적 價値判斷을 배경으로 하고 있다. 최근 센(A. Sen)은 파레토原理를 포함한 厚生主義(welfarism)를 설득력있게 비판하고 있다 [Sen (1979) 참조]. 그렇지만 이 비판은 공리주의에만 해당하는 것이 아니기 때문에 여기서는 그것을 문제삼지 않기로 한다.

(4) 獨立性的의 公理는 여러 가지의 유사한 방법으로 정의된다. 독립성이라는 표현에 적합한 定義는 다음과 같다. 상황 B 의 경우 두 대안이 모두 같은 결과를 갖는다면, 그 결과에 관계없이, 두 대안 사이의 선호는 상황 A 하에서의 결과에 따라 결정된다.

이제 문제가 되는 것은 하사니의 둘째 가정 즉, 社會的 選好도 期待效用理論의 제공리를 만족해야 한다는 가정이다. 그 중에서 특히 독립성의 공리가 비판의 대상이 되어 왔다. 이 문제를 처음으로 제기한 Diamond(1967)에 의하면 독립성의 공리가 社會的 選好에 適用되면 개인적 선호에서와는 달리 강한 倫理的인 意味를 갖게 된다는 것이다. 즉 일진 윤리적으로 중립적으로 보이는 가정이 사실은 그 자체 내던히 강한 윤리적 가치판단을 내포하고 있기 때문에 하사니定理의 공리주의가 걸려드는 것이다. 그 이후 사회적 선호에 있어서 獨立性的의 公理가 과연 타당성을 갖는가 하는 문제에 관해 활발한 논쟁이 전개되어 왔다 [Sen (1977a, 1977b), Harsanyi(1975b, 1977a) 참조]. 그런데 그 논쟁은 주로 특수한 例와 反例(counterexample)를 들어 그것들의 직관적 호소력을 보여주는 형태로 전개되었을 뿐, 어떤 특정한 예가 직관적으로 타당한가에 대한 體系的인 分析이 없었다. 최근 Broome (1984a, 1984b)이 이러한 방향으로의 최초의 시도라고 보여진다. 본연구는 이러한 흐름의 연구성과를 비판적으로 검토함으로써 체계적인 분석을 위한 기본틀을 제시하고자 한다. 본 연구를 통해서 우리가 찾아낸 결론은 다음과 같이 요약될 수 있을 것이다. 즉 사회적 선호에 있어서 獨立性的의 公理의 타당성은 우리가 正義의 개념을 어떻게 이해하느냐에 따라 결정된다고 할 수 있다. 우리는 여기서 두 개의 正義概念 즉 功利主義와 契約論을 대비시킴으로써 그 차이를 분명히 하고 어떤 정의론을 택하느냐에 따라 사회적 선호에서 독립성의 공리가 타당할 수도 혹은 그렇지 않을 수도 있음을 논증한 것이다. 우리의 이러한 분석은 정의론 혹은 社會選擇理論의 주요한 방법의 하나인 公理的 接近方法과는 대조를 이루게 된다. 公理的 接近에서는 특정 정의론이 만족하는 제 성질을 公理로 추출해내어 그 공리들의 타당성을 검토함으로써 그 정의론을 평가하는 방식을 취하고 있는데, 우리는 역으로 공리의 타당성이 어떤 정의론의 입장에 시는가에 의존하는 경우도 있음을 보여주고 있다. 이러한 사실은 또한 公理的 접근방법의 限界를 의미하기도 한다는 점에서 우리는 본연구의 시사점중의 하나를 찾을 수 있을 것이다.

이 논문의 구성은 다음과 같다. 第2節에서는 하사니의 正義論의 개념적 틀을 검토하고, 그와 대립되는 이론으로서 契約論의 개념적 틀이 하사니의 공리주의와 어떠한 점에서 다른가를 보인다. 第3節에서는 사회적 선호에서의 獨立性的의 公理에 대한 기존의 분석들을 정리하고 그 공리에 대한 평가가 正義概念의 어떠한 인간을 갖는가를 분석하고, 第4節에서 결론을 요약할 것이다.

2. 하사니의 正義概念

하사니의 正義概念 혹은 道徳理論(theory of morality)은 하사니定理가 제시된 1955년의 논문보다 2년 앞서 발표된 1953년의 논문에서 그 기본틀이 제시되고 그 후의 여러 논문에서 더욱 精緻化된 형태로 발전하였다 [Harsanyi(1976, 제 1 부) 참조]. 하사니에 의하면 道徳的 價値判斷이란 개인적 의사결정문제의 한 특수한 형태로 간주될 수 있으며 따라서 도덕이론은 합리적 의사결정이론의 한 분과로 분류할 수 있다고 본다. 즉 하사니는 合理的 意思決定理論을 개인적 의사결정이론, 게임이론, 그리고 도덕이론의 세 분과로 나누어서 분석하고 있다. 개인적 의사결정이론은 개인적 선호의 합리성의 문제를 다루는 반면, 도덕이론은 社會的 選好(social preference)의 合理性을 다룬다. 개인적 선호와 사회적 선호는 모두 어떤 특정 개인의 선택행위를 설명한다는 점에서는 같다. 그러나 개인적 선호는 개인의 이익을 추구하는 선호인데 반해서, 社會的 選好는 이 의사결정을 내리는 특정 개인의 이익이 아니라 사회전체의 공동의 이익을 추구하는 선호로 정의된다. 여기서 문제가 되는 것은 社會全體의 利益을 어떻게 概念化하는가 하는 것인데, 하사니는 공평한 외부의 觀察者(impartial outside observer)의 관점에서 본 이익을 사회의 이익이라고 간주한다. 즉 이 외부의 관찰자는 자신이 사회의 구성원중에서 누구의 입장에 놓이게 될지 완전히 불확실한 상황에서 意思決定을 내린다고 할 때, 그의 행위를 지배하는 선호는 어느 특정 개인의 이익이 아닌 사회전체의 이익을 대변하는 선호라 할 수 있을 것이다. 여기에서 이 의사결정자가 배이지안 合理性을 만족한다고 가정하면 아무 정보가 없는 상황에서 각 사회구성원의 입장에 처할 주관적 확률이 모두 같다는 결론을 내릴 수 있다. 바로 이 균등확률의 조건에 의해 각 사회구성원의 이익이 공평하게 반영되는 것이다. 그런데 이러한 公平한 외부적 觀察者란 실제로 존재할 수는 없는, 그렇지만 도덕적 관점을 설명하기 위해 理論적으로 만들어낸 假想의 존재이다. 하사니에게 있어서 도덕이론은 바로 이러한 가상의 관찰자의 행동에 관한 이론이 된다. 이 관찰자의 행동은 不確實性下에서의 개인의 合理的 意思決定理論으로 설명될 수 있는 것이다.

그런데 이러한 가상의 관찰자의 관점을 설정하게 되면 다음과 같은 중대한 문제가 발생하게 된다. 즉 가상의 관찰자의 입장에 시계 되면 각 사회구성원의 主體性(self-identity)은 무시되게 된다. 각 사회구성원은 사회의 利益을 담는 그릇에 불과하게 되고 롤즈(J. Rawls)가 말한 바와 같이 개인간의 구별(distinction between persons)은 신중히 고려되지 않게 된

다 [Rawls(1971, p. 27) 참조]. 도덕의 세계에는 사회의 利益의 極大化라는 하나의 기준이 있을 뿐이고, 생각하고 의사결정을 내리는 단 하나의 주체가 있을 뿐이다. 따라서 사회구성원간의 相互關係는 무시된다. 공리주의가 總效用의 크기에만 관심이 있을 뿐, 효용이 개인들간에 어떻게 분배되는가에는 무관심한 것은 바로 이러한 이유 때문이다. 그렇다면 효용의 平等한 分配를 바람직한 것으로 간주하는 우리의 도덕적 직관을 반영하는 도덕이론은 다른 곳에서 찾아야 할 것이다.

하시니의 도덕이론과는 달리 契約論(contract theory)에서는 외부의 관찰자의 관점을 취하지 않는다. 契約論이 공리주의와 구별되는 차이점은 많이 있지만 가장 중요한 차이점은 사회구성원이 아닌 어떤 초월적 존재의 관점을 취하는가 아닌가에 있다고 보여진다. 계약론적 입장에서는 道德的 判斷이란 특정 개인의 판단이 아니라, 사회구성원이 각자 자신의 주체성을 보유하면서 시로의 협의를 통해서 도출해낸 합의에 따라 내리는 판단이다. 그런데 이러한 合意의 公正性은 어떤 가상의 이론적인 공정한 상태에서 공정하다고 생각되는 질차에 따라 합의가 이루어질 때 보장될 수 있는 것이다. 이 공정한 상태와 공정한 질차를 어떻게 구성하느냐에 따라 다양한 道德理論이 도출될 수 있을 것이다. 롤즈의 원초적 입장(original position)과 無知의 베일(veil of ignorance)은 이러한 公正性을 달성하기 위한 理論的 裝置인 것이다. 롤즈가 자신의 정의론이 公正으로서의 正義(justice as fairness)임을 강조하는 것은 이러한 계약론적 입장에서 서있음을 밝히기 위한 것이다. 계약론에서는 의사결정을 내리는 단일의 주체를 가정하지 않기 때문에, 功利主義에서와 같이 극대화되어야 할 공동의 善(common good)은 존재하지 않는다. 社會構成員 각자는 자신의 정당한 몫에 관심을 가지며 부당한 대우를 받지 않도록 노력할 것이다. 각 개인은 가능한 한 자신의 이익을 극대화하려고 노력하지만 그러나 또 한편으로는 타인과의 관계에 있어서 公正性을 維持하는 범위 안에서 자신의 이익을 추구해야 한다는 사실을 인식하고 있다. 따라서 공정하게 구성된 상태와 질차하에서 개인들은 일종의 게임적 상황에 처하게 되고 그 게임의 균형 혹은 解(solution)로서 道德原理가 도출된다고 할 수 있을 것이다. 즉 契約論的 道德理論은 게임이론을 이용하여 분석하는 것이 자연스런게 된다.

하시니의 공리주의와 계약론은 두 개의 대조적인 도덕적 관점을 대표하는 것이지만 또한 많은 공통점을 갖고 있는 것도 사실이다. 먼저 두 이론 모두 개인의 合理的 選擇理論에 그 바탕을 두고 있다. 하시니의 功利主義는 가상적 불확실성에 처한 超越的 主體의 선택행위의 이론이고, 계약론은 게임적 상황에 처한 개인들의 選擇行爲를 다루고 있다. 이 점에서 두 이론 모두 현대 합리적 의사결정이론의 비약적인 발전에 힘입고 있는 것이다. 둘째로

두 이론 모두 에로우에 의해 정식화된 社會選擇理論의 형식적 틀 안에서 정식화될 수 있다. 이같이 정식화할 때 두 이론의 차이는 效用的 測定可能性과 개인간 比較可能性에 관한 차이로 환원되어 버린다. 즉 Sen(1984)에 의해 시작되고 정리된 사회선택이론의 情報的 접근 방법에 따르면, 기수적 효용의 개인간 비교가능성을 인정하면 공리주의가 도출되고, 서수적 효용의 비교만을 허용하면 契約論의 관점에 입각한 롤즈의 最小極大化原則(maximin principle)이 도출된다는 것이다. 또한 하사니의 공리주의론과 롤즈의 '원초적 입장'의 이론 사이의 방법론적 유사성이 Arrow(1973) 등에 의해 지적되어 왔다. 즉 하사니는 공리주의를 契約論的으로 정당화한 반면에, 롤즈의 이론은 序數的 功利主義라고 불리울 만큼 공리주의적 측면이 있음이 널리 인식되어 왔다.

그러나 사회선택이론에서 강조되는 이러한 공통점 혹은 유사성으로 인하여 두 이론의 本質的인 差異點이 간과되는 경향이 있다. 롤즈는 이러한 문제점을 인식하면서 자신의 이론이 어떻게 서수적 공리주의로 지칭되는 이론과 다른가를 보이려고 노력하고 있다 [Rawls (1982) 참조]. 롤즈의 주장의 요점은 자신의 이론은 效用水準의 比較를 가정하고 있지 않을 뿐 아니라 가정할 필요도 없다고 하는 것이다. 롤즈가 개인간의 비교의 대상으로 사용하는 基礎財(primary good)의 개념은 효용과 다르다는 것이다. 롤즈는 개인들의 효용이란 기본적으로 比較不可能한 것으로 인정함으로써 개인간의 극복불가능한 차이와 다양성을 인정한다는 전제하에서, 단지 모두가 자신의 목적과 요구를 충족시키는 데 필요하다고 인정하는 몇 개의 기초재의 公正한 分配에만 정의론의 대상을 限定해야 한다고 주장한다. 그러나 그는 원초적 입장에서의 선택행위를 대표적 개인의 合理的 選擇問題로 환원해 버림으로써 초월적 개인의 선택을 다루는 공리주의와 같은 관점을 택하게 되었다. 그 결과 하사니의 功利主義와 롤즈의 契約論의 차이는 합리적 선택이론에서의 期待效用理論과 最小極大化原則 중에서 어느 것을 선택하는가 하는 문제로 축소되어 버린 것이다. 롤즈는 개체성을 유지하면서 이론을 전개해야 할 필요성을 충분히 인지하고 있으면서도, 원초적 입장의 구체적 분석과정에서는 개체성을 유지하는 데 실패하였다고 할 수 있다.

하사니의 功利主義論과 契約論의 근본적인 차이는 超越的 主體의 관점을 택하는가 혹은 사회구성원중 한 個人의 관점을 택하는가 하는 데에 있다고 요약할 수가 있다. 이 두 관점의 차이가 어떠한 결과를 가져오는가를 예를 들어 살펴보기로 하자. 지금 관찰될 수 있는 모든 면에서 같은 두 개인으로 구성된 社會를 상징하자. 이제 두 사람 앞에 100원이 주어지고 이것을 나누는 정의로운 방법을 고찰해보자. 문제를 좀더 단순화하기 위해 두 사람의 선호는 危險中立的이라 가정하자. 그리고 어느 한 사람이 100원을 모두 차지하는 분

배는 사회적으로 無差別하다고 가정하자. 이것은 두 사람중 어느 누구도 더 큰 권리를 갖고 있지 못함을 의미한다. 이 문제에 관한 구체적인 분석은 다음 節에서 행할 것이다. 그런데 다음 節에서 증명되는 바와 같이 功利主義에서와 같은 초월적인 주체의 관점을 택한다면 어떠한 분배든지 總效用의 크기는 같기 때문에 社會적으로 무차별하게 된다. 개인은 단지 效用의 그릇에 불과할 뿐이므로 누가 얼마의 效用을 누리는가는 超越的 主體의 관심 밖에 있다. 그러나 사회의 구성원인 개인의 관점에서 본다면 서로 양보할 수 없는 상충되는 이해관계를 갖게 되고 두 사람이 모두 승인할 수 있는 유일한 분배는 50원씩 나누어 갖는 평등한 분배일 것이다. 왜냐 하면 다른 어떠한 不平等한 分配도 적게 받는 사람의 관점에서는 동의할 이유가 없기 때문이다.

3. 獨立性的 公理

우리는 앞에서 공리주의를 비판하기 위해서는 하사니定理의 가정들이 타당하지 않음을 보여야 한다고 주장하였다. 그런데 Sen(1977a, 1977b)은 하사니定理의 가정들을 모두 받아들인다 하더라도 功利主義的 도덕판단에 따라야 될 필연성은 없다고 한다. 우리는 먼저 센의 이 주장이 옳지 않음을 보이고자 한다. 센에 의하면 하사니定理가 보여주는 것은 社會厚生과 개인의 효용이 vNM효용함수와 같다면 功利主義的 사회후생함수를 얻을 수 있다는 것인데 우리가 도덕적 가치판단의 문제에서 사용하는 사회후생과 개인의 효용이 vNM效用函數와 일치할 이유는 없다고 한다. 個人과 社會의 선호가 주어졌을 때 그 선호와 일관된 효용함수는 vNM효용함수뿐 아니라 vNM효용함수의 임의의 단조증가함수도 모두 포함한다. 그리고 이러한 임의의 단조증가함수들간에는 vNM효용함수들간에서와 같은 線形關係가 성립하지 않는다는 것은 자명하다. 그런데 vNM효용함수는 不確實性에 대한 개인들의 태도를 반영하는 것으로서 道德的 의미를 갖는 것은 아니라고 센은 주장한다. Arrow(1963, p. 10)도 같은 견해를 표명하고 있다. “vMN효용은 후생판단과는 아무런 관계가 없다. ……그렇지 않다면 소득의 분배가 개인들의 도박에 대한 嗜好에 의해 지배되어야 한다고 주장하는 것과 다르다.” 이 비판에 대해 Harsanyi(1982)는 vNM效用函數는 단지 위험에 대한 태도만을 나타내는 것이 아니라 진정한 선호의 強度를 나타내는 것임을 강조하고 있다. 즉 어떤 사람이 위험을 무릅쓰는 것은 그가 위험 그 자체를 즐기기 때문이 아니라 성공했을 때 얻는 결과를 매우 높이 평가하고 있기 때문이라고 해석할 수 있다는 것이다. 그리고 이렇게 測定된 效用은 선호의 強度를 적절히 표현하고 있으므로 도덕적 판단에 사용하지 못할 이

유가 없다. 그런데 과연 vNM效用函數가 센이나 애로우의 주장과 같이 순수히 위험에 대한 태도만을 반영하는 것인지, 혹은 하사니의 주장대로 선호의 強度를 나타내는 것인지를 실증적으로 검증할 방법은 없다.

그러나 이 논증에서 어떤 입장을 취하든지간에 한 가지 분명한 것은, 社會的 代案들중에는 결과가 불확실한 選擇代案들도 있다는 점이다. 그 불확실성은 氣候와 같이 외부적인 요인에 기인할 수도 있고, 혹은 의도적으로 도입된 것일 수도 있다. 예를 들어 추첨에 의해 소득분배를 결정하는 방법을 명시적으로 하나의 代案으로 고려할 수 있다. 이와 같이 불확실한 대안들에 대한 선호를 표현하기 위해서는 vNM效用函數를 적용하는 것이 타당하게 된다. 물론 불확실한 대안들을 고려하는 경우에도 vNM효용함수뿐 아니라 그 임의의 單調增加函數도 같은 선호를 표현한다고 할 수 있다. 이러한 임의의 단조증가함수는 단지 기대효용의 성질을 갖지 못할 뿐인 것이다. 그렇다면 이 때도 여전히 센의 비판은 유효한 것처럼 보인다. 즉 vNM효용함수가 아니라면 社會厚生函數는 공리주의의 형태인 선형함수가 될 수 없다는 것이다. 그러나 效用函數가 선호를 표현하는 것이고 같은 선호를 표현하는 효용함수가 여럿 있을 수 있는 것처럼 사회후생함수 그 자체는 배후에 있는 道德的 原理를 표현하는 하나의 수단일 뿐으로서 효용함수의 선택에 따라 그 형태가 달라진다 하더라도 본질적으로 功利主義的 道德原理는 그대로 유지될 수 있는 것이다. 우리는 제 2 절에서 언급된 例를 좀더 구체적으로 살펴봄으로써 이 점을 분명히 할 수 있을 것이다.

이제 甲, 乙 두명으로 구성된 사회에서, 100원을 나누는 문제를 생각해 보자. 100원을 임의의 작은 단위로 나눌 수 있다고 가정하면 무한히 많은 分配方式를 생각할 수 있다. 그 중에서 갑이 100원을 모두 갖는 분배를 A, 을이 모두 갖는 분배를 B, 갑과 을이 50원씩 나누는 분배를 C 라 하고, 당첨확률이 같은 추첨에 의해 당첨된 사람이 100원을 모두 갖는 방식을 D 라 하자. 그러면 D는 A와 B를 價으로 하고 각각의 確率이 1/2인 복권과 같다고 할 수 있어서 $D = 0.5A + 0.5B$ 로 표시할 수 있다. 또한 제 2 절에서와 같이 두 사람의 선호는 모두 危險中立의이라 가정하고, A와 B는 도덕적으로 무차별하다고 가정하자. 이러한 상황에서 하사니의 諸公理가 함축하는 바를 도출해 보기로 한다. 먼저 선호가 危險中立의이므로 갑, 을 각자에게 대안 C와 D는 무차별하다. 그러면 만장일치의 公理에 의해 C와 D는 도덕적으로도 무차별하게 된다. 또한 A와 B가 도덕적으로 무차별하므로, 獨立性的 公理를 적용하면 A, B, D가 모두 도덕적으로 무차별하게 된다. 마지막으로 선호의 移行性에 의해 A, B, C, D가 모두 도덕적으로 무차별한 것이 된다. 같은 논리에 의해 임의의 분배방식이 모두 무차별하다는 것이 증명될 수 있다. 똑같은 결론을 vNM效用函數를

사용하여 얻을 수 있다. 값이 x 원으로부터 얻는 효용은 위험중립성에 의해 $ax + b$ 의 형태로 표시되며 마찬가지로 B가 $(100 - x)$ 로부터 얻는 효용은 $c(100 - x) + d$ 의 형태가 된다. 분배방식 $(x, 100 - x)$ 의 사회후생은 하사니정리에 따르면 두 사람의 효용의 승인 $(ax + b) + (c(100 - x) + d)$ 와 같다. 그런데 분배 A와 B가 무차별하므로 x 가 0일 때와 100일 때의 社會厚生이 같아야 한다. 즉 $b + 100c + d = 100a + b + d$ 이다. 따라서 $a = c$ 가 되고 社會厚生은 x 와 관계없이 $100a + b + d$ 로 일정하게 된다.

이 예에서 우리는 다음과 같은 두 가지의 사실을 확인할 수 있다. 첫째로 vNM효용함수를 사용하지 않더라도 불확실한 代案을 고려함으로써, 하사니의 가정들로부터 功利主義的 道德原理를 도출할 수 있게 된다. 즉 선택과 애로우의 주장과는 달리 vNM효용함수의 해석 방법에 관계없이, 하사니의 假定들을 받아들인다면 공리주의적 도덕판단은 피할 수 없게 되는 것이다. 둘째로 일견 대단히 일반적으로 보이는 이와 같은 分配問題에서 어떻게 분배 하든지 모두 무차별하다고 하는 판단은 우리의 분배적 정의에 관한 직관에 맞지 않는다. 여기에서 우리는 하사니의 假定들을 再檢討하여야 할 필요가 있다.

이제 하사니의 가정들중에서 특히 獨立性의 公理의 타당성을 검토하기 위하여 먼저 하사니 자신의 주장을 살펴보기로 하자. 하사니에 의하면 개인의 選好가 合理的이기 위해서 독립성의 공리를 만족해야 한다면 社會的 選好는 당연히 그것을 만족해야 한다. 왜냐 하면, “우리가 사회의 이익을 위해 행동할 때는 물론 다른 사람들의 이익을 위해 행동할 때에도, 우리가 우리 자신의 이익을 위해 행동할 때보다 훨씬 더 높은 수준의 合理性에 따라 행동해야 할 의무가 있는 것이다.”[Harsanyi(1975b, p. 316)] 이 때 우리는 하사니가 자신의 道德概念에 따라 사회적 선호를 이해함으로써 獨立性의 公理를 정당화하고 있음을 주목해야 한다. 즉 社會的 選好란 사회의 이익을 위해 행동하는 개인의 선호이므로 그 개인이 合理的 行動의 조건인 독립성의 공리를 따르지 않을 이유가 없는 것이다. 이러한 관점이 어떤 의미를 갖는가를 보기 위해 구체적 예를 들어 고찰해보자.

앞에서 든 예에서 분배 A와 D만을 다음의 <表 1>과 같이 비교해보자. 表에서 사건 H와 T는 각각 확률 1/2의 상호배타적인 사건들이다. 또한 (100, 0)은 π에게 100원 진부름

<表 1>

代 案	事 件	
	H	T
A	(100, 0)	(100, 0)
D	(100, 0)	(0, 100)

주는 分配를 표시한다. Diamond(1967)에 의하면 D는 A보다 바람직하다. 왜냐 하면 D의 경우 甲과 乙 모두에게 공평한 기회(fair shake)를 주기 때문이다. 즉 결과적으로는 어느 한 사람이 100원을 모두 갖는 것이므로 不平等이 초래되는 절에서는 A와 D가 마찬가지로이지만, D의 경우 절차상에서 公正性이 달성되고 있다. Sen(1977)도 이 주장에 동의하면서 독립성의 공리는 이 절차상의 공정성을 배제하므로 받아들이기 어렵다고 하였다. 그러나 하사니의 道德理論에 따른다면 D가 더 공정하다고 볼 수 없다. 즉 도덕적 판단을 내리는 超越의 存在의 관점에서 보면 사건 H와 T중 어느 것이 실현될지 모를 뿐 아니라 甲과 乙의 어느 입장에 놓일지 모르는 상황에 있으므로 분배 A와 D 모두 똑같이 0와 100원을 상으로 하는 복권과 같게 된다. 따라서 A와 D는 道德的 觀點에서 볼 때 무차별하게 된다. 이것을 좀 다른 각도에서 보면 다음과 같이 설명할 수도 있다. 사회적으로 이미 甲이 모두 갖는 분배 A나 乙이 모두 갖는 分配 B가 無差別하다고 하자. 그러면 둘중에서 어느 하나가 실제로 실현된다면 그것은 無作爲選擇(random selection)의 결과로 간주할 수 있을 것이다. 따라서 명시적인 추첨방법을 도입한다고 해서 더 공정하게 되는 것은 아니다.

Broome(1984a)은 D가 A보다 사회적으로 선호되어야 한다고 인정하더라도 獨立性的의 公理의 타당성 여부와는 무관하다고 주장하였다. 그는 이것을 다음과 같이 설명한다. 먼저 期待效用은 효용과는 구별되어야 한다. 효용이 무엇을 의미하든지간에, 효용은 그 자체 내재적으로 바람직한 것인 반면, 期待效用은 그렇지 않다는 것이다. 效用이 욕구의 정도를 나타낸다고 하면, 기대효용은 수단에 대한 욕구인 반면 효용은 결과에 대한 욕구의 정도를 나타내는 것이다. 우리가 期待效用이 價値가 있다고 하는 것은 기대효용을 크게 함으로써 효용을 증대시킬 수 있기 때문이다. 따라서 효용의 평등이 어떠한 근거에서든지 바람직한 것이라 하더라도 그 때문에 期待效用의 평등도 바람직하게 되는 것은 아니라는 것이다. D가 選好되는 것은 기대효용이 균등하기 때문이 아니라 누가 100원을 갖는가를 결정하는 공정한 절차가 갖추어져 있기 때문이다. 無作爲選擇이 공정한 절차의 한 방법이라면 이것이 결과에 함께 명시적으로 記述되어야 한다. 브롬의 표를 약간 수정하여 만든 <表 2>가 이러한 방식으로 代案들을 묘사하고 있다.

여기에 獨立性的의 公理를 적용하더라도 공정한 선택이 바람직한 것이라면 代案 D가 A보다 선호될 수가 있는 것이다. 문제는 獨立性的의 公理에 있는 것이 아니라 無作爲選擇 방식이 과연 공정한 절차를 보장하는가에 있다고 브롬은 보는 것이다. 그런데 이와 같이 獨立性 公理의 타당성에 관한 문제와 무작위선택의 공정성의 문제를 분리해서 고찰하는 것이 과연 정당한가 하는 의문을 제기할 수 있다. 먼저 期待效用과 效用이 다른 것으로 보기 어

〈表 2〉

代 案	事 件	
	H	T
A	(100, 0)와 불공정한 선택	(100, 0)와 불공정한 선택
D	(100, 0)와 공정한 선택	(0, 100)와 공정한 선택

〈表 3〉

代 案	선택을 하기 위해 동전을 던져 앞이 나왔음	선택을 하기 위해 동전을 던져 뒤가 나왔음
A	(100, 0)	(100, 0)
D	(100, 0)	(0, 100)

럽다. 기대효용이란 불확실한 대안에서 얻는 효용이다. 不確實한 대안으로부터 얻는 期待效用과 그와 동등한 확실한 대안에서 얻는 효용은 질적으로나 양적으로나 같은 것이다. 따라서 效用的 均等化가 바람직하다면 기대효용의 균등화도 바람직한 것이 된다. 다음에 무작위선택이 공정한 것인지에 관한 판단이 〈表 2〉에서는 의생적으로 결정되어 삼입되고 있다. 그런데 문제를 먼저 객관적으로 記述하고 거기에 道德原理를 적용하여 대안을 선택하는 것이 올바른 이론구성이라고 할 수 있는데 브롬의 경우 문제의 敘述에서부터 이미 道德原理의 適用을 필요로 하고 있는 것이다. 무작위선택의 공정성에 대해 가치판단을 미리 도입하기보다는 불확실한 상태를 보다 완전히 기술해서 모형을 구성하는 것이 더 바람직하다고 생각된다. 그 예로서 〈表 3〉과 같이 문제를 기술할 수도 있을 것이다. 이렇게 하면 獨立性 公理의 정당성은 무작위선택의 공정성 여부에 의해 판정되게 된다.

결국 功利主義의 評價에 있어서 관건이 되는 것은 독립성의 공리의 정당성의 여부에 있고 그것은 또한 무작위선택의 공정성과 직결되는 문제라는 것이 밝혀졌다. 그런데 無作為選擇의 公正性은 어떤 관점을 택하느냐에 따라 달리 판정될 수 있다. 앞에서 본 바와 같이 초월적 관찰자의 관점에서 보면 이미 누구의 입장에 처할지 不確實한 상황에서 선택하는 것이므로 意圖的인 무작위선택의 도입은 선택에 아무 영향을 미치지 못한다. 따라서 무작위선택은 더 公正한 것이 아니다. 그러나 사의적 선택이 超越的 主體에 의해 이루어지는 것이 아니라 구성원들의 合意의 결과라 하면 각자는 자신의 주체적 관점에서 판단하므로 예를 들어 乙에게는 대안 A는 명백히 불공정한 반면, 대안 D는 기회가 주어진다라는 점에서 보다 공정하다고 할 수 있다. 즉 하사니는 超越的 主體의 관점에서 문제를 고찰한 반면에, 센과 다이아몬드는 社會構成員의 觀點에서 평가하기 때문에 서로 다른 결론을 내리고 있는 것이다. 그렇다면 어떤 觀點이 더 타당하다고 할 수 있는가? 이것은 구체적인 문제의 성

질에 따라 결정될 성질의 것이라 할 수 있다. 실제로 우리는 여러 종류의 현실적 문제에 직면할 때마다 무의식적으로 문제에 따라 끊임없이 관점을 바꾸고 있다.

두 개의 관점이 각각 타당한 것으로 보이는 예를 들어서 설명하기로 한다. 첫번째 예는 Harsanyi(1975b)로부터 인용해서 재구성한 것이다. 지금 정부가 두 政策代案을 가지고 있다. 대안 A는 특정 산업에 종사하는 소수에 특혜를 주고 있는 보호관세를 폐지함으로써 대다수의 이익을 증진시키는 정책이다. 代案 B는 代案 A에서와 똑같이 보호관세를 폐지하면서 동시에 누가 어느 입장에 처할지를 추첨에 의해 결정하는 것이다. 정책 A와 B는 동일한 所得分布를 갖게 된다. 이 경우 정책 B가 A보다 낫다고 말할 수 없을 것이다. 정부가 의도적으로 행하는 부작위추첨이 사람들이 특정 위치에 있게 되는 역사적 우연보다 도덕적으로 우월하다고 볼 수는 없는 것이다. 두번째 예는 Broome(1984b)이 들고 있는 例중의 하나이다. 징병제를 실시하는데 적령기에 있는 자들중에서 추첨에 의해 선발할 수도 있고, 혹은 社會的 效率性의 관점에서 체력이 강하고 우수한 자를 선발하는 방법을 택할 수도 있다. 후자의 경우 이 原則에 충실하게 참으로 강하고 우수한 자가 선발된다고 하자. 즉 이 원칙이 모든 자에게 공정하게 적용된다고 가정하자. 그러나 그럼에도 불구하고, 강하고 우수한 자는 예외없이 징집될 것이므로 처음부터 불공정하다고 항의할 것이다.

첫째 例에서는 정책의 評價가, 누가 무엇을 얻는가에는 관심이 없고 단지 소득분배의 일반적 패턴에만 관심이 있는 정부당국자와 같은 外部의 觀察者의 관점에서 이루어지고 있다. 외부의 관찰자는 그 자신 각 사회구성원의 입장에 처할 確率이 모두 같은 가상적 상황에서 意思決定을 내림으로써 공정성을 확보할 것이다. 따라서 명시적인 無作爲選擇 방법을 도입한다고 해서 公正性이 증대되지 않는다. 이러한 종류의 문제에서는 이와 같이 제 삼자의 입장을 택해서 사회적 판단을 내리는 것이 자연스럽고 따라서 사회적 선호는 독립성의 공리를 만족하게 된다. 그러나 둘째 예에서는 사회를 대표하는 제 삼자를 상정하는 것이 부자연스럽거나 불가능한 경우이다. 이 때에는 개별 사회구성원의 관점에서 公正性을 기준으로 도덕판단을 내릴 수밖에 없게 된다.

4. 結 論

본연구의 결론은 다음 세 사실로 요약될 수 있다. 첫째, 두 개의 내립적인 道德的 觀點은 도덕이론의 양대 思想 즉 功利主義와 契約論과 밀접한 관련을 갖고 있다. 功利主義는 초월적 주체의 관점에서 도덕적 판단을 구하는 반면에, 契約論은 사회구성원 각자의 주체

적 관점으로 끊임없이 돌아온다. 공리주의는 超越的 主體의 판단을 다루기 때문에 전통적인 의사결정이론의 틀 내에서 분석이 가능하였으며 체계적인 발전이 이루어졌다. 그러나 契約論의 근원은 수 세기를 거슬러 올라가지만 최근까지 이론적으로는 공리주의에 비해 상대적으로 체계화되지 못했다. 이것은 계약론이 사회구성원 모두의 主體性을 維持하면서 이론을 구성하는 데 어려움이 있기 때문이다. 다양성과 개인의 不可測性(incommensurability)을 인정하면서 어떻게 사회적 통합을 이룰 수 있는가라는 문제는 아직도 완전히 해결되지 못한 상태에 있다.

둘째, 獨立性의 公理의 타당성은 어떤 도덕적 관점을 택하느냐에 달려 있으며 그 역은 아니다. 즉 독립성의 공리의 妥當性 때문에 도덕적 관점이 타당하게 되는 것이 아니다. 기존의 문헌에서는 모두 독립성의 공리의 타당성을 검토하고 그 결과로서 공리주의를 평가하고 있는데 이러한 公理的 接近方法이 적어도 여기서는 재검토되어야 한다.

셋째, 물즈는 契約論에 입각한 자신의 정의론을 하사니와 같이 불확실성하에서의 개인의 의사결정문제로 정식화하여 분석하였다. 하사니는 기대효용이론을 택한 반면, 물즈는 最小極大化原則을 택하고 있는 점이 다른 뿐이다. 그런데 우리가 찾아낸 사실은 契約論은 다양한 의사결정자의 관점을 동시에 채택해서 분석을 행하는 데 그 특징이 있다는 것이다. 개인의 의사결정이론에서는 하사니가 채택한 期待效用理論이 더 일반성과 타당성을 갖고 있다는 것은 널리 알려진 사실이다. 계약론의 고유한 체계를 유지하면서 이론을 전개하는 하나의 방법은, 사회구성원들의 각자의 이익추구가 초래하는 社會的 均衡을 분석하는 게임이론의 분석도구를 이용하는 것이다.

서울대학교 經濟學科 助教授
151-742 서울 관악구 신림동
전화 : (02)880-6378
팩시 : (02)888-4454

參 考 文 獻

- Arrow, K.J. (1963): *Social Choice and Individual Values*, Wiley.
 _____ (1973): "Some Ordinalist-Utilitarian Notes on Rawls' Theory of Justice," *Journal of Philosophy*, **70**, 245~263.
 Broome, J. (1984a): "Uncertainty and Fairness," *Economic Journal*, **94**, 624~632.
 _____ (1984b): "Selecting People Randomly," *Ethics*, **95**, 38~55.
 Butts and Hintikka eds. (1977): *Foundational Problems in the Special Sciences*, Dordrecht,

Reidel.

- Diamond, P.A. (1967): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: A Comment," *Journal of Political Economy*, **75**, 765~766.
- Elster, J., and A. Hylland (1986): *Foundations of Social Choice*, Cambridge University Press.
- Harsanyi, J.C. (1953): "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking," *Journal of Political Economy*, **61**, 434~435.
- _____ (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, **63**, 309~312.
- _____ (1975a): "Can He Maximum Principle Serve as a Basis for Morality? A Critic of John Rawls' Theory," *American Political Science Review*, **69**, 594~606.
- _____ (1975b): "Nonlinear Social Welfare Functions: Do Welfare Economists Have a Special Exemption from Bayesian Rationality?," *Theory and Decision*, **6**, 311~332.
- _____ (1976): *Essays in Ethics, Social Behavior, and Scientific Explanation*, Dordrecht, Reidel.
- _____ (1977a): "Nonlinear Social Welfare Functions: A Rejoinder to Professor Sen," in Butts and Hintikka (eds.), *Foundational Problems in the Special Sciences*, Dordrecht, Reidel.
- _____ (1977b): *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press.
- _____ (1982): "Morality and the Theory of Rational Behavior." in Sen and Williams (eds.), *Utilitarianism and Beyond*, Oxford University Press.
- Jeffrey, R.C. (1971): "On Interpersonal Utility Theory," *Journal of Philosophy*, **68**, 647~656.
- Ng, Y.K. (1981): "Welfarism: A Defence against Sen's Attack," *Economic Journal*, **91**, 527~533.
- Rawls, J. (1971): *A Theory of Justice*, Harvard University Press.
- _____ (1974): "Concepts of Distributional Equity: Some Reasons for the Maximin Criterion," *American Economic Review*, **64**, 141~146.
- _____ (1982): "Social Unity and Primary Goods," in Sen and Williams (eds.), *Utilitarianism and Beyond*, Oxford University Press.
- Samuelson, P.A. (1974): "Complementarity: An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory," *Journal of Economic Literature*, **12**, 1255~1289.
- Scanlon, T.M. (1982): "Contractualism and Utilitarianism," in Sen and Williams (eds.), *Utilitarianism and Beyond*, Oxford University Press.
- Sen, A. (1977a): "Welfare Inequalities and Rawlsian Axiomatics," in Butts and Hintikka (eds.), *Foundational Problems in the Special Sciences*, Dordrecht, Reidel.
- _____ (1977b): "Nonlinear Social Welfare Functions: A Reply to Professor Harsanyi," in Butts and Hintikka (eds.), *Foundational Problems in the Special Sciences*, Dordrecht, Reidel.
- _____ (1979): "Personal Utilities and Public Judgements: Or What's Wrong with Welfare

Economics?” *Economic Journal*, 89, 537~558.

_____ (1981): “A Reply to ‘Welfarism: A Defence against Sen’s Attack,” *Economic Journal*, 91, 531~535.

Sen A., and B. Williams (1982): *Utilitarianism and Beyond*, Oxford University Press.