Doctoral Dissertation

A Critical Re-evaluation of "Welfarism" and a Proposed Alternative from the View Point of Social Justice

Shojiro YASUI

2009

Graduate School of Decision Science and Technology

Tokyo Institute of Technology

Abstract

As the beginning, Chapter 1 describes the definition of “welfarism” and reveals its deficiencies which are caused by limiting information of “revealed preference”, irrespective of the non-utility features of the states. This chapter examines Rawls’s principles of justice and decides to employ Rawls’s “generalized” differential principle instead of the specific one as a benchmark to evaluate improvements of decision making methodology in the whole thesis. Following which, this chapter proposes incorporating the prioritization of “status quo” into decision-making to realize socially justifiable decision-making. The latter chapters examine if the proposed methodology can remedy the deficiencies above in the use of the generalized differential principle, the principle of everybody’s advantage, as the benchmark.

“Arrow’s general possibility theorem” (1963) has revealed that three basic democratic conditions, which consist of, “unrestricted domain” (U: each individuals has un-restricted alternatives.), “independence of irrelevant alternative” (I) and “Pareto optimality” (P), cannot consistent with the condition of “non-dictatorship”. If collective choice is regarded as multiple criteria decision of a single person, Arrow’s theorem also affirms impossibility of implementation of multi-criteria decision, although some psychological studies show opposite results. To resolve this issue, Chapter 2 proposes a social welfare function which employs and prioritizes “status quo”, and relaxes the condition of “independence of irrelevant alternative”. The main result of this chapter reveals that if the proposed social welfare function is applied to the model proposed, non-cyclic social orderings could be an output in accordance with conditions U, P and D. Although the relaxation of the condition I is inescapable, however, the use of such restricted non-utility function enables us to identify the social ordering in which the highly preferred alternative than the status quo shall be highly preferred by all criteria (individuals), than the status quo. Thus the ordering satisfies Rawls’s “generalized” differential principle as well.

The traditional Cost benefit analysis (CBA), which is deeply rooted in welfarism, in particular, utilitarianism, has long been the preferred methodology for evaluating economic factors in policy making. The traditional CBA is also regarded as a social welfare function which can designate a "social ordering" of all alternatives. Chapter 3 proposes incorporating the concept of "time sequence" and "status quo" into the traditional CBA framework whose guiding concept is the "Pareto optimization" of outcomes for the stakeholders. Chapter 3 further proposes Rawls's principles of justice, in particular, "generalized" differential principle which requires everybody's advantage as evaluation criteria for CBA, and thereby critiques the traditional and alternative variants against this standard, ultimately to demonstrate the superiority of the latter. The major findings of Chapter 3 reveals that the alternative can guarantee an advantageous situation compared to the status quo for every stakeholders as well as the alternative can satisfy Rawls's principle of everybody's advantage, which is the benchmark of a social decision-making.

Chapter 4 examines the feasibility of the adoption of the principles of social justice, in particular, the "generalized" differential principle, which require everyone's advantage, to the decision-making process in a strategic situation. The concept of equilibrium, which is a dominated concept of game theory, does not utilize non-utility information and cannot guarantee to benefit all players compared with the status quo. For that purpose, as the application, this chapter introduces the game model in extensive form which has the following characteristics: (a) sub-game perfect equilibrium (SPE) should be employed as a solution, (b) the status quo should be treated as a disagreement point, and (c) a game tree should be divided into several stages (multi-stage approach). Applying "implementation theory" or "mechanism design", Chapter 4 focuses on defining the condition of the subset of "equilibrium outcome" which satisfy SPE as "solution concept" and unanimous decision as "mechanism" to meet "everybody's advantage" principle as "choice correspondence". Thereby, we examine the interrelations of the strategy of players. As the application, the chapter analyzes a collective agreement of industrial relations, in particular, the effects of the "white collar" exemption system on working hours and productivity of agents, in application with the traditional principal-agent theory. The analysis of this chapter confirms the

coherency of Rawls's "generalized" differential principle with a strategic situation. The main result of the application is that the adoption of the "white collar" exemption system may induce shorter working hours compared with those under the current work-hour payment system. The analysis also reveals that the overtime premium has positive effect on working hour elongation.

From the over all analysis in previous chapters, Chapter 5 concludes that the decision-framework which prioritizes the status quo is able to be coherent with traditional decision making methodologies if they are partially modified. The proposed alternative is advantageously able to satisfy the Rawls's principles of justice, in particular the "generalized" differential principle. The results, however, reveal further research needs which include a development of real-time modeling so that the advantage of the proposed alternative can be fully demonstrated. Besides, the results show that the sequence of decisions is also an important factor to be analyzed, because the status quo or a precondition of the first decision can affect the latter decision in a chain reacted manner. For the analysis of the sequence, the further research of setting a boundary of the scope of a decision, or "protected sphere", is required.

Table of Contents

Chapter 3: A Critical Review of the Traditional Methodology of Cost Benefit Analysis and a Proposed Alternative

Chapter 4:A feasibility Study of the Rawls's Principles of Social Justice in Strategic Situations

Chapter 5: Conclusion

Chapter 1

Introduction

Summary

This chapter describes the definition of "welfarism" and reveals its deficiencies which are caused by limiting information of "revealed preference", irrespective of the non-utility features of the states. This chapter examines Rawls's principles of justice and decides to employ Rawls's "generalized" differential principle instead of the specific one as a benchmark to evaluate improvements of decision making methodology in the whole thesis. Following which, this chapter proposes incorporating the prioritization of "status quo" into decision-making to realize socially justifiable decision-making. The latter chapters examine if the proposed methodology can remedy the deficiencies above in the use of the generalized differential principle, the principle of everybody's advantage, as the benchmark.

1. Background: Framework of New/Traditional Welfarism

The so called "new welfarism", defined by Sen with the adoption of Wassily Leontief's summary of the normative properties "on which something like a general consensus of opinion seems to exist" in the formal discussion of public policy, has the following properties. (Sen, 1982), (Leontief, 1966)

- Welfarism: Social welfare is a function of personal utility levels, so that any two social states must be ranked entirely on the basis of personal utilities in the respective states (irrespective of the non-utility features of the states).
- Ordinalism: Only the ordinal properties of the individual utility functions are to be used in social welfare judgments.
- Non-comparable utilities: The social welfare ranking must be independent of the way in which utilities of different individuals compare with each other.
- Pareto Preference Rule: If everyone has at least as much utility in x as in y, and if someone has more utility in x than y, then x is socially better than y.
- Pareto-inclusive welfarism: Social welfare is an increasing function of personal utility levels, thus satisfying both welfarism and the Pareto preference rule.

According to Sen's classification, "utilitarianism" is "the classic approach to welfare

economics and in the usual applications is combined with the use of interpersonally comparable and cardinal individual utilities." (Sen, 1982) Because of methodological critique such as Robbins's (1932), the assumption of interpersonally comparable utilities was replaced by ordinalism.

As explained by Leontief's "general consensus in public policy", the so-called "new welfare economics" accepted all these properties as legitimate. Social choice theory, pioneered by Arrow (1951), also accepted these conditions through the usage of somewhat weaker versions of the Pareto principle and welfarism.

The game theory framework in strategic situation normally adopts most of properties of "new welfarism", the concept of equilibrium point, however, does not guarantee the Pareto preference rule. Also, the ordinalism cannot be maintained in the mixed strategy of the game theory.

2. Limiting information may invite social injustice

According to the principle of "revealed preference", the utility can be defined only in terms of what is chosen irrespective of why it is chosen, or who chose it. (Sen, 1982) Therefore, the satisfaction of any desire is assumed to have some value in itself which must be taken into account in deciding what option should be taken.

There seems to be no doubt the strictness of common sense of justice has certain usefulness in limiting men's propensities to injustice and to socially injurious actions, but the welfarism theorists believes that to affirm this strictness as a first principle of morals is a mistake. Therefore, socially unjustifiable desires may be a cause of the social decision. Thus, limiting non-utility information may invite "the pleasure in other's deprivation" of liberty "which requires the violation of a principle to which he would agree in the original position". (Rawls, 1971)

2.1 Non-comparability of utilities makes it impossible to aggregating competing interests.

In aggregating conflicting interest of different persons, groups or classes, e.g., in

planning decisions, or in comparisons of national welfare in alternative or successive situations (Sen, 1973), the informational limitations of non-comparability are exceptionally telling. Indeed, many acts of political and social judgments, e.g., the personal decision as to what kind of a government or a society one should want, are themselves based on aggregating conflicting interests (Harsanyi, 1955), and in making these judgments, to be constrained by the informational base of non-comparable individual orderings would be peculiarly limiting.

2.2 Limiting information may render distribution of "social goods" unfair.

The severity of the information restrictions in the Arrow's framework can be illustrated by taking up a problem of income distribution judgment. As mentioned above, "revealed preference" is irrelevant from who chooses the preference, not only why it is chosen. These combined informational exclusions make it impossible to give priority to the interest of the poor in the exercise of aggregating the conflicting interests of the poor vis-à-vis the rich. The poor cannot be distinguished for this purpose from the rich – neither in terms of utility, nor in terms of income or other non-utility information.

If we assume, as utilitarian, that it is rational for one man to maximize his/her utility, it is right for a society to maximize the net benefit of the satisfaction taken over all of its members. As results, utilitarian does not matter how this sum of satisfactions is distributed among individuals any more than it matters, how one man distributes his satisfactions over time. The correct distribution in either case is that which yields the maximum fulfillment. (Rawls, 1971)

2.3 "Strategic" actions of players cannot be consistent with Pareto preference rule.

According to Arrow, his framework does not cover the issue of "strategic" behavior of players (von Neumann and Morgenstern, 1947). He only analyses "cooperative" situation in which agreements can be formulated to bind to all players, and which assumes that all players do not disguise their preference in decision-making process. (Arrow, 1971)

Moreover, the concept of Nash equilibrium [1](Nash, 1950) in the game theory framework, which is normally utilized in strategic situation, is not always consistent with the Pareto preference rule.

3. Benchmark for improvement: Application of "Rawls's principles of Justice" to address the deficiency of information limitation of welfarism.

Responding to the above deficiencies by information limitation, this thesis aims to propose a solution to address the problems. From a perspective of information to be used for decision making, Arrow's impossibility theorem can be regarded as the results of combining "welfarism" (ruling out the use of non-utility information) with remarkably poor utility information (especially because of the avoidance of interpersonal comparisons) (Sen, 1982). Therefore, there are two ways for solution. The first is dropping welfarism, which means utilization of non-utility information; the other is enrichment of utility information including the application of comparable cardinal utility.

To do so, the emerging problem is what is a "merkmal" or benchmark to evaluate improvements of the proposed solution. In this thesis, Rawls's "principles of justice", especially in generalized form, would be employed as it. Rawls is well known as the proposer of his "Difference Principle", however, this thesis are not willing to reiterate multiple critique on Rawls's theory from number of scholars. Rather, this thesis would focus on generalized "Difference Principle".

3.1 Rawls's Principles of Justice

Rawls's principles of justice characterize the principles for distribution of "primary social goods" (Rawls, 1971). There are "things that every rational man is presumed to want", including "right, liberties and opportunities, income and wealth, and the social bases of self-respect". The basic liberty has special distinctive priority among other

[1] Nash equilibrium is a solution concept of a game involving two or more players, in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing only his or her own strategy unilaterally (Nash, 1950)

primary social goods. The second principles, demanding equality and efficiency supplement of the first principles (Rawls, 1971), which is based on the idea of "original position"[2], Rawls defines the second principle as the second principle in generalized form at first as follows;

First: each person is to have an equal right to the most extensive basic liberty compatible with a similar liberty for others.

Second: social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone's advantage, and (b) attached to positions and offices open to all.

The first principle is arranged in a serial order prior to the second. This ordering guarantees that equal liberty "cannot be justified by, or compensated for, by greater social and economic advantage". Although the traditional decision theories guarantee the freedom of choice as preconditions of their frameworks, however, the utilitarian decision making weighs assets of individual and may gives a larger asset holder greater decision power than a one of smaller (this issue is argued in Chapter 3.) Besides, this thesis rather focuses on the second one, which we call it as "generalized differential principle" in accordance with the critique made by Arrow (Arrow, 1973), while the distribution of wealth and income need not be equal, but it must be advantageous for all.

---

[2] Rawls embodies his guiding idea of justice as the "original agreement". This is the agreement which "free and rational persons concerned to further their own interests would accept in an initial position of equality as defining the fundamental terms of their association".

This "initial position" is defined as "original position" which "corresponds to the state of nature in the traditional theory of the social contact". This original position is the postulated situation and "not, of course, thought of as an actual historical state of affaires, much less as a primitive condition of culture". In this situation, "no one knows his place in society, his class position or social status, nor does not any one know his fortune in the distribution of natural assets and abilities, his intelligence, strength, and the like".

Rawls argued that in "original position" people would not choose to maximize the utility sum, rather choose his two principles. Many critiques to the "original position" have emphasized that it is unclear what precisely would be chosen in such a situation. It is also not obvious that prudential choice under such uncertainty provides an adequate basis for moral judgment in unoriginal, i.e., real-life, positions. (i.e., Nagel, 1973) However, Rawls's "more direct critiques in terms of liberty and equality do remain powerful" (Sen, 1981).

In other words, unless the result of the distribution realizes benefit for everybody, such distribution would be condemned injustice.

Rawls modifies the generalized differential principle to the specific "differential principle" such that social and economic inequalities are to be arranged so that they are both:

(a) to the greatest benefit of the least advantaged, consistent with the just saving principles, and

(b) attached to offices and positions open to all under considerations of fair equality of opportunity

The differential principle should be lexically prior to the principle of efficiency (Pareto optimality) and to that of maximizing the sum of advantages; and fair opportunity is prior to the difference principle.

3.2 Formal modeling of "Differential Principle" and its generalized version.

Both the generalized and specific differential principle are yielded as the choice of the social contract in "original position". The major discrepancy of them is that the latter has clearer indication of egalitarian for the mitigation of gaps.

"Risk averse" is used as one of the reasons of differential principle by Rawls. It says that individuals tend to avoid the worst situation if there is significant uncertainty. The opposite assertion is done by Harsanyi and Arrow, which is individuals try to maximize expected values. (Harsanyi, 1975) (Arrow, 1973)

In this thesis, however, we examine the differential principle from the view point of information limitation. Rawls is said to insist the differential principle because it is simple and easy to apply to numerous actual decision situation. The concept of maximization or optimization can discover a solution if alternatives keep ordinance relations. Thus the decision can be done without bothering what is the situation prior to

the decision (the status quo).

3.2.1. Formal expression of the decision of Rawls's differential principle

Assuming that it is easy to discover and designate the least advantaged individuals pro forma without examining their social goods, based on the differential principle, the alternative of distribution which can maximize the social goods of the least advantaged individuals is required to satisfy the formulae as follows;

In 2-tuple social system $(A, f_{\min})$, the decision function F can be defined by

$F : (A, f_{\min}) \mapsto a^*$, where a^* is an alternative such that $f_{\min}(a^*) = \max_{a \in A} \{f_{\min}(a)\}$.

Where, let A be the set of alternatives, and $\{f_{\min}(a)\}_{a \in A}$ be the set of payoff (social goods) of the least advantaged individual.

3.2.2. Formal expression of the decision of generalized differential principle (everybody's advantage)

On the other hand, based on the "generalized" differential principle, social goods distribution needs to be identified in comparison with ones at a "reference point" for all individuals. This means that the concept of everybody's advantage needs more information of "a reference point" than differential principle.[3]

[3] Rawls is not quite clear to define the nature of social goods. Goods can be regarded as "asset" which is an accumulation of social good s received, or "flow" which is a continuous incoming of social goods based on a distribution policy. This thesis, however, assume that the social goods are "flow", because normally the poor and the rich are distinguished by an yearly income, and Rawls does not accept the asset inheritance from the view point of the eradication of social unfairness by inherited assets.

Formally expressed, in 4-tuple social system $\left(N, A, a_r, \{f_j\}_{j\in N}\right)$, the decision function is defined as

$$F:\left(N, A, a_r, \{f_j\}_{j\in N}\right)\mapsto a^*,$$

where a^* is an alternative such that $f_j(a^*) > f_j(a_r)$ for $j \in N$.

Where, A is the set of alternatives, $N = \{1,2,...,j,\cdots,n\}$ is the set of individuals, a_r is the alternative at the reference point, $\{f_j(a^*)\}_{j\in N}$ is the set of social goods of individual j under the alternative $a^*, \{f_j(a_r)\}_{j\in N}$ is the set of social goods of individual j in the reference point a_r.

3.2.3. Formal expression of the decision of differential principle, if there is not the least advantaged a priori.

Rawls argues that it is easy to discover the least advantaged individuals, because they are *a priori.,* However, in this industrial society, it is not easy to identify who is the actually least advantaged individual without making comparison of distributions of social goods inter-individually. In this case, the decision function has to be defined to discover the least advantaged as follows;

Formally expressed, in 4-tuple social system $\left(N, A, a_r, \{f_j\}_{j\in N}\right)$, the decision function is defined as

- $F^1:\left(N, a_r, \{f_j\}_{j\in N}\right)\mapsto j_{\min}$, where j_{min} is the least advantaged individual such that $f_{\min}(a_r) = \min_{j\in N}\{f_j(a_r)\}$ where $f_{\min}(a_r)$ is the payoff of j_{min} in the reference point a_r.
- $F^2:\left(A, f_{\min}\right)\mapsto a^*$, where a^* is an alternative such that

$$f_{\min}(a^*) = \max_{a \in A} \{f_{\min}(a)\}$$

Where, A is the set of alternatives, $N = \{1,2,...,j,\cdots,n\}$ is the set of individuals, a_r is the alternative at the reference point, $\{f_j(a^*)\}_{j \in N}$ is the set of social goods of individual j under the alternative a^*, $\{f_j(a_r)\}_{j \in N}$ is the set of social goods of individual j in the reference point a_r.

3.3 Critical evaluation of "difference principle" and its generalized form.

The necessary elements of information for the differential principle in Section 3.2.3 are as same as the ones for the "generalized" differential principle in Section 3.2.2. Moreover, the differential principle needs inter-individually comparable cardinal utilities to discover the least advantaged. Thus the differential principle needs more information than principle of "everybody's advantage", and as a result, the application feasibility of the former is limited than the latter.

In discussing a "strategic" decision-making, "the greatest benefit to the least-advantaged" is not feasible in strategic situation. In strategic situation, the player selects "a strategy" instead of an alternative. If the least advantaged select a strategy to maximize their payoffs, the payoffs depend on the other player's strategy and cannot guarantee the maximum distribution to the least. The decision function to achieve " the greatest benefit to the least-advantaged" needs to establish a "rule" to restraint others, which clearly violates the definition of "strategic situation" and characterizes cooperative games.

From the view point of the social justice, it is not natural that the decision making of "everybody's advantage" is socially injustice or unfair. This does not promise the mitigation of social gaps, but at least give everybody the better, which mitigate gaps as a result in long run from our experience in development of countries. Even for individual in the most risk averse would not object to give a better life to any social rank.

From the above reasons, we employ the "generalized" differential principle instead of specific one as the benchmark of improvement of the decision system, such as Arrow's framework, utilitarian cost benefit analysis and strategic analysis.

4. Proposal of this thesis: Realization of justifiable decision-making

This thesis aims to propose the improvement of new welfarism's framework which compounds its disadvantage by limiting information. Rawls's two principles of justice are employed as the "merkmal" or the benchmark for evaluation. As stated in Section 2, these principles are still controversial but reliable enough to be used as criteria for evaluation. The proposed methodologies of decision-making are evaluated with their fulfillment of Rawls's principles of justice.

For the proposal of improvement, this thesis rather focuses on utilization of non-utility information than enrichment of utility information. As Sen stated in his "liberal paradox", utilization of compatible cardinal utility can not avoid Arrow's impossibility theory. According to Sen, the way to free from the impossible theory, the boundary of the stakeholder's scope, or "protected sphere" have to be defined by the use of non-utility information. [4]

4.1. Proposal: Incorporating "status quo" into decision-making

The generalized differential principle requires that "everyone benefit from economic and social inequalities". To clarify "benefit", the initial arrangement prior to the decision has to be defined as a benchmark. This initial arrangement is called as "status

[4] The "impossibility of the Paretian libertarian" which I have presented elsewhere (Sen, 1970), relates closely to the difficulties with welfarism. The result shows the incomparability of the Pareto principle (even in the weak form) with some relatively mild requirements of personal liberty, for consistent social decisions, given unrestricted domain. The link with welfarism can be seen in the following way. Considerations of liberty require specification of non-utility information as relevant, e.g., whether a particular choice is self-regarding or not, or as falling within a person's "protected sphere". The claim is that this use of non-utility information goes not merely against welfarism, it can go even against Paretianism. (Sen 1970)

quo" in this thesis later on.

The status quo can only be defined as the situation of "social goods" prior to the decision. To identify what situation is status quo, we need to employ the concept of "time sequence" to differentiate the situation of "before" and "after" the decision.

In real-life decision making, it is rare not to have time sequence in decision process. Such decision is limited in the situation where the initial arrangement has ceased to exist such as the general election after parliamentary dissolution. Aside of it, status quo normally exist.

Status quo can also act as "disagreement point", if the decision cannot be reached. It is natural and reasonable to keep status quo until decision has done to alter it, as stated by Nash (Nash, 1950). In the real-life decision making process, incomplete decision frequently happen, and the preparation for it is unavoidable.

In formal expression, the proposal of "incorporating the status quo" introduce 5-tuple social system including "time sequence" such that $S = (N, A, \{f_j\}_{j\in N}, T, \{a_t\}_{t\in T})$, where, $T = \{-m, \cdots, t, -2, -1, 0\}$ is the set of retrospective timings which specify stages of decision. The status quo of each stage is defined as unique precondition of decision in each stage. Timing 0 identifies the stage in which the latest decision takes place, and specifies the status quo as the unique precondition in the latest stage. Then timing -1 identifies the decision stage which is one stage prior to the latest one (if any) and specifies the status quo in that stage. The concept of T is retrospective, and does not foresee a prospective situation. a_t is the alternative under the specific timing of t. This system has clear contrast to the 4-tuple system $(N, A, a_r, \{f_j\}_{j\in N})$ defined in Section 3.2, which incorporates "reference point" instead of "time sequence."

In comparison with the system in Section 3.2, the proposed social system can trace back the history of the decision. It opens the application to retrospective analysis of decisions, or step by step approach in prospective analysis.

4.2. Proposed decision rules: Prioritizing the status quo

This thesis proposes the following decision rules to make improvements of the traditional methodologies;

a) Incorporating concept of "time sequence" into the decision making framework. This could identify a precondition prior to a decision (status quo). The status quo should not be restricted the "initial" situation of the sequence, could be defined as preconditions of each decision stage in multi-stage approach.

b) Conflicting situation (such as situation which may not reach unanimity) should be included in the alternatives to be chosen. The status quo should be treated as disagreement point of the conflicting situation.

c) The distribution of "social goods" by the decision must satisfy "Pareto Optimal" for all stakeholders. Any single person who is deteriorated his social goods cannot be legitimate in the decision. This rule is particularly formulated to satisfy the principle of justice of Rawls.

The rule a) is formally defined in the above section. The rule b) is formally defined as the decision function such that $F^b : (N, A, \{f_j\}_{j\in N}, T, \{a_t\}_{t\in T}) \mapsto a^*$, where a^* is the alternative such that if there is a' which satisfies $f_j(a') \succ f_j(a_{sq})$ for all $j \in N$, then $a^*=a'$, else $a^* = a_{sq}$

The rule c) is formally expressed as the decision function such that $F^c : \{a^*\} \mapsto a^{**}$, where a^{**} is the alternative satisfy the following;

There is no a'' such that $f_j(a'') \succ f_j(a^{**})$ for all $j \in N$.

5. Organization of this thesis

This thesis consists of four chapters other than this chapter. The problem arises in Sec. 2.1 are discussed in Chapter 2. Deficiencies on distribution stated in Sec. 2.2 are analyzed in Chapter 3. The application of strategic actions into the decision appealed in Sec. 2.3 is attempted in Chapter 4. The summarization and conclusion of this thesis

appears in Chapter 5.

Reference

Arrow, K.J. (1951). Social Choice and Individual Values, 1st edition, Wiley, New York

Arrow, K.J. and Hahn, F. (1971). General Competitive Analysis, Holden-Day, San Francisco

Arrow, K.J. (1973). Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice, Journal of Philosophy Vol.70,9, pp.245-263

Arrow, K.J. and Raynaud, H. (1986). Social Choice and Multicriterion Decision Making, MIT Press, Cambridge

Harsanyi, J.C. (1975). Can the Maximin Principle Serve as a Basis for Morality? a Critique of John Rawls's Theory, American Political Science Review Vol.69, 2, pp. 594-606

Nash, J.F. (1950). The Bargaining Problem, Econometorica, Vol.18

Sen, A.K. (1970). Collective Choice and Social Welfare, Holden-Day, inc

Sen, A.K. (1973). On Economic Inequality, Clarendon Press, New York, Norton

Sen, A.K. (1982). Choice, Welfare and Measurement, Harvard University Press, Cambridge, Massachusetts

Leontief, W. (1966). Essays in Economics: Theories and Theorizing, Oxford University Press, New York

Robbins, L. (1932). An Essay on the Nature and Significance of Economic Sciene, Allen & Unwin, London

Rawls, J. (1971). A Theory of Justice, Clarendon Press, Oxford

Harsanyi, J.C. (1955). Cardinal Welfare, individualistic ethics and interpersonal comparisons of utility, Journal of Political Economy, 63: 309-321

von Neumann J. and Morgenstern, O. (1944). Theory of Games and Economic Behavior, Princeton University Press, Princeton

Chapter 2

Re-examination of Arrow's Framework by Restricted Use of Non-utility Information

Summary

"Arrow's general possibility theorem" (1963) has revealed that three basic democratic conditions, which consist of, "unrestricted domain" (U: each individuals has un-restricted alternatives.), "independence of irrelevant alternative" (I) and "Pareto optimality" (P), cannot consistent with the condition of "non-dictatorship". If collective choice is regarded as multiple criteria decision of a single person, Arrow's theorem also affirms impossibility of implementation of multi-criteria decision, although some psychological studies show opposite results. To resolve this issue, this chapter proposes a social welfare function which employs and prioritizes "status quo" , and relaxes the condition of "independence of irrelevant alternative". The main result of this chapter reveals that if the proposed social welfare function is applied to the model proposed, non-cyclic social orderings could be an output in accordance with conditions U, P and D. Although the relaxation of the condition I is inescapable, however, the use of such restricted non-utility function enables us to identify the social ordering in which the highly preferred alternative than the status quo shall be highly preferred by all criteria (individuals), than the status quo. Thus the ordering satisfies Rawls's "generalized" differential principle as well.

1. Introduction

When you have to choose a hotel you stay, how do you select it? Someone focuses on charges; the other person focuses ambiences or else. Most of you decide the hotel by utilizing multiple criteria, such as charges, location, ambiences, services, and so on. In this situation, what do you think if you are allowed to employ only a single particular criterion for your decision and cannot utilize other numerous indicators? That is what J.K. Arrow proved in 1950, and it is called as "Arrow's general impossibility theorem" (Arrow 1950, 1951, 1963).

Arrow's social welfare function is a particular type of collective choice rules that specifies orderings for the society as "output" by utilizing the information about multiple individuals' orderings as "input". An individual's ordering is assumed to be reflexive, transitive and complete.[1] We should note that the collective choice by Arrow's social welfare function specifies an "ordering" for the society, not only the "best alternative". A binary relation on the possible alternatives has to be determined by the function.

The above Arrow's framework generally accepts four conditions of "new welfarism." The first condition embodies "welfarism" such that any social state ranking must be based entirely on personal utility. Secondly, Arrow's framework does not only utilize "ordering" as the output of the social welfare function, but also exclusively use "ordinal properties" as the basement for social welfare judgement. Thirdly, Arrow strictly criticised the assumption of interpersonally comparable utility attributed to the "utilitarianism", and postulated the incomparability. Lastly, Pareto preference rule is weakly introduced in Arrow's framework such that if everyone has higher utility in x than y, x have to be socially better than y.

As stated above, if we cannot use non-utility information, all individual orderings or preferences have to be treated equal. For example, the ordering of Mr. A cannot be prioritized to the ordering of Mr. B, nor a specific relation among the alternatives in the ordering cannot be prioritized to the other.

[1] : "An ordering is a ranking of all alternatives vis-à-vis each other. Consider the relation "at least as good as." First, it must be "transitive," i.e., if x is at least as good as y, and y is at least as good as z, then x should be at least as good as z. Second, the relation must be "reflexive," i.e., every alternative x must be thought to be at least as good as itself. Third, the relation must be "complete," i.e., for any pair of alternatives x and y, either x is at least good as y, or y is at least as good as x (or possibly both)" (Sen, 1970) .

This limitation has great impact on aggregating conflicting interests of different persons, groups, or classes, e.g., in planning decisions, or in comparisons of national welfare in alternatives or successive situations (Sen, 1973). Actually, aggregating conflicting interests are the basement of many acts of political and social judgments, e.g., the personal decision as to what kind of a government or a society one should want (Harsanyi, 1955). However, it is the aggregation that the Arrow's general possibility theory stated as impossible.

Although the first principle of welfarism is not to evaluate values of utility or satisfaction, we hypothesize that it is possible to utilize non-utility information in restricted manner to evaluate the justification of specific utility, and thereby, we propose the social welfare function, which can decide the social ordering even if there is conflict in orderings among individuals.

1.1 Arrow's framework and Arrow's general impossibility theorem

Before we examine the proposal of the utilization of non-utility function, let us re-define the basics of the Arrow's framework and his theorem. In typical, X is the set of all alternatives or outputs. It has at least two members. We let $P(X)$ denote the set of orderings on X. The set N of individuals whose preferences are to be consulted is the finite set $\{1, 2,\ldots, n\}$ with n is more than 1. Then $P(X)^N$ is the set of ordering on X for the set of N. A domain P is some non-empty subset of $P(X)^N$, a member p of $P(X)^N$ is called a profile, and it assigns the ordering $p(i)$ to individual $i \in N$, where $p(i)$ is interpreted as i's preference ordering at profile p. A social welfare function for output set X and domain P is a function f from P into the set $P(X)$ of orderings on X.

The "Arrow's framework" has five characteristics as follows. (Campbell and Kelly 2002)

(1) The set X of alternatives is unstructured.

(2) There is a finite set *N* of "individuals", which is fixed. Typically, the members of *N* are different people, who have preference over the alternatives in *X*. However, May (1954) and Arrow and Raynaud (1986) regard *N* as decision criteria. For example, *X* is a set of hotels and the set *N* is the set of scheme of preference. For each attribute, the ranking reflects the tastes of a single individual. In this interpretation, we can analyze the single person's decision making which is based on more than 2 criteria. This interpretation of *N* has been used in many papers employing "multi-criteria decision making" (van Delft and Nijkamp, 1977) .

(3) Social choice is sensitive only to the ordinal property of individual preference. Specifically, the informational base for a social choice procedure is a family of profiles, where a profile is an assignment of a preference over *X* to each individual *N*. The input to the social choice procedure is a preference over *X* to each individual *N*.

(4) The output of a social choice procedure is an ordering of the alternatives in *X*, called a social ordering, rather than the selection of one or more members of *X*.

(5) Preference revelation is non-strategic. We do not consider what happens when "the ideals of the just society meet with the play of self interest" (Arrow 1997).

Arrow proved general possibility theorem in 1950, which consists of imposing certain conditions on a social welfare function *f*, and showing that these conditions are mutually incompatible. In other words, rule *f* which can satisfy condition U, P and I must be dictatorial. These conditions are as follows ;(Campbell and Kelly 2002)

Condition U (unrestricted domain): The domain of the rule *f* must include all logically possible combinations of individual orderings.

Condition P (weak Pareto principle): the rule *f* must satisfy the Pareto principle in the weak form, i.e. if everyone prefers *x* to *y*, then society must also prefer *x* to *y*.

Condition I (independence of irrelevant alternative): Social choice over a set of alternatives must depend on the orderings of the individuals only over those

alternatives, not on anything else, e.g., on rankings of "irrelevant" alternatives that are not involved in this choice.

Condition D (non-dictatorship): Rule *f* should not be dictatorial. In other words, there should be no individual such that whenever he prefers *x* to *y*, society must prefer *x* to *y*, irresponsible of the preference of everyone else.

1.2 Application of Arrow's Theorem to Multi-criteria Decision Making

If we regard *N* as the set of schemes of preference of single person, the interpretation of Arrow's theorem tells us that any person has to make his/her decision depending on only a single particular criterion. In other words, multi-criteria decision making cannot be done. For example, in case you have to decide the hotel to go, you cannot utilize combined information of charges and ambience, but have to be based on only charges. This conclusion is not acceptable instinctively, and also contradicts some psychological studies, mentioning that human can combine information from seven criteria. (Miller, 1956)

2. Proposal: Enabling multi-criteria decision making by incorporating non-utility information into decision-making

2.1 Rawls's "generalized" differential principle of justice and a prioritization of the status quo

Recalling Rawls's principles of justice explained in Chapter 1, the generalized differential principle , which requires "everyone's benefit," if economic and social inequalities are not avoidable (Rawls 1951). "Benefit" is normally interpreted as better than the situation prior to the decision, or "status quo." To respond these two requirements, this thesis proposed to employ the concept of Pareto optimality and "time

sequence", the sequence which enables to differentiate "before" and "after" of the decision to identify "status quo." Thereby, to satisfy "everyone's benefit," the situation accredited to the decision has to satisfy Pareto optimality for everyone compared with the status quo, unless that, the society should keep status quo.

2.2. An individual/criteria as a tie breaker

Even if the prioritization of the status quo is introduced, it is possible that there is an unresolved conflict between individual/criteria. In response to this, we propose the concept of a "tie-breaker". The concept is to prioritized binary relations in a specific individual / criteria among others. For designating the tie-breaker, we propose and utilize the concept of "common shared criteria".

To limit the universal domain and keep away from the Arrow's theorem, the concept of "single peakedness" has been proposed (Arrow 1951). This concept postulates that all preferences of alternatives can be on the single axis of value, such as "right wing" or "left wing" of the political behaviour, and constrains to individuals to have only one "peak" of their preference on that axis (selection of peak is free for individual).

As an extension of the concept of "single peakedness," we would like to propose the concept of "common shared criteria," in which all individuals share not only the axis of value of a criterion, but also a ranking of the preference in the criterion. In this chapter, the item of "hotel charge" is employed as the example of this value. In this value, even if the unit of the indicators (e.g. $, Yen, Euro, etc) are transformed, the ordering itself does not change, i.e., charges of Hotel A are more expensive than those of Hotel B, whatever the customers use USD, JPY, GBP or EUR. Therefore, the order of the criterion 1 cannot prefer B to A whoever the person employs this criterion, unless the customers like better expensive than inexpensive in general.

According to its "common shared" nature, it is reasonable that the criterion which consists of "common shared criteria" is prioritized than other criteria.

2.3 Proposed Social Welfare Function

Incorporating the concepts listed in 2.1 and 2.2, let me propose the social welfare function as follows. The decision process from (I) to (III) in order.

(I) If every criteria (every individual) prefers x to y, then society shall also prefer x to y (Condition P (weak Pareto principle)),

(II) If there is a conflict of ordering between criteria (individuals), i.e., criterion 1 prefer x to y, and criterion 2 prefer y to x, and if x is a certain social state (the "status quo"), the society shall prefer x to y, and

(III) If there is a criterion designated as a "tie-breaker", the criterion shall be select as a "prioritized criterion". The ordering of the prioritized criterion shall be prioritized than others.

In this case, the society is assumed not to select the criterion by which the status quo is located in the lowest order as "common shared" criterion. Because such a criterion appeals strong dissatisfaction to the status quo, and thereby it is moderate enough not to be commonly shared by all individuals.

As for rule (I), if at least one individual strictly prefers x to y, and every individual regards x to be at least as good as y, then the society should prefer x to y. This criterion has an obvious appeal.

As for rule (II), in predicting the actual output of a conflict of the preference, the status quo is clearly relevant, for it defines what will happen in the absence of the parties agreeing to a cooperative solution, as stated in Nash's "bargaining problem" (Nash 1950).

As for rule (III), if the conflict could not be resolved even if priority is given in status quo, the "common shared criteria" can be a final tie-breaker.

2.4 Formal expression of the traditional and the proposed social welfare function

Let us assume the social system $S = (N, A)$. In this system, the social welfare function which aggregates individuals'/criteria's into a single preference order (social ordering) on A is defined as;

$$F : L(A)^N \rightarrow L(A)$$

Where, $A = \{a \in A\}$ is the set of alternatives, $N = \{1,2,..., j, \cdots, n\}$ is the set of individuals or decision criteria, *L(A)* is the set of all full linear orderings of A (the set is equivalent to the set of permutations on the elements of A).The n-tulpe $(R_1, R_2, \cdots R_n)$ of individual's or criteria's preference is called a "preference profile". The decision of the social welfare function is expressed as;

$F(R_1, R_2, \cdots R_n) = R^*$, where, R^* is a "social ordering."

Where, $(R_j)_{j \in N} \in L(A)^N$ and $R^* \in L(A)$

In contrast, the proposed alternative, the modified social welfare function is formally expressed as the social system $S = (N, A, T, \{a_t\}_{t \in T})$. In this system, let F' denote the social welfare function such as;

$$F' : L(A)^N \times A^T \rightarrow L(A).$$

Where, $T = \{-m, \cdots, t, -2, -1, 0\}$ is the set of retrospective timings which specify stages of decision. The status quo of each stage is defined as unique precondition of decision in each stage. Timing 0 identifies the stage in which the latest decision takes place, and

specifies the status quo as the unique precondition in the latest stage. Then timing -1 identifies the decision stage which is one stage prior to the latest one (if any) and specifies the status quo in that stage. The concept of T is retrospective, and does not foresee a prospective situation. a_t is the alternative under a timing t. The decision of the social welfare function is expressed as;

$$F'((R_1, R_2, \cdots R_n), \{a_t\}_{t \in T}) = R'^*, \text{ where } R'^* \in L(A)$$

3 The Model

The following decision-making models are formulated to examine whether our proposal can enable multi-criteria decision making, and can satisfy Rawls's second principle of justice. The first examination is correspondent to find a "social welfare function" which is not constrained by Arrow's impossible theory.

In the model we use in this chapter, we employ single-stage decision model such that the set of timing $T = \{0\}$ for simplification. We also denote X as a set of hotels X={A,B,C}, and the set N={1,2} as the set of scheme of preferences denoting charges and ambiences, respectively. Criterion 1 orders the hotels according to their charges; criterion 2 orders them by ambiences as shown in Table 1. A single person has social welfare function f such that determines the ordering of hotels based on the preferences of each criterion.

Table 1: Ranking of preference in each specific criterion

1: hotel charges	2: ambiences
Hotel A ($100)	Hotel B

Hotel B ($200)	Hotel C
Hotel C ($ 300)	Hotel A

You are assumed to be staying in Hotel B, and want to decide whether you should move to Hotel A or Hotel C, or keep staying Hotel B. That is, Hotel B is regarded as the status quo (SQ).

3.1 Adaptation to Aforementioned Conditions

The proposed social welfare function meets the Condition P clearly. The function also satisfies the Condition U, because an identification of the "status quo" is only a "marking" of alternatives, and does not restrict any selection of orderings in each criterion.

To examine the adaptation to the Condition D, let me show the application of the rule to the model shown in Table 1.As the first step, binary relations to meet weak Pareto principle are identified. B (status quo) is preferred to C in both criteria, thus the society must prefer B (status quo) to C. As for the relation between A and B, there is no consensus among criteria (individuals). Therefore, rule (II) is applied, the society must prefer B to A. Lastly, two criteria have conflict over the binary relation between A and C, rule (III) stipulates us to prefer A to C. Through these steps, the social ordering appears as $B \succ A \succ C$. This result clearly remarks there is no dictator in this decision-making. Also, this relation of three alternatives complies with reflexivity, transitivity and completeness. This result can be described as Table 2. The decision process for all possible combinations is shown in Table A in ANNEX.

It is obvious that, if the rule (II) does not exist, the social ordering equals to the ordering of the criterion 1 and the criterion 1 becomes a dictator. The rule (II) plays a key role to keep the proposed social welfare function away from dictatorship.

Table 2: Decision Process of the proposed social welfare function (B is regarded as SQ)

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	B	B > C	B> A	A> C	B
B	C				A
C	A				C

The function, however, contradicts to Condition I. The essence of Condition I (independence of irrelevant alternatives) is that any binary relation of a social ordering has to be decided based only on binary relations of each individual (criterion). The rule II intervenes in the decision of the social ordering by prioritizing the status quo. The decision process showed in Table 3 explains this contradiction clearly. The ordering of the two criteria in Table 2 and 3 is the same, except the location of SQ in each table (alternative B is regarded as SQ in Table 2, and so as A in Table 3). The social orderings of tables, however, are different. (See Table A in ANNEX for decisions in all possible combinations)

Table 3: Decision Process of the proposed social welfare function (A is regarded as SQ)

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	B	B > C	A> B	A> C	A
B	C				B
C	A				C

For discussion, let me show that, in the decision processes when alternative C is regarded as SQ (see Table C in ANNEX), there is one "cyclic" ordering, which unable to determines a social ordering. This situation, however, cannot happen because the society does not accept the criterion as "prioritized criterion" by which the SQ is allocated in the lowest order of alternative, as stated in Rule (III).

4. Conclusion

The analysis above reveals that if the proposed social welfare function is applied to the model which comprises three alternatives and two criteria, the multi-criteria decision model can output the non-cyclic social orderings in accordance with conditions U, P. Because multi-criteria itself denies the condition D, Rawls's first principle of justice (preservation of individual liberty) is satisfied.

Moreover, the restricted non-utility function enables us to identify the social ordering which can satisfy Rawls's generalized differential principle. The highly preferred alternative than the status quo in the ordering shall be preferred by all criteria (individuals) than the status quo. Thus, the ordering is consistent with "everybody's advantage."

On the other hand, the use of non-utility information such as "status quo" or "commonly shared preference" have to make the condition I to be relaxed. However, this proposal does not relax the condition I arbitrarily, but limit to the utilization of non-utility information such as "status quo" or "tie-breaker criteria". This non-utility information exists generally, but only in strictly limited conditions. In most cases, a decision has a pre-decision situation or status quo. The decision which does not have pre-situation is

limited such as an election of candidates. Thus the information of status quo is widely applicable in almost all decisions.

Since this chapter examines limited situation, the study to seek further generalization is desired.

Reference

Arrow, K.J.(1950). A difficulty in the concept of social welfare, Journal of Political Economy 58:328-346

Arrow, K.J. (1951). Social Choice and Individual Values, 1st edition, Wiley, New York

Arrow, K.J. (1963). Social Choice and Individual Values, 2nd edition, Wiley, New York

Arrow, K.J. and Raynaud, H. (1986) . Social Choice and Multicriterion Decision Making, MIT Press, Cambridge

Arrow, K.J. (1997). Introduction, Social Choice Re-examined, eds by Arrow, K.J., Sen, A.K. and Suzumura, K. pp.xvi-xvii, Macmillan, London

Campbell, D.E., Kelly, J.S. "IMPOSSSIBILITY THEOREMS IN THE ARROVIAN FRAMEWORK, Handbook of Social Choice and Welfare, Vol.1, Edited by Arrow, K.J., Sen, A.K., Suzumura, K. Elsevier Science B.V., 2002 pp. 36-85

Gaertner, W., "DOMAIN RESTRICTIONS", Handbook of Social Choice and Welfare, Vol.1, Edited by Arrow, K.J., Sen, A.K., Suzumura, K. Elsevier Science B.V., 2002 pp. 132-165

May, K.O. (1954). Intransitivity, utility and the aggregation of preference patterns, Econometorica Vol.21:1-13

Miller, G.A., (1956) The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information, Psychological Review, Vol.63, No.2, pp.81-97

Nash, J.F. (1950). The Bargaining Problem, Econometorica, Vol.18

Sen, A.K., (1970) COLLECTIVE CHOICE AND SOCIAL WELFARE, Holden-Day, inc

van Delft, A.and Nijkamp, P. (1977) "MULTI-CRITERIA ANALYSIS AND REGIONAL DECISION-MAKING", H.E. Stenfert Kroese B.V., Leiden.

ANNEX

Table A: Decision process of the proposed social welfare function in all possible combinations. (The alternative B is regarded as SQ.)

Case 1

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	A	A > B			A
B	B	B > C			B
C	C	A > C			C

Case 2

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	A	A > B	B > C		A
B	C	A > C			B
C	B				C

Case 3

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	B	A > C	B> A		B
B	A	B > C			A
C	C				C

Case 4

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	B	B > C	B> A	A> C	B
B	C				A
C	A				C

Case 5

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	C	A. > B	B > C	A > C	A
B	A				B
C	B				C

Case 6

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	C		B > A	A > C	B
B	B		B > C		A
C	A				C

Table B: Decision process of the proposed social welfare function in all possible combinations. (The alternative A is regarded as SQ.)

Case 1

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	A	A > B			A
B	B	B > C			B
C	C	A > C			C

Case 2

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	A	A > B	B > C		A
B	C	A > C			B
C	B				C

Case 3

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	B	A > C	B> A		B
B	A	B > C			A
C	C				C

Case 4

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	B	B > C	A> B	A> C	A
B	C				B
C	A				C

Case 5

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	C	A. > B	B > C	A > C	A
B	A				B
C	B				C

Case 6

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	C		A > B	A > C	A
B	B		B > C		B
C	A				C

Table C: Decision process of the proposed social welfare function in all possible combinations. (The alternative C is regarded as SQ.)

Case 1

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	A	A > B			A
B	B	B > C			B
C	C	A > C			C

Case 2

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	A	A > B	C > B		A
B	C	A > C			C
C	B				B

Case 3

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	B	A > C	B> A		B
B	A	B > C			A
C	C				C

Case 4

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	B	B > C	C> A	A> B	-
B	C				-
C	A				-

Case 5

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	C	A. > B	C > B	A > C	A
B	A				C
C	B				B

Case 6

Criterion 1	Criterion 2	Rule I	Rule II	Rule III	Social Ordering
A	C		C > A	A > B	C
B	B		C > B		A
C	A				B

Chapter 2: Re-examination of Arrow's Framework

Chapter 3

A Critical Review of the Traditional Methodology of Cost Benefit Analysis and a Proposed Alternative

Summary

The traditional Cost benefit analysis (CBA), which is deeply rooted in welfarism, in particular, utilitarianism, has long been the preferred methodology for evaluating economic factors in policy making. The traditional CBA is also regarded as a social welfare function which can designate a "social ordering" of all alternatives. This chapter proposes incorporating the concept of "time sequence" and "status quo" into the traditional CBA framework whose guiding concept is the "Pareto optimization" of outcomes for the stakeholders. This chapter further proposes Rawls's principles of justice, in particular, "generalized" differential principle which requires everybody's advantage as evaluation criteria for CBA, and thereby critiques the traditional and alternative variants against this standard, ultimately to demonstrate the superiority of the latter. The major findings of this chapter reveals that the alternative can guarantee an advantageous situation compared to the status quo for every stakeholders as well as the alternative can satisfy Rawls's principle of everybody's advantage, which is the benchmark of a social decision-making.

1. INTRODUCTION

The traditional CBA can determine which alternative can produce the maximum sum of utility for the society questioned by confirming whether or not the "benefit" of a given regulatory option is greater than its "cost." The traditional CBA has long been the widespread methodology as the collective choice rule (CCR) to identify an alternative which maximizes the sum of utility compared with the status quo, in accordance with "utilitarian" approach rooted in J. Bentham. The traditional CBA can also specify the sum of utilities on each alternative, therefore, the traditional CBA is also regarded as a social welfare function which can designate a "social ordering" of alternatives.

The fundamental critique to utilitarianism, which is stated by Rawls, is that "the satisfaction of any desire has some value in itself which must be taken into account in deciding what is

right, ” even if the satisfaction is illegal, unethical, immoral or injustice. (Rawls 1971) The part of this issue is discussed in Chapter 2.

The other major critique to the utilitarianism or the traditional CBA is the issue of utility distribution, which this chapter is mainly concerned with. The correct distribution in utilitarian approach is that which yields the “maximum fulfillment” and not any more than that. The utilitarian “does not matter, except indirectly, how this sum of satisfactions is distributed among individuals any more than it matters, except indirectly, how one man distributes his satisfactions over time.” (Rawls, 1971)

In the approach above, the benefit of the weak, who has small amount of utility, is likely to be outweighed by the benefit of the strong, who has large amount of utility. As a result, the alternative which is not beneficial for the weak can be the best choice for “maximum fulfillment.” This result clearly violates Rawls’s “the second principle of justice,” the principle that “social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone’s advantage, and (b) attached to positions and offices open to all.” (Rawls 1971)

In addition, the traditional CBA which is in accordance with the utilitarian approach can only analyze cardinal and interpersonal comparable utility. This assumption was widely criticized such as by Robbins (1932), and its difficulty of application especially appeals on the comparison between the value of life and the value of goods.

1.1 The Purpose of this Chapter

This chapter aims to propose the alternative CBA methodology which complies with the followings;

a) Incorporating concept of “time sequence” into CBA framework. This could identify the situation before the decision (status quo). The status quo should not be restricted the “initial” situation of the sequence, could be defined as the situation prior to the “last”

decision.

b) The distribution of "social goods" by the decision must satisfy "Pareto Optimal" compared to the status quo for all stakeholders. Any single person who is deteriorated his social goods cannot be legitimate in the decision. This distribution can also be consistent with Rawls's principles of social justice, in particular the "generalized" differential principle.

In doing so, this chapter proposes an alternative CBA that attempts to "optimize the payoffs" to any organization, governmental entity, or individual that has a stake in or may be impacted by a given approach to regulation (here in after, I call it as "stakeholders"). This paper thereby critiques the use of traditional CBA as the methodology of choice for making decisions in policy. The traditional and alternative methodologies are appraised in terms of evaluation criteria for enhancing CBA performance.

1.2 The Organization of this Chapter

In this chapter, I employ the decision of environmental regulation as the field of application. Mathematical definitions of the traditional and proposed alternative methodologies and an analysis of their respective strengths and weaknesses follow. I propose the evaluation criteria for which the stakeholders can evaluate the usefulness of CBA as it is variously applied in actual risk management. In particular, I formulate a simple model based on the hypothetical situation where the government aims to adopt the new regulation and demonstrate how the traditional and proposed alternative CBA are applied in actual decision making. Finally, both are critiqued in terms of the aforementioned criteria to which CBA should subscribe.

2 RISK ASSESSMENT FRAMEWORKS

2.1 Risk Assessment for the Protection of Human Health

The central goal of regulations for protecting human health is the reduction of risk. To

this end, risk assessment is carried out within the overarching framework of risk management. Specifically, risk assessment attempts to ascertain the required reduction of exposure to potentially hazardous substances in the environment in order to identify a tolerable level of risk. This is accomplished by "the use of the factual base to define the health effects of exposure of individuals or populations to hazardous materials and situations."(US NAS/NRC. 1983) In this paper, carcinogens serve as the model target of regulation.

Risk assessment for cancer in humans takes into account the exposure to known carcinogenic agents, the dose absorbed by the body, and the cancer-slope factor (i.e., potency) that is based on the dose-response relationship for various types of cancer. From the analysis of these factors, risk assessment can specify the probabilities of contracting cancer. Therefore, this technique calculates how much reduction of exposure is required to achieve a certain reduction of risk. (James *et al*. 2000)

2.2 Functions of Risk Management and CBA

Risk management is a means by which government agencies evaluate various regulatory options and choose among them. It utilizes information generated by risk assessment and integrates relevant "political, social, economic and engineering information in[to] the decision process."(Beck *et al*. 1994)

Various studies show that the perceptions of people, who know that they are being subjected to risk, are normally not influenced by economic factors (Wilson and Crouch 2001), although difficulties arise in applying expected utility model for cases where discrete catastrophic financial impact affect individual survival. (Roy, 1952) Therefore, unless the financial impact is so severe, within the context of risk management, CBA is mainly used to justify the imposition of regulations on enterprises.

2.2.1. De minimus risk and de manifestus risk, and ALAPR

Countries such as the United Kingdom not only employ the concepts of de manifestus risk and de minimus risk but also define the decision criteria for their application. The

former represents the maximum threshold of risk to which humans may be subjected. Risk levels that reach this maximum cannot be justified on any grounds. The latter represents the minimum threshold of risk exposure, and, therefore, risks satisfying this criterion are broadly acceptable. The range between them is designated the ALARP (as low as reasonably acceptable). Accordingly, risk "is tolerable only if the risk reduction is impractical or if its cost is grossly disproportionate to the improvement gained." (Marszal 2001) The choice of regulatory options within this range usually relies upon CBA, the traditionally preferred methodology of economic analysis.

Most countries, including the United States and Japan, do not define de minimus risk and de manifestus risk so rigorously. The broad consensus holds that formal decision criteria "cannot explicitly reflect uncertainty about risks, population within, variation in susceptibility, community preferences and values, or economic considerations." (Presidential Commission on Risk and Risk Management 1997) However, it appears that most federal regulatory decisions in the U.S. do invoke the ALARP criterion in conjunction with CBA.

A survey of 132 U.S. federal regulatory decisions involving environmental carcinogens to which large populations were exposed showed that regulatory action was always undertaken if the risk was greater than 10^{-4}. (Travis *et al.* 1987) This value represents de manifestus risk, and for a small population, it rises to 10^{-3}. Furthermore, if the risk level for a large population was less than 10^{-6} to 10^{-7} and for a small population less than 10^{-4} to 10-5, no regulatory action was taken. These values represent de minimus risk.

The foregoing results strongly suggest that U.S. regulatory authorities apply CBA inexplicitly, that is to say, the tolerable risk level for a small population is greater than that for a large population. However, this outcome is all quite rational in view of the logic of traditional CBA: the smaller the population, the smaller the net benefit to be realized from a given regulation; in other words, relax the tolerable level of risk in order to hold down the cost of regulation.

3. THE EVALUAITION CRITERIA FOR CBA

The evaluation of CBA methodology should be done from multiple aspects. Each stakeholder has different stakes to the regulation; therefore each stakeholder's focus of evaluation has diversity. The following characteristics are selected as the evaluation criteria that cover all points of view from various stakeholders.

a. Decision-making by means of CBA should be advantageous to all stakeholders.

As stated by Rawls's "generalized" differential principle, distribution of payoffs should be "everyone's advantage".

b. Within the context of public decision-making, the stakeholders under all circumstances should be able to influence the outcomes of CBA.

All stakeholders should be able to affect the decision-making process regardless of the scale or magnitude of their respective interests. This is related to the Rawls's first principle of justice, which guarantees "equal right to all". This criterion is especially important in the case of the smaller or weaker stakeholders.

c. The CBA should void the use of inter-individual comparable cardinal utility for valuing cost and benefit.

It is desirable if the CBA methodology does not need a unit conversion for valuing cost and benefit. For the use of inter-individually comparable cardinal utility for valuing cost and benefit, the unit conversion from risk base unit (i.e. incident/year) to monetary unit (i.e. dollar or yen) is not easy process in many cases; the value of the rich tends to overestimate the one of the poor. Thus the unit conversion is likely to violate the Rawls's equal right principle.

4. AN EVALUATION OF THE TRADITIONAL CBA METHODOLOGY

An examination of the utility of the traditional CBA methodology for measuring economic factors in risk management follows.

4.1 A Definition of Traditional CBA

CBA is the umbrella term for those forms of economic analysis that directly compare the costs and benefits of regulatory actions. In many of its documented applications,

traditional CBA employs a decision rule that confirms whether or not net benefit (gross benefit minus total cost) is greater than 0, or a decision rule that confirms whether or not gross benefit divided by total cost is greater than 1. (Oka 2003) From the utilitarian view, the latter is the decision rule most commonly used, and, therefore, is employed in the analysis of traditional CBA.

Definitions of the decision rules of the traditional CBA methodology

With respect to payoffs of the stakeholders, the decision rules of the traditional methodology are defined as follows:

Rule: the decision selects the *l* such that the one that maximizes the sum of the payoffs of all stakeholders

In formal expression, the above rule is the decision function in the social system is defined as the 3-tuple system as follows:

$$S = \left(L, N_r, \left\{f_j\right\}_{j \in Nr}\right)$$

Where, let $L = \{0 \le l \le 1\}$ be the feasible set of the regulatory level, $N_r = \{1,2,...,n\}$ be the set of elements that are influenced by the regulation (stakeholders), and f_j is the payoff function, and $f_j(l)$ be the payoff of each $j \in N_r$.

The decision function F is formally defined as;

$$F : \left(L, N_r, \left\{f_j\right\}_{j \in Nr}\right) \mapsto l^*.$$

Where, l^* is the laudatory level such that

$$\sum_{j \in N} f_j(l^*) = \max_{l \in L} \sum_{j \in N} \left\{f_j(l)\right\}.$$

4.2 Major Deficiencies of the Traditional CBA Methodology

In many countries that employ risk management, the traditional methodology is not rigorously applied (Marszal, 2001). The main reason for this shortcoming is that the results of economic analyses have a negligible impact on the risk perceptions of persons who are exposed to environmental hazards. In concise terms, the relaxation of a regulation based on the results of traditional CBA is usually regarded as a compromise that redounds only to the benefit of the enterprise. Such an outcome is likely to be widely criticized and, therefore, ultimately to make the level of risk much more difficult to decide.

However, the non-rigorous application of the traditional methodology is also due to its intrinsic weaknesses. These are mathematically described as follows:

Proposition T-1: It is possible that traditional CBA decision-making might result to certain stakeholders in payoffs which are less than the status quo.

According to the decision rule, the traditional methodology must accept the new regulatory level if it maximize the sum of the payoffs. However, this decision cannot guarantee positive net-payoffs to every stakeholder, because the decision rule above pays attention only to the sum of the gross payoffs, but distributions to each individual. This fact clearly violates Rawls's generalized principle, everybody's advantage. Although the methodology may uphold in general terms the rationality of a social system's regulatory environment, it might not be able to sustain the rationality of any or all of a system's stakeholders, and thereby a consensus among all stakeholders may become impossible.

Proposition T-2: If a certain stakeholder has exceptionally large net-payoffs, the traditional CBA methodology might overlook those stakeholders that have relatively small payoffs.

Where a certain stakeholder has very large net-payoffs, stakeholders having relatively

small profits are likely to be overlooked by the traditional methodology. (Mathematical proof is in APPENDIX.)

This result clearly violates the Rawls's first principle of justice, the principle of equal right. The deficiencies are highlighted in the following example. Where the immediate region surrounding a coal-fired power plant is sparsely populated, a disproportionately large cost is incurred by the enterprise to purify the airborne effluents. According to the traditional CBA decision-making, the persons living in close proximity to the plant are forced to endure a greater health risk than if the region were densely populated. Furthermore, if the de manifestus risk level is not precisely determined, the decision might indicate that, in extreme cases, regulations protecting human health are not cost-effective for very small populations.

Proposition T-3: The traditional CBA methodology needs a consistency of unit for all stakeholders' payoffs for calculating the net benefit.

The outcome or effect of the regulatory option appears in variable fields in social system and is expressed in various unit, such as monetary bases (i.e. dollar, yen or euro), or risk bases (i.e. incident/year, life/year and so on). However, the traditional CBA need a consistency of unit to sum up all payoffs of stakeholders. Therefore, unit conversion is required to be carried out. From the perspective the quality of information, the traditional CBA requires inter-individual comparable cardinal payoffs, which is hardly criticized by many scholars such as Arrow. Moreover, the value of the rich tends to overestimate the one of the poor. Thus the unit conversion is likely to violate the Rawls's equal right principle.

5. AN ALTERNATIVE CBA METHODOLOGY

In spite of the importance of economic considerations in risk management, the traditional methodology has deficiencies. Therefore, I propose an alternative CBA methodology, which is intended to enhance and strengthen the decision-making procedure.

Under the traditional approach, the summation of net payoffs is employed as the decision criterion. Conversely, the proposed alternative approach selects regulatory options that realize "the optimization of the list of payoffs." Specifically, the payoffs to stakeholders are calculated when a new regulatory level comes into play. Thereupon, the alternative procedure chooses the particular payoff in the list that meets the "Pareto optimality." The Pareto optimality (PO) is the condition such that "no one can be made better off without someone else being made worse off".(Rutherford 1995) If there are certain lists of payoffs that satisfy PO, this procedure opts for that list which maximizes the multiplicative product of all payoffs.

5.1 The Decision Rules of the Alternative CBA Methodology

The decision rules of the alternative approach are described as follows: Formula (1) includes the set of elements of the model social system which is subject to regulatory decisions. Therefore, the alternative CBA methodology is formally defined as a procedure which selects the list of payoffs engendered by a new level of regulation that satisfy the following criteria:

Rule 1: The list of payoffs to all stakeholders must satisfy PO.

Rule 2: The payoffs to all stakeholders must greater than the payoffs in status quo. (The net-payoffs of all stakeholders must be greater than 0.)

Rule 3: If there are two or more lists of payoffs that satisfy the first two criteria, then the list of payoffs that maximizes the product of all payoffs must be chosen.

The Rule 1 gives us efficient results which are guaranteed by the concept of Pareto optimality. The Rule 2 is required to satisfy the Rawls's generalized differential principle which emphasizes "everybody's advantage". The Rule 3 is necessary to preserve the Rawls's first principle of justice or "equal right" principle. The advantages of the proposed alternatives are formally proved by the propositions in section 5.2.

5.2 Formal expression of the alternative CBA.

In the alternative CBA, the social system is defined as the 5-tuple system of

$$S' = \left(T, L, N_r, \{f_j\}_{j \in Nr}, \{l_t\}_{t \in T}\right).$$

Where,

$L = \{0 \le l \le 1\}$: The feasible set of the regulatory level,

$N_r = \{1,2,...,n\}$: The set of elements that are influenced by the regulation (stakeholders),

$T = \{-m, \cdots, t, -2, -1, 0\}$ is the set of retrospective timings which specify stages of decision. The status quo of each stage is defined as unique precondition of decision in each stage. Timing 0 identifies the stage in which the latest decision takes place, and specifies the status quo as the unique precondition in the latest stage. Then timing -1 identifies the decision stage which is one stage prior to the latest one (if any) and specifies the status quo in that stage. The concept of *T* is retrospective, and does not foresee a prospective situation. In this model, we employ a single stage model such that $T = \{0 = sq\}$.

$\{l_t\}_{t \in T}$ is the regulatory level under a specific timing.

$\{f_j \mid f_j(l) = payoff\}_{j \in Nr}$: The set of payoff function f_j of each $j \in N_r$.

5.2.1. Definitions of benefit and cost

In most risk-assessment scenarios, "benefit" connotes a "reduction of risk" (Oka, 2003). Since "reduction" implies how much a regulation might diminish the risk level of the status quo, it follows that CBA measures the "difference" (or "gap" between the payoffs under the status quo and the payoffs that follow the implementation of a regulation.

In mathematical terms, this "difference" is described as follows: *lsq* is the strength of the regulatory option that is imposed on the status quo. If a hypothetical social system is

subjected to a new regulatory level *l*, the system will change as it adjusts to *l*. The payoffs to the stakeholders will change as well.

The "benefit" and "cost" are defined in general terms as the difference (or gap) between the payoffs under the status quo and the payoffs after the regulation is implemented. When a certain regulation $l \in L$ is put into play, the net payoffs to a stakeholder $j \in Nr$ are defined as

$$NetPayoff_j = f_j(l) - f_j(l_{sq})$$

If the net payoffs are a positive number, a "benefit" is realized. Conversely, if the net payoffs are a negative number, a "cost" is incurred. The "benefit" and the "cost" are computationally defined as follows:

$$Benefit_j = NetPayoff_j = f_j(l) - f_j(l_{sq}), \textbf{ if } \; f_j(l) - f_j(l_{sq}) > 0$$

$$Cost_j = \left|NetPayoff_j\right| = \left|f_j(l) - f_j(l_{sq})\right|, \textbf{ if } \; f_j(l) - f_j(l_{sq}) < 0$$

To assess the change in payoffs to all stakeholders $j \in Nr$ that results from the implementation of a new regulatory level $l \in L$, the net benefit (NB) is defined as the summation of the net payoffs to all stakeholders or

$$NetBenefit\,(NB) = \sum_{j \in Nr} f_j(l) - \sum_{j \in Nr} f_j(l_{sq}) = \sum_{j \in Nr} \left\{f_j(l) - f_j(l_{sq})\right\} = F(l)$$

5.2.2. Formal expression of the decision rules of the alternative.

The first rule implies that a certain stakeholder's attempt to improve its payoff results in the reduction of payoffs to the other stakeholders. That is to say, when a certain payoff list meets PO, there is no better list than that. Therefore, choosing the Pareto optimal list maximizes the efficiency of payoffs to all stakeholders. The payoffs under PO are described as follows: If *l** satisfies the following condition, the list of payoffs for *l** satisfies PO.

For all $l \in L$**, there exists a** $j \in Nr$ **such that** $f_j(l) \leq f_j(l*)$**.**

Or

There does not exist an $l \in L$ **such that** $f_j(l) > f_j(l*)$ **for all** $j \in Nr$**.**

The second rule attempts to guarantee that the payoffs to all stakeholders will always be more than the payoffs in status quo. This rule is set because enforced levels of regulation that compel releasers of hazardous substances to incur disadvantageous results have little or no rational basis for persuading them to bear such costs. The second rule is mathematically described as follows:

$f_j(l^*) > f_j(l_{sq})$ **for all** $j \in Nr$

The third rule intends for the regulatory analyst to employ those decision criteria which select the list of payoffs that maximizes the product of all payoffs, provided that two or more lists satisfy PO. Specifically, if there are two or more *l**s that satisfy Pareto optimality, then the *l** which maximizes the product of payoffs is chosen. Mathematically, invoking this rule is equivalent to choosing the $l^{**} \in L$ which satisfies the following Formula:

$$\prod_{j=1}^{n} f_j(l^{**}) = \max_{l^* \in L^*} \prod_{j=1}^{n} f_j(l^*)$$

Where, $L^* = \{l^* \in L \mid f_j(l^{**}) > f_j(l^*) \text{ for all } j \in N\}$

5.3 The Advantages of the Alternative CBA Methodology

The advantages of the alternative methodology are assessed in terms of the deficiencies of the traditional methodology (see Section 3.2).

Proposition A-1: The alternative CBA methodology stipulates that payoffs to all

stakeholders are more than status quo. This means that the alternative does not choose a regulatory level *l* that gives rise to payoffs which is less than status quo for any stakeholders.

Proposition A-2: Since the alternative CBA methodology maximizes the product of all payoffs to stakeholders, even if there is the stakeholder whose payoff is much greater than others, the other stakeholders are still able to affect decisions regarding to the level of regulation. (Mathematical proof is in APPENDIX.)

Proposition A-3: Since the alternative CBA methodology maximizes the product of all payoffs to stakeholders in its decision rules, the product of payoffs can be comparable even if payoffs are described in any kind of unit, such as case/year, $ or Yen.

For example, if the first stakeholder's payoffs $f_1(l)$ is described in case/year, and the second payoffs $f_2(l)$ employs monetary unit, i.e. dollar, the unit of the product of the both payoffs $f_1(l) * f_2(l)$ is expressed in "case/year*dollar". Therefore, the product of payoffs is comparable to find out maximum product of payoffs $\max_{l \in L} \{f_1(l) * f_2(l)\}$. Hence, any unit conversion is not necessary to find out *l* such that maximize the product of all payoffs.

6. AN APPRAISAL OF THE TRADITIONAL AND ALTERNATIVE CBA METHODOLOGIES BY MEANS OF MATHEMATHECAL MODELING

The section 6 intends to demonstrate detailed analysis to rigorously critique the traditional and alternative methodologies. Thus I employ a series of mathematical models which entails two-stakeholder payoff functions to overcome this shortfall.

The analysis to follow has three aims:

1. To test the feasibility of the alternative methodology's decision-making in real situation by formulating the comprehensive payoff functions such that include the cost to adjust

the change of regulation, and consumers' reactions such as boycott related with capability of substitution.

2. To find the limitation of the alternative in real-world situations.
3. To demonstrate the efficacy of a methodology of "relocation of payoffs by the government (subsidy)" to expand the applicability of the alternative, which implemented under the situation where the proposed alternative cannot work.

6.1. A Hypothetical Scenario of the Payoff Function Model

The following scenario illustrates how the analyst might construct a series of payoff function models. A company manufactures product X intended for daily use by consumers. Product X, however, contains a known carcinogenic substance. Consumers thereby are exposed to this carcinogen when they use the product. The regulatory authority estimates the level of cancer risk upon the ascertainment of the exposure concentration, deliberates, and finally drafts and enacts a new regulation in an effort to eliminate or minimize the risk to human health.

6.2 The Payoff Function of the Model that Employs U.S. EPA Methodology.

To construct a payoff function model, I employ a risk assessment methodology for modeling payoffs as the results from the implementation of new regulations. Risk assessment attempts to ascertain the required reduction of exposure to potentially hazardous substances in the environment in order to identify a tolerable level of risk (US NAS/NRC, 1983).

With carcinogens operating as the model's target of regulation, the risk of cancer is assessed in terms of the lifetime risk, calculated by multiplying the cancer-slope factor by the duration of the lifetime exposure (U.S. Environmental Protection Agency, 1992). The risk of cancer for an exposed individual is described as follows:

$$r = \beta d \qquad \textbf{(6-1)}$$

where r is the risk level (usually lifetime risk), d is the lifetime average daily dose (LADD

= Total Dose/(Body Weight * Lifetime)) expressed as mg/kg/day, and β is the potency or cancer-slope factor i.e., the slope of the dose-response curve expressed in units of cases per mg/kg/day). In more precise terms, the risk factor r is the elevated probability of developing cancer specifically associated with exposure to chemical substances cited in the risk assessment (not the general or background probability of developing cancer).

In the context of cancer risk, l is defined as the level (i.e., strength) of the regulatory option measuring how much the exposure concentration e decreases with the implementation of that option. If l = 0, then the rate of reduction of exposure becomes 100%. If l = 1, then the rate becomes 0%.

Stated in more formal terms, l is the level of regulatory options that is required to reach a certain target level of risk r_l. If d becomes d_l upon the implementation of a new regulation, then each variable can now be expressed as follows:

$$r_l = \beta d_l = \beta dl = rl \qquad \textbf{(6-2)}$$

6.2.1. The Payoff Function Model for the Consumers

The payoff function for a certain population of consumers $f_p(l)$ models the payoffs to these consumers when a reduction in the exposure concentration of the carcinogenic substance in product X is obtained by the imposition of a new regulation. These payoffs are defined as the multiplicative product of the reduced risk level and the total welfare of the individual consumer. In accordance with the USEPA definition (US Environmental Protection Agency, 1999), the payoffs can be expressed as follows:

$$f_p(l) = W_p(r - r_l) \qquad \textbf{(6-3)}$$

where W_p is the total welfare of a consumer having an average life expectancy, r is the risk to health, r_l is the risk when level of regulatory option l are in force. Substituting Formula (6-2) to Formula (6-3),

$$f_p(l) = W_p(1 - l)\beta d \qquad \textbf{(6-4)}$$

6.2.2. The Payoff Function Model for the Producer

The payoff function for a certain enterprise (i.e., the producer $f_e(l)$) models the net benefit resulting from product X sales. The payoffs are calculated by subtracting the profit made in the absence of regulation from the cost after a new regulation is imposed (C_e). In general, the C_e escalates dramatically as the level of release approaches 0, and can be expressed as follows:

$$C_e = c(1/l - 1) \quad \textbf{(6-5)}$$

where c is a coefficient that reflects the socioeconomic significance of these costs.

In accordance with the foregoing result, the payoff function for the enterprise (i.e. the producer) can be expressed as follows:

$$f_e(l) = P_e - C_e = P_e - c(1/l - 1) \quad \textbf{(6-6)}$$

where P_e is the profit made in the absence of regulation. Since l can only take on values from 0 to 1, $f_e(l)$ is always greater than 0 for the case $P_e > c$.

6.2.3 The Boycott Coefficient Describing the Consumer's Action

A payoff function can be tailored to accurately reflect the actions of consumers. In general, if consumers obtain information regarding the health risks of certain products, it is assumed that they will boycott such products. It is further assumed that the larger the residual risk r_l is, the more widespread the consumer boycott will be. Therefore, the proportion of consumers who join a boycott B can be expressed as function in the following manner:

$$B = f(r_l) \quad \textbf{(6-7)}$$

where $0 \le B \le 1$. If $B = 0$, then all consumers will join the boycott. If $B = 1$, then no consumers will boycott such products.

According to many studies of risk perception in the public domain, the overwhelming

majority of consumers tend to be "risk averse" (U.S. Environmental Protection Agency, 1997), (Cropper and Oates, 1992). The results of studies focusing on consumers' "willingness to pay" or WTP (the amount of money consumers are willing to spend to avoid risk) are typically summarized in Figure 2. The figure reveals that the general public is inclined to be more sensitive to lower levels of risk than to higher levels.

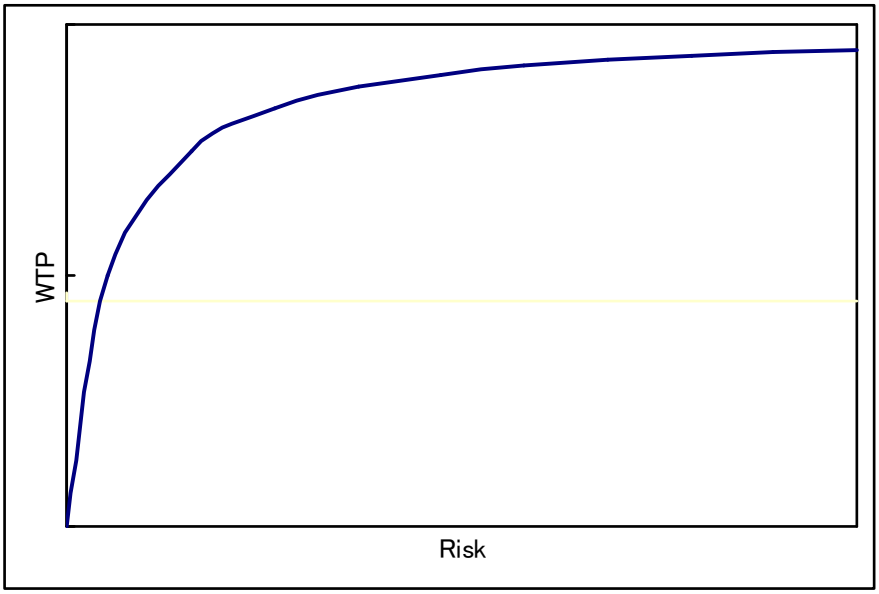

Figure 2: The Relationship between Willingness to Pay (WTP)and Risk Level

Motivation to boycott, risk aversion, and "willing to pay" are closely interrelated. Therefore, *B'* is defined as the proportion of consumers who are willing to boycott any product to pose a health hazard, and can be expressed as follows:

$$B' = f'(r_l) = 1 - l^2 \qquad \textbf{(6-8)}$$

Since *B'* expresses only the "willingness" to boycott, "product substitution" should be considered. For example, tap water, a product, is a daily necessity for all consumers and is relatively accessible, abundant, and inexpensive. Conversely, bottled water is much less abundant and accessible, and certainly much more expensive. Therefore, if the tap water supply is found to pose serious health risks, then many consumers in all probability would not participate in a tap-water boycott because of the cost. Therefore, *b* is defined as a

coefficient expressing the capability of product substitution. This coefficient takes on values $0 \leq b \leq 1$. If b = 0, then substitution is not possible. If $b = 1$, then "perfect" capability of substitution obtains. Incorporating b into Formula below gives rise to the following expression:

$$B = f(r_l) = 1 - bl^2 \qquad \textbf{(6-9)}$$

Since the residual risk r_l decreases linearly with respect to the proportion of boycotting consumers B, substituting Formula (6-9) into Formula (6-917) generates the public's payoff function expressed as follows:

$$f_p(l) = W_p(r - r_l B) = W_p(1 - lB)\beta d = W_p(1 - l(1 - bl^2))\beta d \qquad \textbf{(6-10)}$$

Since the initial profit P_e decreases linearly with respect to the proportion of boycotting consumers B, substituting Formula (6-10) into Formula (6-4) generates the enterprises' (producers') payoff function expressed as follows:

$$f_e(l) = P_e B - C_e = P_e B - c(1/l - 1) = P_e(1 - bl^2) - c(1/l - 1) \qquad \textbf{(6-11)}$$

6.2.4. Modeling the Government-Oriented Redistribution of Payoffs (the Subsidy)

The stakeholders are estimated to act so as to maximize their respective payoffs. However, the government may mandate the redistribution of payoffs among the stakeholders. In particular, the government may reduce the payoff to one stakeholder and simultaneously increase the payoff to another stakeholder. This pattern of shifting costs and benefits is commonplace, for example, among industrial nations. With respect to Japan's Power Source Funding Special Account, the government imposes a tax on the electrical power companies and uses this revenue to subsidize local private entities as well as public facilities located near the power plants. The U.S.'s Super Fund obtains tax revenue from enterprises producing industrial wastes and allocates it to environmental protection programs and toxic-waste cleanup projects.

In the payoff redistribution model, S represents the payoffs transferred from some stakeholders to other stakeholders. Normally, the magnitude of S varies as a function of the

production, such as the amount of generated electric power. In this paper, however, the value assigned to *S* is in direct proportion to the regulatory level *l*. This policy is reasonable because the weaker the regulatory level is, the more profit the enterprises will earn and the greater the health risk will be to the public. Therefore, S can be expressed as follows:

$$S = sl \tag{6-12}$$

where *s* is a coefficient that reflects the significance of the subsidy (s > 0).

By subtracting *S* from the profits of the enterprises and substituting Formula (6-12) into Formula (11), the payoff function of the enterprises can be expressed as follows:

$$f_e(l) = P_e B - C_e - S = P_e(1 - bl^2) - c(1/l - 1) - sl \tag{6-13}$$

Furthermore, by adding *S* to the payoffs to the consumers and substituting Formula (6-12) into Formula (6-10), the payoff function of the consumers can be expressed as follows:

$$f_p(l) = W_p(r - r_l B) - S = W_p(1 - l(1 - bl^2)\beta d + sl \tag{6-14}$$

6.3. Conditions for the proposed alternative to function

Reflection on what we seen in former sections will make clear the conditions for the ADM to work properly. These conditions can be classified as following cases.

6.3.1. The Effects of a Boycott

Can the proposed alternative be effectively utilized in conjunction with the payoff functions introduced in previous sections, recall that Formula (6-3) is the consumers' payoff function, and Formula (6-5) is the enterprises' payoff function. These functions, regardless of what values are assigned to their respective coefficients, the consumers' payoffs monotonously decrease as the regulatory level increases, because the slope is always negative. Conversely, the enterprises' payoffs monotonously increase as the regulatory level increases, because the slope is always positive. Therefore, the consumers' payoffs decrease as the enterprises' payoffs increase. In this situation, all regulatory levels satisfy the PO

criterion to both consumers' and enterprises' payoffs. In this case, the proposed alternative will always have to choose the level of the status quo in accordance with Decision Rule 1. Therefore, it appears that the proposed alternative's decision-making process cannot identify possible regulatory levels in given situation.

6.3.2. The Case Which Considers the Effects of a Boycott

If the effects of a boycott by consumers are taken into account, the foregoing conclusion changes dramatically. Recall that Formula (6-9) expresses the consumers' payoffs, and Formula (6-10) expresses the enterprises' payoffs. If a certain boycott level is in play, then the value of coefficient b has a certain value greater than 0. In case that the slope of $f_e(l)$ change from negative to positive within the range $0 \le l \le 1$, a relationship between $f_p(l)$ and $f_e(l)$ looks like as figure 3 such that the weakest regulatory level among those satisfying the PO obeys Decision Rules 1 and 2. This outcome is equivalent to the fact that there does not exist an l^* satisfying the criteria of the following formulas:

$$\frac{\partial^2 f_e(l^*)}{\partial^2 l} < 0 \qquad \textbf{(6-15)}$$

$$\frac{\partial f_e(l^*)}{\partial l} = 0 \quad and \; 0 \le l^* \le 1 \qquad \textbf{(6-16)}$$

An l^* satisfying the criteria of Formulas (6-14) and (6-15), and the values of coefficient b that can satisfy the criteria of Formula (6-16) can be expressed as Formula (6-17).

$$l^* = (c / P_e b)^{1/3} \qquad \textbf{(6-17)}$$

$$b \ge c / 2P_e \qquad \textbf{(6-18)}$$

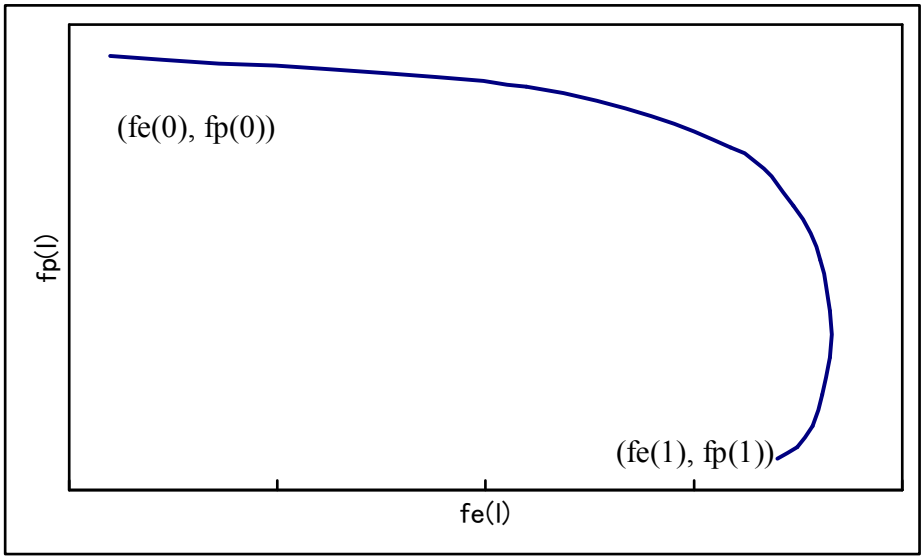

Figure 3: The Relationship between f_e(l) and f_p(l) , in case of $b \geq c/2P_e$.

The preceding analysis identifies the values of coefficient *b* that can satisfy the criteria of Formula (6-17) as well as facilitate the application of the proposed alternative to regulatory issues. Moreover, Formula (6-16) shows that decision-making of the proposed alternative does not depend on the value of W_p. Therefore, the determination of regulatory level options by the proposed alternative methodology is independent of population density factors and the incomes of consumers, and treats each consumer equitably.

Formula (6-13) also shows that the proposed alternative selects more strict regulatory levels as coefficient *b* increases. Thus, the capability of product substitution has a strong effect on decision-making. Additionally, the smaller the initial profit of an enterprise is, the weaker the regulatory level tends to be decided. These results are rational decisions intended to foster the survival of enterprises.

6.4. The traditional CBA's Criteria for Obtaining the Support of the Stakeholders

In addition to the analysis of proposed alternative in former sections, the proposed alternative does not function optimally unless consumers have at their disposal the ability to

substitute products. However, certain products (e.g., clean air, tap water, electricity) are not substitutable in an absolute sense or in terms of economic efficiency. A further example is the situation where a consumer boycott cannot reduce the exposure to harmful substances released by enterprises that are located in close proximity to the consumers. In this instance, it is imperative that the governmental intervention to re-distribute payoffs among enterprises and consumers or general public around factories, using its subsidy system. The subsidy system can assure the traditional CBA to pick up the regulatory level that maximizes payoffs of stakeholders as mentioned in Section 6.2.4.

6.4.1. Detailed Conditions of the Subsidy for the traditional CBA to Work Properly

For the formal definition of the subsidy system, the capability of product substitution is assumed to be nonexistent ($b = 0$), then the government should raise the amount of subsidy per regulatory level s until the slope of $f_p(l)$ becomes 0. In this instance, s satisfies the following conditions:

$$\frac{\partial f_p(l)}{\partial l} = -W_p \beta d + s = 0\,, \quad then,\ s = W_p \beta d \qquad \textbf{(6-19)}$$

where the weakest level among the regulatory levels meeting the PO criterion satisfy the criteria of Formulas (6-14) and (6-15) as well. Therefore, substituting Formula (26) and (6-18) to (6-14), then the l^* can now be expressed as follow:

$$l^* = (c/s)^{1/2} = (c/W_p \beta d)^{1/2} \qquad \textbf{(6-20)}$$

If l^* satisfies the criteria of Formula (6-19), Formula (6-14) is satisfied. Therefore, $f_e(l)$ have a peak when $l = l^*$. Furthermore, l^* can be the regulatory level that maximizes payoffs to both the enterprises and the consumers. Application of the traditional CBA herewith strengthens this result.

6.4.2 The Deficiencies of the Subsidy Model

The subsidy model, however, has some major shortcomings. For example, Formula

(6-18) indicates that the required amount of subsidy per regulatory level s is proportional to the amount of W_p. Formula (6-19) shows that as W_p becomes smaller, a smaller s is required, and a smaller s necessitates a larger (i.e., more robust) l^*. Therefore, smaller populations of consumers are subjected to higher levels of risk than larger populations of consumers. On the other hand, Formula (6-12) and (6-18) reveals that if W_p is sufficiently large, $f_e(l)$ becomes less than 0, and economic activity stagnates. This scenario implies that facilities releasing harmful substances can only be located in regions of low population density.

7. DISCUSSION

This chapter proposed the alternative approach of CBA for enabling socially justifiable distribution of social goods. To be consistent with the second principle of justice, or "everybody's advantage", the concept of the status quo is introduced to justify the "advantage". The propositions above and the feasibility study of the alternative in realistic situation reveal that the alternative can enable the CBA to distribute social goods in accordance with Rawls's principles of justice.

This chapter tried to express the social system as the set of mathematical elements. This enables us to describe the social system change by imposing the new regulation in mathematical manner. The mathematical expression can evaluate "cost" and "benefit" objectively, although the cost and benefit are likely to be subjectively defined. Moreover, this mathematical evaluation can provide the objective basis for the further research of CBA.

The most important mathematical modeling in this paper is the stakeholders' payoffs as the function of the regulatory option level l. This modeling enables the mathematical evaluation of deficiencies and advantages of the traditional CBA and the proposed alternative. Moreover, employing the practical modeling as the payoff function allows calculating and comparing the feasible set of the actual regulatory option level.

The applied model in Section 6 aims to test the feasibility of the alternative in real situation by formulating the comprehensive payoff functions such that include the cost to adjust the change of regulation, and consumers' reactions such as boycott related with capability of substitution. The results reveal that the proposed alternative can work in those realistic situations. Moreover, the model succeeded to the alternative's applicability can be broadened if the methodology of "relocation of payoffs by the government (subsidy)" implemented.

The following sections show the result of critical evaluation of the traditional CBA and the proposed alternative on each evaluation criteria.

7.1 A Critical Evaluation of the traditional CBA and the proposed alternative

From the foregoing analyses (Sections 4, 5, and 6), a critical appraisal of the traditional CBA methodology and the proposed alternative methodology follows. Comparisons are made by applying the criteria which were introduced in Section 3.

a. Decision-making by means of CBA should be advantageous to all stakeholders.

The traditional methodology upholds the desirability of a social system's regulatory environment with respect to the stakeholders collectively. However, it does not guarantee to uphold the desirability of the stakeholders individually. In certain cases, the chosen level of regulation may lead to payoffs less than those in status quo for some stakeholders. (see Proposition T-1 in Section 4.2) Conversely, the alternative approach employs decision rules which generate only payoffs more than those in status quo (see Proposition A-1 in Section 5.2). Thus the designated level of regulation leads to payoffs that are always favorable and sustains the system's rationality with respect to the stakeholders individually and collectively.

Furthermore, the alternative approach enables the regulatory authority to reliably choose the most efficient list of payoffs, because the alternative requires that all lists of

payoffs satisfy the Pareto optimal.

b. Within the context of public decision-making, the stakeholders under all circumstances should be able to influence the outcomes of CBA.

The traditional methodology employs decision rules which enable the determination of the net benefit (the summation of payoffs) to all stakeholders. However, stakeholders who have relatively small payoffs are likely to be ignored in the regulatory level selection process (see Proposition T-2 in Section 4.2). Conversely, the alternative approach employs decision rules which facilitate the calculation of the product of all payoffs to stakeholders (see Section 5.2). Therefore, stakeholders who have relatively small payoffs are potentially able to influence regulatory decision-making.

c. CBA should be unit-free for valuing cost and benefit.

The use of traditional CBA requires that the health risks that a given regulation might engender are evaluated exclusively in monetary terms, as Proposition T-4 stated. In many cases, monetary valuation is based on an individual's preference or "willingness to pay" (WTP) to reduce health risks. The EPA method stated in Section 6.2. is also widely used. Although various methodologies for evaluating benefits have been advanced, most are still in need of further testing and refinement. (Cropper. 2000)

In author's view, the most important deficiency of these monetary valuations is the dependence of social economic level. Whichever the WTP or EPA methods are employed, monetary valuation of a single person's life in developed countries is higher than in developing countries, because of income level gap. Moreover, the cost to adjust the regulation is normally higher in developing countries than in developed countries, because of transportation and import cost. (Trqwen, Maraste, and Persson 2002) Therefore, the traditional CBA is likely to result in higher risk level to people in developing countries than in developed.

7.2 Further Research Needed to Improve the Proposed Alternative

The proposed alternative CBA approach, however, has certain shortcomings that need to be addressed. Its principal weakness is that it cannot be applied if there is no list of payoffs where all such payoffs are positive. Furthermore, since the regulatory level which maximizes the multiplicative product of all stakeholders is not always guaranteed to fall within the ALARP range of risk, this approach as presently configured is not able to provide a decision rule for consistently selecting the appropriate level. In view of these drawbacks, further study into and refinement of this methodology are necessary.

8. CONCLUSION

This chapter proposed a new methodology for decision-making on economic and social issues in which everyone's net-beneficial advantage compared to the status quo and an optimization of net-payoffs are guaranteed. The traditional CBA methodology has long been considered the dominant method, however, the foregoing analyses and modeling clearly demonstrate that the proposed alternative CBA methodology possesses a number of unique advantages, which include the avoidance of the decision by which disadvantageous net-payoffs yields to some stakeholders, the direct impact that relatively small stakeholders can have on the decision-making process, and no need to convert the unit of payoffs, which is consistent with the principle of equal right and the principle of everybody's advantage of John Rawls. The traditional approach does not share these attributes. Therefore, the alternative approach, which incorporate the concept of time sequence into the framework and whose conceptual basis is Pareto optimization of every payoff occasioned compared to those in the status quo, offers a potentially more effective and versatile decision-making tool. Accordingly, further research into its utility for direct application in actual decision-making is warranted.

REFERENCES

Beck B. D., Dudel R., and Calabrese E. J. 1994. The Use of Toxicology in the Regulatory Process. In: Hynes A.W. (eds), Methods of Toxicology, 3rd ed. Pp. 19-58. Raven Press, Ltd., New York, USA.

Cropper M. L. 2000. Has Economic Research Answered the Needs of Environmental Policy? Journal of Environmental Economics and Management 39, pp.328-350

James R., Warren D., Helmes N., Roberts S. 2000. Risk Assessment. In: Williams P, James R, Roberts S. (eds), Principle of Toxicology: Environmental and Industrial Applications, Second Ed. pp.437-477, John Wiley and Sons, NY, USA

Marszal E. M. 2001. Tolerable Risk Guidelines. ISA Transactions 40:391-399.

Oka T. 2003. Cost-effective Analyses of Chemical Risk Control Policies in Japan. Chemosherer 53:413-499.

Rawls, J. 1971. A Theory of Justice, Clarendon Press, Oxford

Roy A. D. 1952. Safety first and the holding of assets. Econometrica 32, 431-449.

Rutherford D. 1995. The Routledge Dictionary of Economics, p.343, Routledge Press, London and New York, UK and USA.

Travis C., Richter S., Crouch E., Wilson R., and Klema E. 1987. Cancer risk management. A review of 132 federal regulatory decisions", Environ.Sci.Technol.21(5)

Trawen A., Maraste P., and Persson U. 2002. International comparison of costs of a fatal causality of road accidents in 1990 and 1999. Accident Analysis and Prevention 34:323-332.

U.S. Environmental Protection Agency. 1992. Guideline for Exposure Assessment, EPA/600Z-92/001. Pp. 16-49, 124-135. National Center for Environmental Assessment, Washington DC, USA.

US Environmental Protection Agency. 1999. The Benefit of and Costs of the Clean Air Act: 1990-2010. EPA-410-R-99-001, Washington DC, USA

US NAS/NRC. 1983. Risk Assessment in the Federal Government: Managing the Process.

US Presidential Commission on Risk and Risk Management 1997. P/C Commission,

Washington DC, U.S.A.

Wilson R. and Crouch E. A. C. 2001. Risk Benefit Analysis, 2nd ed. Pp. 101-133. Harvard University Press, Cambridge, USA.

APPENDIX

PROOF OF PROPOSITION T-2

$f_1(l)$ can be described as a linear function of $f_2(l)$ since both are a function of the regulatory level *l*. Therefore,

$$f_1(l) = G\{f_2(l)\} \qquad \textbf{(A1)}$$

We assume that $f_1(l)$ is very much greater than $f_2(l)$ for any $l_A \in L$, hence the following relationship:

$$f_1(l_A) = kf_2(l_A) \text{ or } \quad f_2(l_A) = \frac{1}{k} f_1(l_A) \qquad \textbf{(A21)}$$

Where, $|k| >> 0$, then,

$$f_1(l_A) + f_2(l_A) = (1 + \frac{1}{k}) f_1(l_A) \approx f_1(l_A) \qquad \textbf{(A32)}$$

Formula (A3) means that the relatively small stakeholders can be ignored in decisions made by applying the traditional methodology.

PROOF OF PROPOSITION A-2

If there are two stakeholders, the following Formula can stand for a value of K:

$$\max_{l \in L} \prod_{j=1}^{n} f_j(l) = \max_{l \in L} F'(l) = \max_{l \in L} f_1(l) f_2(l) = K \qquad \textbf{(A4)}$$

In this situation, I assume that the absolute value of $f_1(l)$ is much larger than $f_2(l)$ for all $l_A \in L$. The following Formula expresses this relationship.

$$f_1(l_A) = kf_2(l_A) \quad \textbf{(A5)}$$

where, $|k| >> 0$.

In the case of $\max_{l \in L} F'(l) = K$ (K > 0), the pair of $f_1(l')$, $f_2(l')$ satisfies the first and second decision rules and is equivalent to the intersection of Formulas (A4) and (A5). Substituting Formula (A4) into Formula (A5) yields the following:

$$\frac{1}{k}(f_1(l_A))^2 = K \quad \textbf{(A6)}$$

It is further assumed that the payoffs to the second stakeholder always equals $k' \approx 0$ for all $l_B \in L$. This expression implies that the second stakeholder is virtually nonexistent. Therefore, in the case of $\max_{l \in L} F'(l) = K$ (K >0), it follows that Formula (A4) is equivalent to

$$F'(l_B) = k' f_1(l_B) = K \quad \textbf{(A7)}$$

By Formulas (A6) and (A7), $f_1(l_A)$ and $f_1(l_B)$ can now be described by the following Formula:

$$\left(\frac{1}{k}\right)(f_1(l_A))^2 = K = k' f_1(l_B) \quad \textbf{(A8)}$$

If k has large enough absolute value, then $\frac{1}{k} \approx k' \approx 0$ and $(f_1(l_A))^2 \approx f_1(l_B)$. Because the dimension of $(f_1(l_A))^2$ and $f_1(l_B)$ is different, $l_A \neq l_B$ is generally obtained.

According to the foregoing analysis, even if the first stakeholder's payoffs are much greater than the second stakeholder's, under the alternative methodology the latter is still able to affect decisions regarding the level of regulation.

Chapter 4

A Feasibility Study of the Rawls's Principles of Social Justice in Strategic Situations

Summary

This chapter examines the feasibility of the adoption of the principles of social justice, in particular, the "generalized" differential principle, which require everyone's advantage, to the decision-making process in a strategic situation. The concept of equilibrium, which is a dominated concept of game theory, does not utilize non-utility information and cannot guarantee to benefit all players compared with the status quo. For that purpose, as the application, this chapter introduces the game model in extensive form which has the following characteristics: (a) sub-game perfect equilibrium (SPE) should be employed as a solution, (b) the status quo should be treated as a disagreement point, and (c) a game tree should be divided into several stages (multi-stage approach). Applying "implementation theory" or "mechanism design", this chapter focus on defining the condition of the subset of "equilibrium outcome" which satisfy SPE as "solution concept" and unanimous decision as "mechanism" to meet "everybody's advantage" principle as "choice correspondence". Thereby, we examine the interrelations of the strategy of players. As the application, the chapter analyzes a collective agreement of industrial relations, in particular, the effects of the "white collar" exemption system on working hours and productivity of agents, in application with the traditional principal-agent theory. The analysis of this chapter confirms the coherency of Rawls's "generalized" differential principle with a strategic situation. The main result of the application is that the adoption of the "white collar" exemption system may induce shorter working hours compared with those under the current work-hour payment system. The analysis also reveals that the overtime premium has positive effect on working hour elongation.

1. INTRODUCTION

In this chapter, we investigate whether the decision-making process in non-cooperative situation can preserve the social justice such as Rawls's. As widely acknowledged, the game theory does not utilize non-utility information, and the concept of equilibrium, a core concept of the game theory, is defined based on the payoffs of the players of a game. Through this process, thereby, there is no guarantee that the alternative which is an equilibrium point in strategic situation can satisfy the concept of social justice which is advocated such by John Rawls as well.

1.1. Proposed decision framework

We focus on the Rawls's principles of justice, in particular the "generalized" differential principle such that "social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone's advantage, and (b) attached to positions and offices open to all" (Rawls 1971) . We employ this principle as a benchmark to evaluate the realization of social justice in the decision process on the basis of the game theory.

In this regard, we need the decision frameworks in which we can identify if the alternative becomes Pareto superior compared with respect to the status quo. We, thereby, propose the following decision framework by which we can assess the conformity to the Rawls's principles of everybody's advantage, and examine its feasibility to the strategic situation;

a) Incorporating concept of "time sequence" into the decision making framework. This could identify the situation before the decision (status quo). The status quo should not be restricted the "initial" situation of the sequence, could be defined as the situation prior to the "last" decision. (multi-stage approach)
b) In conflicting situation (such as situation which cannot reach unanimity), the status quo should be treated as disagreement point of the conflicting situation.

c) Identifying the list of payoffs which all stakeholders can adopt (sub-game perfect equilibrium SPE) in comparison with the payoffs in the status quo. Any single person who is deteriorated his payoffs (social goods) cannot be legitimate in the decision.

Having a normal form game been difficult to coherent with the requirement of the framework, we employ an extensive form game in which if one of the players of the game denies an alternative, then the play goes back to the status quo. In this game tree, the list of payoffs of the alternatives which satisfy sub-game perfect equilibrium cannot be Pareto inferior with respect to the status quo from the definition.

In the application of this extensive form game, we can construct the game in which we can simulate a multi-stage decision process by recurrent use of the basic game tree. Also, the game can be extended to n-person game.

1.2. "Mechanism design" and proposed framework

The methodology employed this chapter can be explained in the framework of "mechanism design" or "implementation theory" (Hurwicz 1960), (Maskin, 1999), in particular "multi-stage mechanism" introduced by J. Moor and R. Repullo (Moor and Repullo 1988), (Moor 1992). In this framework, A denotes a "set of outcomes", H denotes the direct product of H_i, the set of payoff function of player i. If H is given as the "set of profile" of payoff functions of players, the correspondence $F : H \to\to A$ is called "choice correspondence". In this chapter, the generalized differential principle, or "everybody's advantage" is employed as the correspondence. S denotes the direct product of the set of strategy of all players S_i. If we employ sub-game perfect equilibrium as "Solution concept", $SPE_g(h)$ denotes the set of strategy profiles which are sub-game perfect equilibrium. Then, the correspondence $SPE_g : H \to\to S$ is called sub-game equilibrium correspondence. Lastly, the function $g : S \to A$ is called "mechanism"

which is defined as a game tree in this model. Specifically, the mechanism in this model employs "unanimous decision making" and treats the status quo as "disagreement point". The feasibility study implemented in this chapter can be understood as "implementation" problem whether the strategies of "equilibrium outcomes" $SPE_g(h)$ (that is, solutions of the game) can satisfy Rawls's generalized differential principle. Formally this problem is described as; whether (or not) $g(SPE_g(h)) \subseteq F(h)$. Following which, we try to identify the set of profile H which can satisfy the above condition. In the proposed model, it means finding out the set of strategies which can be "agreeable" by all players. This model can graphically be described as Figure 1.

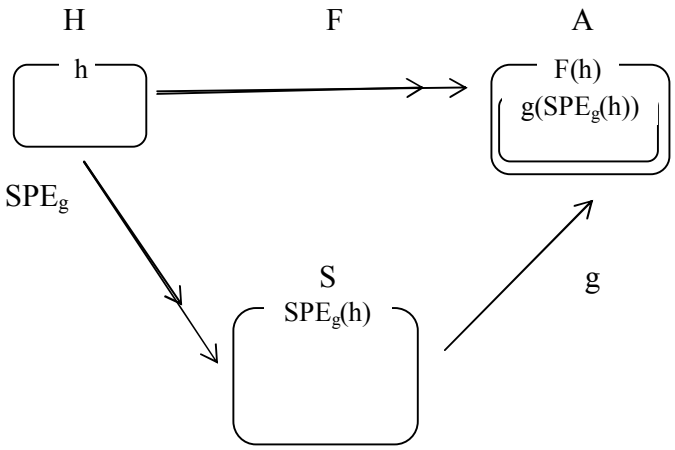

Figure 1

In the above analysis, the mechanism is not designed so that all equilibrium outcomes can satisfy the choice correspondence, and analysis which is different from the standard implementation theory. In the scheme in which the status quo is treated as a disagreement point, the mechanism is only required to give opportunities that at least one equilibrium outcome can satisfy the choice correspondence. If a player cannot reach a decision, then they just keep status quo and nothing happens. In this case, we should focus on the conditions or restrictions to enable the subset of equilibrium outcomes meet a choice correspondence.

2. APPLICATION: Analysis of Effects of "White Collar" Exemption System on Working Hours

A "white-collar" exemption system is the system that exempts workers who are regarded as "white-collar" from overtime payments protection. This system has been employed in the U.S. since late 1900s. Recently, adoption of this system becomes a controversial issue in Japan. Employers groups which intend to introduce this system insist that "in line with a performance-based payment system, creating a "white collar" exemption for certain employees would facilitate a shift towards compensating employees for the actual work that they do, not for the number of hours they spend at the work place" (The American Chamber of Commerce in Japan, 2007). However, workers groups who oppose the proposal insist that the system increases incidences of "karoshi", a sudden death caused by excessive overtime working, because the system gives workers an incentive to work longer in order to get better wages (Fujino and Matsuda, 2007). Also, it is said that a "white collar" exemption needs specific labour conditions to work such as well-functioning external labour markets and dissemination of performance-based payment scheme, which are rarely seen in Japan (Ishige, 2005). On this point, employers groups argue that the system would motivate "white-collar" workers to work more efficiently and productively. Unfortunately, there is no formal analysis of the effect of the "white collar" exemption system on working hours and workers' productivity. This present chapter aims to fulfil this gap.

In order to analyse an incentive system which can motivate workers to conduct more efficient work, the principal-agent model has been widely used. The principal-agent theory was first presented by Alchian and Demsetz in 1972. The economics of the principal-agent relationship were further developed, among others, by Sharvell,

Holmstrom, Grossman and Hart in early 1980s. In the traditional principal-agent model, however, only two individual can be analyzed. Also, the model assumes that the principal can propose only one specific type of payment schemes and cannot choose the best type of payment schemes with the consideration of agent's decision, although an agent can choose its effort to maximize its payoffs.

The present chapter employs the game in extensive form as a model in order to overcome above deficiencies. The concept of subgame perfect equilibrium (SPE) (Selten, 1975) enables us to analyze conditions in which the strategy to adopt the "white collar" exemption becomes the best response for both a principal and agents. Moreover, the game in extensive form can analyze bargaining problems among more than two players with due consideration of player's rationality to maximize their payoffs. Applications of the game theory to principal-agent bargaining are viewed such as an analysis of Stackelberg equilibrium in oligopoly situation (Basu, 1995). Also, several recent studies employed the extensive game model and utilized the concept of SPE for the analysis of principal-agent bargaining including wage setting (Ruiz-Verdu, 2007, Tasnadi, 2005), union form bargaining (Appelbaum, 2007).

The present chapter aims to find out effects of the "white collar" exemption system on agent's working hours under the strategy such that a principal selects "propose" and agents select "accept", which satisfies SPE of the game. In doing so, working hours of agents under the exemption system can be obtained. Also, by the use of multi-agent model, this chapter tries to expand the analysis of the effect of the "white collar" exemption to unlimited number of agents.

Our main research questions can be described as follows:

1. Does adoption of the "white collar" exemption system contribute to extend working hours?
2. Does a rise in overtime premium contribute to extend working hours?

This present chapter is organized as follows. Basic concepts of the employed model are presented in Section 3. Section 4 covers a single-person model under the assumption that productivities of workers are constant. In Section 5, all results of former sections are evaluated.

3. MODEL

Mathematical definitions of the game in extensive form to analyze effects of the "white collar" exemption system are as follows: the game in extensive form is a five-tuple model (*K, P, U, h, T*), where *K* is the game tree, $P = \{p, a\}$ is the set of all players, where *p* is the principal, and *a* is an agent. *U* is the set of all information sets, and $h = (h_{p,} h_a)$ is the list of payoff functions of players, where h_p is the payoff function of the principal, and h_a is the payoff function of the agent. $T = \{-m, \cdots, t, -2, -1, 0\}$ is the set of retrospective timings which specify stages of decision. The status quo of each stage is defined as unique precondition of decision in each stage. Timing 0 identifies the stage in which the latest decision takes place, and specifies the status quo as the unique precondition in the latest stage. Then timing -1 identifies the decision stage which is one stage prior to the latest one (if any) and specifies the status quo in that stage. The concept of *T* is retrospective, and does not foresee a prospective situation. $h_p(s_t)$ and $h_a(s_t)$ are payoffs of the principal and the agent respectively under the strategy s_t. Where, s_t is a precondition of strategy under a specific timing of *t*.

3.1. Assumptions to apply the game theory

In this present paper, we additionally assume the followings:

Assumption 1: We do not consider any exogenous risk to the principal's gross wealth for simplicity. In other word, risk is zero. Therefore, the principal and the agent can know their wealth based on the agent's efforts precisely. Hence this chapter does not consider an "expected value" of their wellness and does not have to employ "utility functions" to evaluate the agent's "risk averse" behaviour.

Assumption 2: We assume rationality of the principal and agents. The principal can select to offer a payment scheme to agents or not. If the principal does not offer a payment, no employment happens and the principal does not gain any outcome from agents. Agents accept a payment offered only if a payment outweighs reservation utilities of each agent. In case of rejection of an offer, agents have options to gain reservation utilities from other source. Agents also select their effort to gain their maximum payoffs, and the principal selects a payment scheme which can maximize the principal's payoff by comparing with the payoff under the current scheme. This assumption enables us to utilize the concept of game theory including the concept of equilibrium.

Assumption 3: We assume that status quo functions as a disagreement point. This assumption combined with "everybody's advantage" logically has the decision to be unanimous decision, because even a single player's disadvantage has a veto power to a decision. Thus if and only if the principal and all agents selects the new "white collar" exemption payment scheme, the new payment scheme would be accepted. By this assumption, we can adopt the concept of sub-game perfect equilibrium (SPE) as a solution. Otherwise, current work-hour payment scheme (the status quo) would be kept. Unanimous decision making is efficient in the sense that it achieves Pareto optimal outcomes under some realistic conditions. (Yasui and Inohara, 2007)

3.2. The Payoffs of Players

The payoffs of the agent in each strategy are described formally as the following formulae:

$$h_a(s_{-1}) = R,$$

$$h_a(s_0) = p_c(o) - c(e), \text{ and}$$

$$h_a(s_n) = p_n(o) - c(e),$$

where R is a reservation utility of the agent, o is an output function of the agent, e is an effort of the agent, and $P(o)$ is a payment scheme, and $P_c(o)$ denotes the current scheme, $P_n(o)$ denotes the new scheme. $c(e)$ is a disutility function of the agent. Where, s_{-1} is the precondition of strategies in the timing *-1*, which identify the decision stage of the basement model in Section 4.1, and s_0 is the precondition of strategy in decision stage at timing *0* for the proposed game tree in Section 4.2, and s_n is the situation decided by the proposed game tree.

In the same manner, a payoff of the principal in each strategy is prescribed as following formulae:

$$h_p(s_{-1}) = 0,$$

$$h_p(s_0) = o - p_c(o), \text{ and}$$

$$h_p(s_n) = o - p_n(o).$$

3.3. Payment Scheme

The payment scheme of the "white collar" exemption system P_n and the scheme of current work-hour system are denoted by:

$$P_n(o) = r + q(o), \text{ and}$$

$$P_c(o) = b + k(e),$$

where r is the fixed salary of agents, $q(o)$ is the performance based salary of the agent, b is the basement salary of agents and $k(e)$ is the overtime payment of the agent.

3.4. Assumptions to Analyze the "White collar" Exemption System

3.4.1. Output function of agents

Let $o = f(e,\theta)$ be the output function of the agent. This type of output functions is used in the traditional Linear-Exponential-Normal-Model (LEN model) (Spremann, 1987). The output of the agent is expressed as:

$$o = f(e,\theta) = e + \theta .$$

Because of Assumption 1, risk $\theta = 0$. This output hypothesizes that agent's productivity is expressed by the productivity constant. In this model, the agent has each productivity constant equal to *1*.

3.4.2. Overtime payment function

Overtime payment function is defined as:

$$k(e) = v(e - h) ,$$

where h is regular working hour (i.e. 40 hours per a week), and v is overtime payment per an hour. Let g denote the overtime premium ($g > 1$). Then, v can be described as follows:

$$v = gb / h .$$

3.4.3. Performance based salary function

Performance based salary function is:

$$q(o) = do ,$$

where d is the distribution constant ($0 < d < 1$).

3.4.4. Agent's disutility function

This paper employs the LEN model's disutility function as follows:

$$c(e) = e^{2}.$$

4. ANALYSIS OF SINGLE-AGENT MODEL UNDER CONSTANT PRODUCTIVITY

This section analyzes the model under the assumption that there is just one agent and it keeps constant productivity.

4.1. Basement Model for Single Agent Model

In order to simplify the model and its analysis, we employ multi-stage approach. We firstly define the basement model to analyze the basement process until the principal and the agent select the current work-hour payment scheme. Then, we construct the model to analyze the process until the principal and the agent selects the new payment scheme after the basement process. In this model, the status quo is utilized as the disagreement point of each decision node. Also, the status quo itself is defined multi-staged, such that the status quo in the first stage is unemployment, and the second stage's is the employment in hourly paid scheme.

4.1.1. Proposed Game Tree of Basement Model for Single Agent Model

The process that the principal and the agent select the current work-hour payment scheme is modelled by the traditional principal-agent theory (Spremann, 1987). Such model can be described as a game in extensive form as shown in Figure 1.

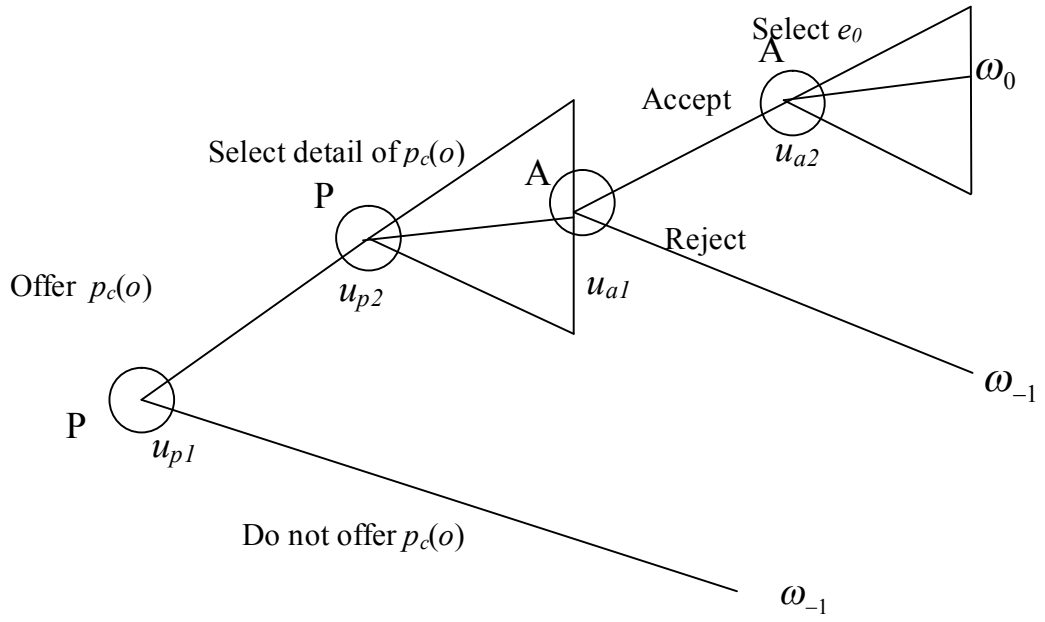

Figure 1 Game tree of the basement model (timing -1) for single-agent model

In this game, the principal and the agent are players of the game. There are two type of terminal nodes $\omega_t \in \{\omega_{-1}, \omega_0\}$. The game ends at ω_0, if the principal offers the current payment scheme $p_c(o)$, the agent accept the offer, and the agent and the principal engage the contract under the current work-hour payment scheme. Otherwise, the game ends at ω_{-1}, the agent is not employed by the principal, and the agent gains reservation utility. Agent's decision depends whether the payment offered outweighs reservation utility.

A list of strategies to reach a specific terminal node $\omega \in \{\omega_{-1}, \omega_0\}$ is denoted by $s_t \in \{s_{-1}, s_0\}$. In each terminal node, payoffs of all players $H(s_t) = \{h_p(s_t), h_a(s_t)\}$ are defined. Also, there are several information sets of players. The set of all information sets in which player *j*'s selects an action is denoted by $U_j = \{u_{jk}\}$, where $j = \{a, p\}$ and *k* is an index for the information sets. The set of all information sets in the game tree *K* is denoted by *U*.

According to the concept of subgame perfect equilibrium (SPE), if there exists the set of strategy s_t which reaches to ω_t, and if s_t satisfies SPE, the selection of payment scheme $p_c(o)$ is rational for both the principal and the agent.

4.1.2. Analysis of conditions for s_t to be SPE

Following the "backward induction" method, let us start from the agent's determination of its effort at u_{a2}. Next, at the move of u_{a1}, the agent decides "accept" or "reject" the proposed new payment scheme. Then, at move of u_{p2}, the principal determines details of proposed payment scheme (value of b) in order to maximize the principal's payoff $h_p(s_0)$. Lastly, at move of u_{p1}, the principal chooses "propose" or "not propose" the new payment scheme.

Through the above analysis, the following formulae are derived as the conditions for s_t to be SPE (See detailed analysis in Appendix A):

$$b^* = \frac{h^2}{g^2}\left(\frac{gt}{2h} - 1 + g\right) \text{ and } e_0 = \frac{t}{4} + \frac{2}{h} - \frac{h}{g},$$

where b^* is the basement salary of agent which maximize principal's payoff $h_p(s_0)$. e_0 is the effort of agent which maximize its payoff under the current payment scheme. According to this formula, working hours expand as overtime premium rises.

4.2 Proposed Game Tree and Set of Strategies for Single Agent Model

The following game model defined in this section is assumed to start from the terminal node ω_0. This means that the agent has been hired under the current payment scheme $p_c(o_0)$ and b^*, and the agent selects the effort e_0 as defined in Section 3.2. This condition is the precondition (or the status quo) of the game defined in this section.

The single-agent traditional principal-agent model can be described as games in extensive form as shown in Figure 2.

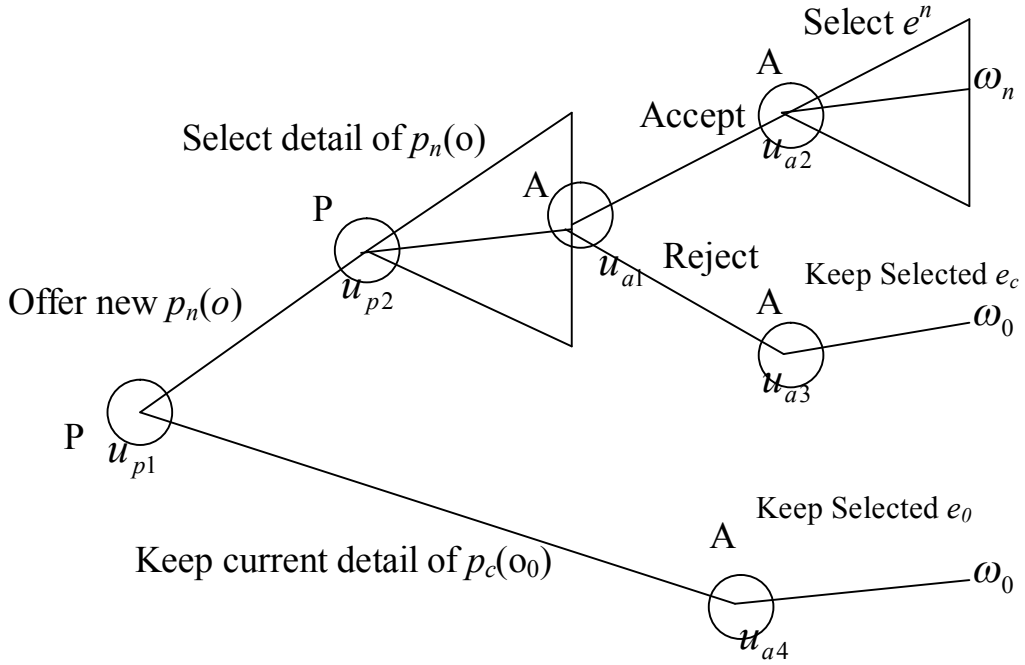

Figure 2. Game tree of single-agent model (timing 0)

In this game, the principal and the agent are players of the game. There are two type of terminal nodes $\omega = \{\omega_n, \omega_0\}$. The game ends at ω_n if both the agent and the principal select a new payment scheme $p_n(o)$ such as the "white collar" exemption system. Otherwise, the game ends at ω_0 and the current work-hour payment scheme $p_c(o)$ is selected. A list of strategies to reach a specific terminal node $\omega = \{\omega_n, \omega_0\}$ is denoted by $s_t \in \{s_n, s_0\}$. In each terminal node, payoffs of all players $H(s_t) = \{h_p(s_t), h_a(s_t)\}$ are defined. Also, there are several information sets of players. The set of all information sets in which player *j*'s selects an action is denoted by $U_j = \{u_{jk}\}$, where $j = \{a, p\}$ and *k* is an

index for the information sets. The set of all information sets in the game tree *K* is denoted by *U*.

4.2.1 Analysis of conditions for s_n to be SPE

Following the "backward induction" methodology, under the assumption that the agent's productivity is constant ($o = f(e_n, \theta) = e_n$), the relationships between working hours under the "white collar" exemption system e_n and those under the current work-hour payment system e_0 can be described as follows (See Appendix B for detailed analysis):

$$e_0 = e_n + h/2(1-1/g),$$

where $g > 1$, and $e_0 > e_n$ no matter what *g* is.

According to the above analysis, if the "white collar" exemption is introduced, working hours under the system may be shorter than those under the work-hour payment system.

Also, fixed salary of new payment scheme *r* has to satisfy the following:

$$r > \frac{h^2}{4g^2}\left(\frac{gt}{2h} - 3(g-1)\right)\left(\frac{gt}{2h} + (g-1)\right) - \frac{t^2 d^2}{4} \quad \text{and}$$

$$r > \frac{t^2}{8} - \frac{h^2}{2g^2}\left(g\left(\frac{t}{2h}+1\right)-1\right)^2.$$

5. CONCLUSIONS

Through the analysis above, we conclude that the Rawls's "generalized" differential principle can be preserved in the model in which the status quo is kept unless all players agreed the option offered unanimously, whose model employs "multi-stage" approach and the concept of sub-game perfect equilibrium (SPE).

In the above sections, we design the mechanism of the proposed decision model under the assumption of unanimous decision in which all agents and the principal have to agree to introduce the system. Following which, we examine the condition that the subset of equilibrium outcome which satisfy "everybody advantage" principle can exist. Though the analysis of the condition, we examine the effects on working hours by adopting the "white collar" exemption system.

6.1 Effect on working hours

In Section 3 and 4, we assumed that a worker's productivity is constant and does not change. This is an optimistic assumption such that a worker always works hard as much as possible under any payment scheme.

Under constant productivity assumption, analysis in Section 4.2.1 reveals that working hours under the "white collar" exemption system is shorter than those under current work-hour payment scheme. This result shows that the adoption of new payment scheme has potential to contribute to reduce working hours.

6.2 Effects of overtime premium on working hours

Elements which influence working hours under the work-hour payment scheme are examined in Section 4.1.2. The analysis reveals that the working hours contribute to expand as the over time premium rises. Also, the higher productivity contributes to expand them.

6.4 Defence of assumptions

Let me defend the assumption of the unanimous decision making. As for the decision making in industrial relations, the principal can negotiate a payment scheme with each agent. In this case, worker's bargaining power is inferior to the principal. Therefore, the

workers group usually requires a collective agreement with the principal when the modification of major working conditions, such as the adoption of new payment scheme, is proposed. Actually, the draft amendment of Japanese labour standard law requires collective agreement to introduce the "white collar" exemption system. Therefore, the assumption of unanimous decision is reasonable enough to be employed in the models.

The other main assumption we employed in the models is the individual decision of the principal and the agents is done in order, not simultaneously (multi-stage approach). As stated in Section 3.2, each agent's payoff is independent from other agent's payoff, and decision is made unanimously. Therefore, the game in extensive form which is composed of a sequence of single-agent games is appropriate. Also, in reflection of the actual decision making, in formal discussion i.e. a general assembly in an international organization, participants' pros and cons on an agenda are normally expressed in turn. Likewise, discussion of workers group normally employs this procedure. The game in extensive form is an analogy of this decision making process.

The simultaneous decision making such as vote may reach different result from the one in this consecutive decision making process. However, under the assumption of unanimous decision, the difference is not so serious. Because if all players know each player has the veto power, a player expresses his/her rejection to the proposal whatever decision process is a simultaneous vote or a consecutive process. Therefore, the model in the use of the game in extensive form is reliable enough to analyze the effects of the new payment scheme on working environment.

6.4 Further research needs

This thesis is assumed productivity of each agent is uniform. In actual situation, however, a worker is highly productive and the other maybe not. For the principal, such

productivity is hardly observed a priori. The principal may want to select payment schemes in accordance with each worker's productivity, i.e. the principal may want propose the white collar exemption system to low productive workers, and hourly-paid system to high productive workers. Such the principal's intention can be realized under the condition of unanimous decision? To address these issues, we need to employ "mechanism design" or "implementation theory", in particular the "multi-stage mechanism" proposed by Moor and Repullo (Moore and Repullo 1988) .

Also, as stated in Assumption 1 of Section 3.1, we assumed that any exogenous risk to the principal's gross wealth is zero. In this assumption, agents can estimate perfectly sure not only their output but also their payment based on their output. Normally, output is influenced by exogenous factors, and agent's effort dose not necessary provide positive outcome to the principal. Because of agents' risk aversion nature, Assumption 1 is assumed to have positive bias for agents to select the new payment scheme than the work-hour payment scheme. This bias is not examined in this present paper, and it highlights the needs of further research.

In addition, we should take into account the effect of "team working," which is common in Japanese working environment on working hours and productivity. In the "team working," it is difficult to clarify each team member's contribution to outcome, and it is likely that some members "free ride" the other member's contribution. Such phenomenon decreases incentives of workers. Moreover, the "team working" restricts each member's decision of working hours to maximize an outcome of the team. These factors give negative effects to adopt the performance based payment scheme.

Furthermore, functioning external labour market is also a factor to be examined. Ill-functioned labour market decreases reservation utilities of workers and their

bargaining power. Numerical simulation is suggested to analyze this effect on working hours.

REFERENCES

Alchian, A. (1972). Production, Information Costs and Economic Organization, *American Economic Review* 62 (5) :777-795.

Appelbaum, E. (2007). Strategic militancy and the probability of strikes in union-form bargaining. *Labour Economics*, doi:10.1016/j.lbeco.2007.06.013

Basu, K. (1995). Stackelberg equilibrium in oligopoly: An explanation based on managerial incentives, *Economic Letters* 49: 459-464.

Fujino, Y., Matsuda, S. (2007). Health Impact Assessment of "White-Collar Exemption" in Japan. *Journal of Occupational Health* 49 (2): 45-53

Grossman, S., Hard, O. (1983). An Analysis of the Principal-Agent-Problem, *Econometorica* 51(1): 7-45.

Holmstrom, B. (1979). Moral Hazard and Observability, *Bell Journal of Economics* 13: 324-340.

Hurwicz, L. (1960). Optimality and infomational efficiency in resource allocation processes, Mathematical Methods in the Social Science, Edited by Arrow, K. J., Karlin, S., Suppes, P., Stanford: Stanford University Press

Ishige, A. (2005). A Study on White-Collar Exemption, *Waseda Business Review* 38: 23-39

Maskin, E. (1977). Nash equilibrium and welfare optimality, Review of Economic Studies 66 pp.23-38

Moore,J.,Repullo, R., (1988). SUBGAME PERFECT IMPLEMENTATION, *Econometorica,* Vol.56, No.5, pp.1191-1220

Moor, J., Implementation, contracts, and renegotiation in environments with complete information. In Laffont, J., eds, *Advances in Economic Theory; Sixth World Congress*, Vol.I, Chapter 5, pp.182-282. Cambridge University Press, Cambridge, 1992

Rawls, J. (1971) A Theory of Justice, Clarendon Press, Oxford

Ruiz-Verdu, P. (2007). The economics of union organization: Efficiency, information and profitability. *Labour Economics* 14:848-868.

Selten,R. (1975). Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Game, *International Journal of Game Theory*, 4:25-55.

Shavell, S. (1979). Risk-sharing and Incentives in the Principal-Agent Relationship, *Bell Journal of Economics* 10: 55-73.

Spremann, K. (1987). Agent and Principal, in *Agency Theory, Information,and Incentives*, (G. Bamberg and K.Spremann ed.), Springer Verlag, Berlin-Hydelberg

Tasnadi, A. (2005) A way of explaining unemployment through a wage-setting game. *Labour Economics* 12: 191-203.

The American Chamber of Commerce in Japan. (2007). Modernize Work Hours Regulation and Establish a White-Collar Exemption System, *ACCJ Viewpoint*.

Yasui, S. Inohara, T. (2007). A Graph Model of Unanimous Decision Systems. *Proceedings of 2007 IEEE International Conference on Systems, Man and Cybernetics*, Montreal, Quebec, Canada: 1752-1757.

APPENDIX A: Detailed Analysis for Section 4.1.2

A.1. Determination of the Agent's effort

Following the "backward induction" methodology, let us start from the agent's determination of its effort at u_{a2}. In this move, the agent chooses its effort to maximise its payoff $h_a(s_0)$. Such e_n satisfies the following:

$$\frac{\partial h_a(s_0)}{\partial e_0} = \frac{\partial(b + gb/h(e_c - h) - (e_0)^2)}{\partial e_0} = 0 \quad \text{and} \quad e_0 \geq h.$$

Then,

If $h \leq gb/2h$, $e_0 = gb/2h$, then $\max h_a(s_0) = g^2b^2/2h^2 - gb + b$, and

If $h > gb/2h$, $e_0 = h$, then $\max h_a(s_0) = b - g^2b^2/4h^2$.

In this case, we can assume $h \leq gb/2h$, because if $h > gb/2h$, then the maximum payoff of the agent can be given at $e_0 < h$. This means that the full-time employment is disadvantageous for an agent, and it is unlikely make a contract with the principal under this condition. Therefore, we assume $h \leq gb/2h$ in the latter part of the analysis.

A.2. Selection of "accept" or "reject" the Proposed New Payment Scheme

Next, at the move of u_{a1}, the agent decides "accept" or "reject" the proposed new payment scheme. The agent chooses larger payoffs in this move. Therefore, the following formula is one of the necessary conditions under which the strategy s_n is an SPE of the game:

$$\max h_a(s_0) \geq h_a(s_{-1}).$$

Then,

$g^2b^2/2h^2 - gb + b \geq R_0$.

A.3 Determination of the Proposed Work-hour Payment Scheme

Then, at move of u_{p2}, the principal determines details of proposed payment scheme (value of *v*) in order to maximize the principal's payoff $h_p(s_0)$ under the condition that the agent accepts the proposed new payment scheme ($e_0 = gb/2h$). Then,

$$\max h_p(s_0) = te_0 - (b + v(e_0 - h)) = -g^2b^2/2h^2 + b(gt/2h - 1 + g).$$

The value of *d* which maximize w_p^{sq} satisfies following:

$$\frac{\partial h_p(s_0)}{\partial b} = \frac{\partial(-g^2b^2/2h^2 + b(gt/2h - 1 + g)}{\partial b} = 0.$$

Then, $b^* = h^2/g^2(gt/2h - 1 + g)$ and $e_0 = t_1/4 + 2/h - h/g$.

A.4 Selection of "propose" or "not propose" the New Payment Scheme

Lastly, the principal chooses "propose" or "not propose" the new payment scheme. Therefore, $\max h_p(s_0) > \max h_p(s_{-1})$ is one of the necessary conditions under which the strategy s_n is an SPE of the game. Then,

$$\max h_p(s_0) = h^2/2g^2(gt/2h - 1 + g)^2 > 0 = \max h_p(s_{-1}).$$

Thus, the principal selects "propose" no matter what overtime premium and worker's productivity.

APPENDIX B: Detailed Analysis for Section 4.2.1

B.1 Determination of the Agent's effort

Following the "backward induction" methodology, let us start from the agent's determination of its effort at u_{a2}. In this move, the agent chose its effort to maximise its payoff $h_a(s_n)$. Thus, e_n satisfies the following:

$$\frac{\partial h_a(s_n)}{\partial e_n} = \frac{\partial(r + dte_n - e_n^2)}{\partial e_n} = 0.$$

then, $e_n = td/2$ and $\max h_a(s_n) = r + t^2 d^2/4$.

According to the analysis of the basement model, agent's payoff $h_a(s_0)$ satisfies the following:

$e_c = gb/2h$, $\max h_a(s_0) = g^2 b^2/2h^2 - gb + b$, where $h < gb/2h$, and

$b^* = h^2/g^2(gt/2h - 1 + g)$.

Then, $e_c = gb^*/2h = t/4 + h/2(1 - 1/g)$. Therefore,

$$\max h_a(s_0) = \frac{h^2}{4g^2}\left(\frac{gt}{2h} - 3(g-1)\right)\left(\frac{gt}{2h} + (g-1)\right).$$

B.2. Selection of "accept" or "reject" the Proposed New Payment Scheme

Next, at the move of u_{a1}, the agent decides "accept" or "reject" the proposed new payment scheme. The agent chooses larger payoffs in this move. Therefore, the following formula is one of the necessary conditions under which the strategy s_n is an SPE of the game:

$$\max h_a(s_n) > \max h_a(s_0).$$

Then,

$$r > h^2 / 4g^2 (gt / 2h - 3(g-1)(gt / 2h + (g-1) - t^2 d^2 / 4 .$$

B.3. Determination of the Proposed New Payment Scheme

Then, at the move of u_{p2}, the principal determines details of proposed payment scheme (value of *r* and *d*) in order to maximize the principal's payoff $h_p(s_n)$ under the condition that the agent accepts the proposed new payment scheme ($e_n = td/2$). Then,

$$h_p(s_n) = te_n - (r - dte_n) = -t^2 d^2 / 2 + t^2 d / 2 - r .$$

The value of *d* which maximize w_p^n satisfies following:

$$\frac{\partial h_p(s_n)}{\partial d} = \frac{\partial(-t^2 d^2 / 2 + t^2 d / 2 - r)}{\partial d} = 0$$

Then, $d^* = 1/2$ and $e_n = t/4$. Thus,

$$\max h_p(s_n) = -t^2 d^{*2} / 2 + t^2 d^* / 2 - r = t^2 / 8 - r .$$

According to the analysis in Section 2.2.1, $e_0 = t/4 + h/2(1-1/g)$. Then,

$$e_0 = e_n + h/2(1-1/g) .$$

B.4. Selection of "propose" or "not propose" the New Payment Scheme

Lastly, at the move of u_{p1}, the principal chooses "propose" or "not propose" the new payment scheme. Therefore, $\max h_p(s_n) > \max h_p(s_0)$ is one of the necessary conditions under which the strategy s_n is an SPE of the game. Then,

$$\max h_p(s_0) = te_0 - (b + gb / h(e_0 - h))$$

$$= h^2 / 2g^2 (gt / 2h - 1 + g)^2 .$$

More, $\max h_p(s_n) > \max h_p(s_c)$ is equivalent to the following:

$$t^2/8 - r > h^2/2g^2(gt/2h - 1 + g)^2 \text{ or } r > t^2/8 - h^2/2g^2(g(t/2h + 1) - 1)^2.$$

Therefore, fixed payment *r* has upper limit, and such limit will rise as overtime premium *g* increases.

Chapter 5

Conclusion

Summary

This chapter summarizes the results of the previous chapters. From the over all analysis in previous chapters, the decision-framework which prioritizes the status quo is able to be coherent with traditional decision making methodologies if they are partially modified. The proposed alternative is advantageously able to satisfy the Rawls's principles of justice, in particular the "generalized" differential principle. The results, however, reveal further research needs which include a development of real-time modeling so that the advantage of the proposed alternative can be fully demonstrated. Besides, the results show that the sequence of decisions is also an important factor to be analyzed, because the status quo or a precondition of the first decision can affect the latter decision in a chain reacted manner. For the analysis of the sequence, the further research of setting a boundary of the scope of a decision, or "protected sphere", is required.

1. Research question and the proposal

This thesis is purposed for seeking a socially justifiable decision-making process by taking into account the non-utility information. This object is motivated by our discomfort from the fact that consequences of the Arrow's impossible theory and the traditional utilitarian welfarism cannot guarantee everybody's advantage in decision-making.

As the solution of the above issue, this thesis proposes the methodology in which the status quo is prioritized, based on the simple question; although the most of decision-making has a status quo as a precondition, why "welfarism" including Arrow's framework and the game theory ignore neither the concept nor the effect of status quo. In the traditional decision theory i.e. the game theory in normal form, an equilibrium point is identified at once, without any recognition of time sequence. The Arrow's social welfare function is also expected to define the solution with no consideration of preconditions of decisions. We, however, know that almost all decision makings are processed in relation to, and affected by, the pre-conditions when a decision takes place, or, status quo.

Based on the above instinctual acknowledgement, we propose putting decision rules such as “prioritizing the status quo” into the traditional decision frameworks. As a benchmark to evaluate the improvements of our proposal, we employ the “Rawls’s principles of justice”, in particular the “generalized” differential principle, by which we could evaluate the deficiency of information limitation of welfarism. In this thesis, we rather focused on the Rawls’s second principle of justice, because it refers to advantageous decisions for all. (Rawls 1971)

The previous chapters examine the proposal applied in different situations and methodologies, and recognize advantages of the proposal compared with the traditional methodologies. In particular, Chapter 2 examines the application to the Arrow’s framework and its social welfare functions. Chapter 2 reveals that the utilization of non-utility information, such as the concept of the status quo, makes it possible to aggregating competing interests in social welfare functions. Chapter 3 focuses the traditional cost benefit analysis to address the issue that limiting the non-utility information may render distribution of “social goods” unfair. Chapter 4 tested whether utilization of non-utility information is feasible in “strategic” actions of players.

2. Summary of the results of previous chapters

Chapter 2 takes care of “Arrow’s general possibility theorem” which has revealed that basic democratic conditions such as “unrestricted domain” (U: each individuals has un-restricted alternatives.), “independence from irrelevant alternative” (I) and “Pareto optimality” (P) cannot satisfy the condition of “non-dictatorship” (D). Arrow’s theorem also affirms impossibility of implementation of multi-criteria decision, which contradicts our instinct and results of some psychological studies (Miller 1956). This chapter examines the proposal which integrates a concept of prioritizes “status quo” and “common shared value” into the traditional Arrow’s framework. The analysis revealed that if the proposed framework is applied to the model, the non-cyclic social orderings could be output in accordance with conditions U, P and D. Although the relaxation of the condition I is inescapable, the use of such restricted non-utility function enables to identify the social ordering which can satisfy Rawls’s second principles of justice.

Chapter 3 focuses on issues of unfair distribution of "social goods" in the use of the traditional cost benefit analysis (CBA) methodology based on Bentham's utilitarianism. In Chapter 3, we propose an alternative CBA methodology which incorporates the concept of the status quo into the traditional CBA and whose guiding concept is the "optimization" of outcomes for the stakeholders. Chapter 3 further proposes evaluation criteria for CBA and thereby provokes advantages of the proposed alternative. The major findings of this chapter reveals that the alternative can guarantee an advantageous condition compared to the status quo for every stakeholders, the second principle of justice, as well as the alternative can satisfy Pareto optimality, which is the benchmark of a social efficiency.

Chapter 4 tested whether the proposed concept which incorporates the concept of status quo can fit to "strategic" actions of players. This chapter introduces the game model in extensive form for the analysis of a unanimous decision in which the status quo is set as the disagreement points; and then finds out the list of payoffs in sub-game perfect equilibrium (SPE). This chapter also introduces the step by step approach, in which the alternatives of players are identified as the list of payoffs in SPE at the first step; and then the identified alternatives are used as the status quo of the next step. The analysis of this chapter confirms the coherency of the proposed framework with principles of social justice in a unanimous decision making in strategic situation.

3. Implication of the generalized differential principles to actual policy making

3.1. Is the principle utopianism?

Rawls's generalized principle guaranties "everybody's advantage" to all stakeholders. There is, however, a possibility that there is no solution which can satisfy the principle. Chapter 3 deals this issue in the application of regulational decision-making model. According to the analysis, possible solutions exist, if payoff functions of players are realistically defined with reflecting actual situation. Even if such solution does not exist in some situations, the analysis of "subsidy model" reveals that transferring payoffs in the use of taxation and subsidy can settle the set of possible solutions which satisfy the

principle.

"Transferring payoffs" may imply "zero-sum game" situation, but actually it does not. The additional payoffs though the subsidy can be provided from "general account" from taxation sources outside of the said game, thus every stakeholder in the game can have benefit from subsidy. We, therefore, can conclude that the principle of "everybody's advantage" is not utopianism.

3.2. Comparison with Utilitarianism and Maxi-min

The decision making with "utilitarianism" and "maxi-min" needs inter-personally comparable cardinal utility. As stated in Chapter 1, this precondition does not apply in Arrows framework and strategic situation, although the decision based on the principle of "everybody's advantage" is possible to be applied in those situations. This is a unique advantage of the principle in comparison with "utilitarianism" and "maxi-min".

Then, if cardinal utility is available, which is better? For example, in the policy making of the international society, "consensus-based decision", which is virtually a unanimous decision making, is widely employed on the basis of the principle such that each country has one vote. In the international society, we can see a lot of examples of stuck decisions, including the climate change agreement, a global trade agreement in the WTO, an introduction of CO2 taxation, although those decisions are expected to provide "maximum utility" for the society as a whole, and should be accepted from the view point of utilitarianism.

Ironically, although a global trade agreement in the WTO does not proceed; bilateral economic packaged agreements (EPA) are widely established between many countries. From the fact that EPA can provide "win-win" relationship, the principle of "everybody's advantage" seems to be more practical than utilitarianism. The decision with "maxi-min" is also estimated to be unsustainable if the decision cannot provide merits to all.

4. Conclusion and needs for further studies

From the overall analysis above, we can conclude that the decision frameworks in which the status quo is prioritized are able to coherent with the traditional decision making frameworks with small modifications. We can also conclude that the proposed alternative is advantageously able to satisfy the Rawls's generalized differential principle.

The results of the thesis emphasize the importance of a formation of "retrospective time sequence" modeling, although most of models employed in traditional decision-making frameworks are "static" modeling. Having revealed that the status quo has an influence to the decision, the timing of decision making should be an important factor of analysis of decision makings. The traditional decision, however, do not involve information of the status quo as a parameter. Since factors which determine the "status quo" are numerous and difficult to be identified in most realistic cases, the status quo should be determined independently as a precondition of a decision. Thereby, the involvement of timing of a decision as a parameter is the best way to identify a specific status quo. The level of environmental regulation, for instance, has largely affected by technical capacities for detecting, controlling, reducing harmful substances, which are largely depending on the level of technologies that is available at a time of a decision.

Besides, a sequence of decisions is also an important factor to be analyzed. This thesis reveals that preconditions of decision-makings can affect the results of decision frameworks in which the status quo is prioritized. Thereby, if a decision-making has a sequence, a chain reaction of the influence of the first preposition could reach a latest decision. This fact underlines the necessity of "multi-stage" modeling which can consider the effect of chain reactions of the questioned decision.

For involving time sequence in the model, another question arises how we can set a boundary of our scope of a decision, or "protected sphere" of interests affected by a decision. Unless that, the sequence of the decision becomes limitless and have a decision be impossible. This question was well pronounced by Sen in his "liberal paradox" theory (Sen 1982). The further study is necessary to indentify how to use non-utility information to determine the scope of social interests to be accomplished by a

decision.

Reference

Miller, G.A., (1956) The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information, Psychological Review, Vol.63, No.2, pp.81-97

Rawls, J. (1971). A Theory of Justice, Clarendon Press, Oxford

Sen, A.K. (1982). Choice, Welfare and Measurement, Harvard University Press, Cambridge, Massachusetts

Bibliography

Alchian, A. (1972). Production, Information Costs and Economic Organization, American Economic Review 62 (5) pp.777-795

Appelbaum, E. (2007). Strategic militancy and the probability of strikes in union-form bargaining, Labour Economics 15 (3) pp.315-333

Arrow, K. J. (1950). A difficulty in the concept of social welfare, Journal of Political Economy 58 pp.328-346

Arrow, K. J. (1951). Social Choice and Individual Values, 1st ed., Wiley, New York

Arrow, K. J. (1963). Social Choice and Individual Values, 2nd ed., Wiley, New York

Arrow, K. J., (1973). Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice, Journal of Philosophy 70 (9) pp.245-263

Arrow, K.J. (1997). Introduction, Social Choice Re-examined, Edited by Arrow, K. J., Sen, A. K. and Suzumura, K. pp. xvi-xvii, Macmillan, London

Arrow, K. J., Hahn, F. (1971). General Competitive Analysis, Holden-Day, San Francisco

Arrow, K. J., Raynaud, H. (1986). Social Choice and Multicriterion Decision Making, MIT Press, Cambridge

Basu, K. (1995). Stackelberg equilibrium in oligopoly: An explanation based on managerial incentives, Economic Letters 49 pp.459-464

Beck B. D., Dudel R., Calabrese E. J. (1994). The Use of Toxicology in the Regulatory Process, Methods of Toxicology, 3rd ed., Edited by Hynes A.W., Raven Press, Ltd., New York, pp.19-58

Campbell, D. E., Kelly, J. S. (2002). Impossibility Theorems in the Arrovian Framework, Handbook of Social Choice and Welfare, Vol.1, Edited by Arrow, K. J., Sen, A.K., Suzumura, K. Elsevier Science B.V., pp.36-85

Cropper M. L. (2000). Has Economic Research Answered the Needs of Environmental Policy? Journal of Environmental Economics and Management 39 pp.328-350

Fujino, Y., Matsuda, S. (2007). Health Impact Assessment of "White-Collar Exemption" in Japan, Journal of Occupational Health 49 (2) pp.45-53

Gaertner, W. (2002). Domain Restrictions, Handbook of Social Choice and Welfare,

Vol.1, Edited by Arrow, K. J., Sen, A. K., Suzumura, K. Elsevier Science B.V., pp. 132-165

Grossman, S., Hard, O. (1983). An Analysis of the Principal-Agent-Problem, Econometrica 51(1) pp.7-45

Harsanyi, J. C. (1955). Cardinal Welfare, individualistic ethics and interpersonal comparisons of utility, Journal of Political Economy 63 pp.309-321

Harsanyi, J. C. Can the Maximin Principle Serve as a Basis for Morality? a Critique of John Rawls's Theory, American Political Science Review 69 (2) pp. 594-606

Holmstrom, B. (1979). Moral Hazard and Observability, Bell Journal of Economics 13 pp. 324-340

Hurwicz, L. (1960). Optimality and infomational efficiency in resource allocation processes, Mathematical Methods in the Social Science, Edited by Arrow, K. J., Karlin, S., Suppes, P., Stanford: Stanford University Press

Ishige, A. (2005). A Study on White-Collar Exemption, Waseda Business Review 38 pp.23-39

James R., Warren D., Helmes N., Roberts S. (2000). Risk Assessment, Principle of Toxicology: Environmental and Industrial Applications, 2nd ed., Edited by Williams P, James R, Roberts S., John Wiley and Sons, New York, pp.437-477

Leontief, W. (1966). Essays in Economics: Theories and Theorizing, Oxford University Press, New York

Marszal, E. M. (2001). Tolerable Risk Guidelines, ISA Transactions 40 pp.391-399

Maskin, E. (1977). Nash equilibrium and welfare optimality, Review of Economic Studies 66 pp.23-38

May, K. O. (1954). Intransitivity, utility and the aggregation of preference patterns, Econometrica 21 pp.1-13

Miller, G. A., (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information, Psychological Review 63 (2) pp.81-97

Moore, J., Repullo, R. (1988). Subgame perfect implementation, Econometrica 56, pp.1191-1220

Nash, J. F. (1950). The Bargaining Problem, Econometrica 18 (2) pp.155-162

Oka T. (2003). Cost-effective Analyses of Chemical Risk Control Policies in Japan,

Chemosherer 53 pp.413-499

Rawls, J. (1971). A Theory of Justice, Clarendon Press, Oxford

Robbins, L. (1932). An Essay on the Nature and Significance of Economic Science, Allen & Unwin, London

Roy A. D. (1952). Safety first and the holding of assets, Econometrica 32 pp.431-449

Ruiz-Verdu, P. (2007). The economics of union organization: Efficiency, information and profitability, Labour Economics 14 pp.848-868

Rutherford D. (1995). The Routledge Dictionary of Economics, Routledge Press, London and New York, p.343,

Selten, R. (1975). Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Game, International Journal of Game Theory 4 pp.25-55

Sen, A. K. (1970). Collective Choice and Social Welfare, Holden-Day Inc

Sen, A. K. (1973). On Economic Inequality, Clarendon Press, New York

Sen, A. K. (1982). Choice, Welfare and Measurement, Harvard University Press, Cambridge, Massachusetts

Shavell, S. (1979). Risk-sharing and Incentives in the Principal-Agent Relationship, Bell Journal of Economics 10 pp.55-73

Spremann, K. (1987). Agent and Principal, in Agency Theory, Information, and Incentives, Edited by G. Bamberg, G., Spremann K., Springer Verlag, Berlin-Hydelberg

Tasnadi, A. (2005). A way of explaining unemployment through a wage-setting game, Labour Economics 12 pp.191-203

The American Chamber of Commerce in Japan (2007). Modernize Work Hours Regulation and Establish a White-Collar Exemption System, ACCJ Viewpoint

Travis C., Richter S., Crouch E., Wilson R., Klema E. (1987). Cancer risk management. A review of 132 federal regulatory decisions, Environmental Science and Technology 21(5) pp.415-420

Trawen A., Maraste P., Persson U. (2002). International comparison of costs of a fatal causality of road accidents in 1990 and 1999, Accident Analysis and Prevention 34 pp.323-332

US Environmental Protection Agency (1992). Guideline for Exposure Assessment,

EPA/600Z-92/001, pp. 16-49, 124-135. National Center for Environmental Assessment, Washington DC

US Environmental Protection Agency (1999). The Benefit of and Costs of the Clean Air Act: 1990-2010. EPA-410-R-99-001, Washington DC

US NAS/NRC (1983). Risk Assessment in the Federal Government: Managing the Process

US Presidential Commission on Risk and Risk Management (1997). P/C Commission, Washington DC

van Delft, A. and Nijkamp, P. (1977). Multi-criteria Analysis and Regional Decision-making, Springer

von Neumann J. and Morgenstern, O. (1944). Theory of Games and Economic Behavior, Princeton University Press, Princeton

Wilson R. Crouch E. A. C. (2001). Risk Benefit Analysis, 2nd ed., pp. 101-133. Harvard University Press, Cambridge

Yasui. S. (2005). A Critical Review of the Traditional Methodology of Cost Benefit Analysis and a Proposed Alternative, Journal of Human and Ecological Risk Assessment 11 pp.411-432

Yasui, S. (2007). The Analysis of a “White Collar” Exemption System on Working Hours, Proceedings of the 51st Meeting of the International Society for the Systems Sciences, Tokyo, #2007-705

Yasui, S. (2008). Conflict between Personal Freedom and Democratic Decision-making, Proceedings of 3rd Tokyo Conference on Argumentation, Tokyo, Japan, pp.291-294

Yasui, S. Inohara, T. (2007). A Graph Model of Unanimous Decision Systems, Proceedings of 2007 IEEE International Conference on Systems, Man and Cybernetics, Montreal, Quebec, Canada: 1752-1757

Acknowledgement

In acknowledgement of the devoted support to me over the past seven years, I would like to express my greatest appreciation to Dr. Takehiro Inohara for so generously forgiving my dragging-feet achievement, for giving so freely and cordially of his time, for sharing his expertise of mathematics and game theory, and for being the kind of editor, without whom this thesis could never have been written.