Disinformation, Benefit of Group Membership, and Conscience: An Economic Model

Harashima, Taiji

Syouei Economic Research Institute

28 March 2025

Disinformation, Benefit of Group Membership, and Conscience: An Economic Model

HARASHIMA Taiji*

March 2025

Abstract

Mutual trust and confidence are important to succeed not only in business but also in many other activities, which means that people do not necessarily behave totally selfishly, even in profit-maximizing activities. Non-selfish behaviors, in other words, behaviors induced by conscience and guilt, are particularly important to actualize the benefit of a group, team, organization, nation, or society as a whole. In this paper, I examine the role of conscience for the benefit of group membership and construct an economic model of conscience. Combining this model with a model of disinformation and mutual trust, I show that the level of conscience in an economy (society) is positively correlated with its efficiency and productivity. Furthermore, on the basis of this model, I discuss the political impact of conscience on, and argue that it will not be easy to establish, a democracy in a nation of low average conscience.

JEL Classification code: D80, D91, E71

Keywords: The benefit of group membership; Conscience; Democracy; Disinformation; Economic efficiency

*Correspondence: HARASHIMA (Family name) Taiji (First name), Syouei Economic Research Institute, Kanazawa, Ishikawa, Japan. Email: t-harashima@mve.biglobe.ne.jp.

1 INTRODUCTION

In economics, it is usually assumed that each person behaves non-cooperatively to maximize his or her own expected utility, which means that people are assumed to be entities who behave selfishly, where "selfishly" here means behaving only considering one's own welfare and not cooperating with or helping other persons. Even under this assumption (i.e., that people are assumed to behave totally selfishly), economics draws an essential conclusion that harmony and stability are naturally established in an economy thanks to the "invisible hand". However, in actuality, mutual trust and confidence are important to succeed in business and many other activities. The importance of mutual trust and confidence implies that, in reality, people do not necessarily behave totally selfishly.

Furthermore, reciprocal and altruistic behaviors have been observed and reported in many psychological laboratory experiments (see e.g., Camerer and Thaler, 1995; Nowak et al., 2000) and have been studied in many research fields, particularly in behavioral game theory (Fehr et al., 1997; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Fehr and Rockenbach, 2003; Dufwenberg and Kirchsteiger, 2004; Dana et al., 2007; Battigalli and Dufwenberg, 2007; Miettinen, 2013; Di Tella, et al., 2015; Exley, 2016; Dufwenberg and Dufwenberg, 2018; Garcia et al., 2020). Behavioral game theory is a derivative of game theory that deviates from game theory in that players are not necessarily assumed to be selfish. In behavioral game theory, some kinds of disutility derived from selfish behaviors, which originate in guilt or remorse (i.e., the conscience of a player), are assumed in utility functions that represent non-selfish behavior in game.

However, behavioral game theory deals with only "micro" level exchanges, contracts, or dealings among a few people in strategic situations because its analyses basically depend on experimental games. Hence, it usually does not consider economy-wide or "macro" level phenomena. Furthermore, to the best of my knowledge, not only behavioral game theory but studies in economics in general have not examined the role of reciprocal and altruistic behaviors (or behaviors influenced by conscience) from a macroeconomic point of view.

In addition, behavioral game theory has not considered the factor of disinformation, at least not directly, even though disinformation seems to be an important element in examining reciprocity, altruism, conscience, and guilt because the effect of disinformation is greatly influenced by mutual trust, as Harashima (2024) indicated. It is highly likely that the level of mutual trust among people is positively correlated with their propensities to engage in reciprocity and altruism as well as their degrees of conscience. The aim of this paper is to examine the role of conscience in the benefit of group membership from a macroeconomic point of view on the basis of the models of

disinformation and mutual trust developed in Harashima (2023a, 2023b, 2024).

An important nature of the benefit of group membership is that self-sacrifice is required of each member of a group for all the members to collectively obtain and enjoy the benefit of group membership. That is, there is a trade-off relationship between the benefit of group membership and each member's "personal gain" (i.e., the gains that could be individually and separately obtained if the member did not sacrifice him/herself for the sake of the group). Because of this trade-off relationship, each member should consider which is more important between the benefit of group membership and the personal gain and how much more important it is. Because people's preferences are crucial in this process, a utility function with regard to them can be assumed.

Here, I assume a simple utility function with regard to the benefit of group membership and personal gain and construct an economic model of conscience on the basis of this utility function. The preference between the benefit of group membership and personal gain in this utility function is influenced by two factors. One is a congenital factor that represents the intrinsic degree of conscience or guilt and is given exogenously and intrinsically, and it is almost unchanged throughout a person's life. The other is an environmental factor that can change depending on surrounding circumstances (e.g., the ratio of selfish people surrounding a member). By maximizing utilities with regard to the benefit of group membership and personal gain, the "optimal" amount of personal gain can be determined.

Combining this economic model of conscience with the model of disinformation and mutual trust presented in Harashima (2023a, 2023b, 2024), I show that the level of conscience in an economy (society) is positively correlated with its level of mutual trust. On the other hand, Harashima (2024) showed that the level of mutual trust and economic efficiency in an economy (society) are positively correlated. Combining these two correlations, I show that the level of conscience in an economy (society) is positively correlated with its level of economic efficiency. As an economy's (society's) level of conscience is higher, its economic efficiency (e.g., productivity) is higher. In this sense, a high level of conscience of people in an economy (society) is important for the economy (society) to prosper.

The above shown relationship means that although the "invisible hand" can establish harmony, stability, or equilibrium even in an economy that consists only of totally selfish people, it cannot necessarily guarantee prosperity in the sense that the levels of production and consumption are high at the achieved equilibrium. For an economy (society) to prosper, people who possess high levels of conscience, empathy, sympathy, and compassion are necessary.

Finally, I discuss the impact of conscience on political aspects as an extension of the economic model of conscience. The model indicates that if many people in a society behave too selfishly, the benefit of group membership becomes negative, which will

cause the society fall into disorder, chaos, and anarchy unless somebody, usually a dictator, forcibly establishes order. This means that it will not be easy to establish a democracy in a nation of low average conscience. Nevertheless, even in a nation of high average conscience, a dictatorship may prevail over democracy during a period of economic or political turmoil (e.g., the Great Depression in the 1930s).

2 DISINFORMATION AND MUTUAL TRUST

In this section, I explain the concept of ranked information and the model of disinformation on the basis of Harashima (2023a, 2023b). In addition, I also explain the relationship between disinformation and mutual trust on the basis of Harashima (2024).

2.1 Ranked information

The model of disinformation in Harashima (2023a, 2023b) was constructed on the basis of the concept of ranked information presented in Harashima (2022). Hence, I first briefly explain this concept in this section.

I refer to a piece of information as an "Inf-piece". Let $IP_{i,w}$ be an Inf-piece with the serial number *w* for purpose *i*. I also refer to a set of Inf-pieces as an "Inf-set". All Inf-sets consist of *n* Inf-pieces. Let IS_i be the Inf-set that is selected for purpose *i* from among all existing Inf-pieces. Let $IS_{i,w}$ indicate that Inf-piece *w* (i.e., $IP_{i,w}$) is included in IS_i.

Let $y(\cdot)$ be the Inf-set production function, where the production function represents the probability to achieve a purpose. A higher value of *y* for an Inf-set corresponds to a higher probability that the Inf-set will achieve the purpose. For purpose *i*, if the Inf-pieces in $IS_{i,s}$ and $IS_{i,r}$ are identical except for IP_s and IP_r and $s < r$, then

$$y(IS_{i,s}) > y(IS_{i,r})$$

for any *s* and *r*.

Each Inf-piece has a particular value, and the value of an Inf-set is equal to the sum of values of the Inf-pieces of which the Inf-set consists. The value of $IP_{i,w}$ will likely be described by an exponentially increasing function of $N - w$. Here, let $\widetilde{IS}_{i,w}$ be the average value of Inf-sets in which the Inf-piece with rank *w* is included. The value of the Inf-set can be approximated by an exponentially increasing function of $N - w$; that is, $\widetilde{IS}_{i,w}$ increases exponentially as the rank of Inf-piece *w* rises.

The distance between each Inf-set and the correct Inf-set (i.e., the top-rank Inf-

set) can be defined as follows. Let $\Theta_{i,h}$ be the Inf-set with the number $h(\in \mathbb{N})$ for purpose i. Here, let $IS_{i,w}\big|_{\Theta_{i,h}} = \sum_{IP_{i,w}\in\Theta_{i,h}} IP_{i,w}$ and $IS_{i,w}\big|_{w=1,2,\dots,n} = \sum_{w=1}^{n} IP_{i,w}$ The distance of Inf-set (DIS) of Inf-set $\Theta_{i,h}$ is defined by

$$D_{i,h} = 1 - \frac{y\left(IS_{i,w}\big|_{\Theta_{i,h}}\right)}{y\left(IS_{i,w}\big|_{w=1,2,\dots,n}\right)} = 1 - \frac{y\left(\sum_{IP_{i,w}\in\Theta_{i,h}} IP_{i,w}\right)}{y\left(\sum_{w=1}^{n} IP_{i,w}\right)} \ .$$

2.2 Disinformation

I define disinformation as a part of misinformation that is deliberately disseminated by a person to obtain utility by making other people's behaviors change, as presented in Harashima (2023a, 2023b). As a result of dissemination of disinformation, the Inf-pieces ranks are distorted.

Suppose that for purpose *i*, a person selects Inf-set *x* if a piece of disinformation *z* is not disseminated, but selects Inf-set *z* if it is. Disinformation will degrade the value of the Inf-set and increase DIS, and therefore,

$$D_{i,x} \leq D_{i,z} \ . \tag{1}$$

Inequality (1) means that the probability of achieving a purpose decreases because of disinformation, and therefore,

$$y\left(IS_{i,w}\big|_{\Theta_{i,x}}\right) \geq y\left(IS_{i,w}\big|_{\Theta_{i,z}}\right) \ . \tag{2}$$

Let $\boldsymbol{\Theta}_{i,m}$ be the set of all Inf-sets in which the highest rank Inf-piece is commonly $IP_{i,m}$. In addition, let $\boldsymbol{D}_{i,m}$ be the average DIS of $\Theta_{i,h} \in \boldsymbol{\Theta}_{i,m}$ such that

$$\boldsymbol{D}_{i,m} = E\left(D_{i,h}\big|_{\boldsymbol{\Theta}_{i,m}}\right) \ ,$$

where *E* is an operator and means that $\boldsymbol{D}_{i,m}$ is the average DIS of all Inf-sets that are included in $\boldsymbol{\Theta}_{i,m}$. Evidently, if $m > l$,

$$\boldsymbol{D}_{i,m} < \boldsymbol{D}_{i,l} \ .$$

That is, $\boldsymbol{D}_{i,m}$ is a decreasing function of the value of $IP_{i,m}$, which means that it is an increasing function of *m*.

Inequality (2) indicates that, because of disinformation, the levels of efficiency in not only individual economic activities but also the entire economy decrease. If the reductions in efficiency indicated by inequality (2) occur in the process of production, total factor productivity (TFP) decreases, and if they occur in the process of investment, the success rate of investment is lowered (see Harashima, 2021a).

2.3 Dissemination of disinformation

A person who disseminates disinformation (i.e., the "disseminator") behaves to maximize rewards obtained by manipulating other persons (i.e., by distorting their inf-sets). Let m be the highest rank inf-piece in the inf-set of a person. Suppose that m is continuous ($0 \leq m$), and therefore, $m = 0$ indicates the top rank, and that initially $m = 0$ for any person. I define the level of manipulation such that the level of manipulation is ψ if the highest rank inf-piece m is aimed to be changed from 0 to ψ (> 0). A larger value of ψ means more manipulation.

As shown in Section 2.2, $\boldsymbol{D}_{i,m}$ is most likely an increasing function of m, and as $\boldsymbol{D}_{i,m}$ increases (i.e., as m increases), it is more apparent that disinformation is present. Hence, the probability a person becomes aware of the manipulation (i.e., the probability that disinformation is uncovered) will increase as ψ increases. Considering the nature of $\boldsymbol{D}_{i,m}$ shown in Section 2.2, the probability of uncovering disinformation (the "uncovering rate"; P) can be most simply modeled as

$$P = 1 - e^{-\delta\psi}\,, \tag{3}$$

where δ is a positive constant.

As ψ increases, the rewards obtained by a disseminator when he or she is successful (i.e., disinformation is not uncovered) will increase in proportion to the corresponding increase in probability that a person naively and wrongly believes the disinformation that is aimed to be included in the person's Inf-set. The reward to the disseminator per piece of disinformation when the dissemination of disinformation succeeds (R) can be most simply modeled as

$$R = \alpha\left(1 - e^{-\zeta\psi}\right),$$

where α and ζ are positive constants.

A disseminator sets a level of manipulation ψ so as to maximize the expected reward. Harashima (2023b) showed that the expected reward is maximized if ψ is set to satisfy

$$\psi = \zeta^{-1}\ln\left(1+\frac{\zeta}{\delta}\right)(>0)\ .$$

2.4 Externality of uncovering disinformation

Suppose that a person makes a judgement about whether a piece of information is disinformation by observing other people's judgements, and

$$\delta = \delta_0(1-\eta)^{-\lambda}\ , \tag{4}$$

where λ (> 0) is a parameter, η ($0 < \eta \leq 1$) is the ratio of the number of people who uncover disinformation to all people, and δ_0 is the value of δ when a person judges a piece of information is disinformation without considering other people's judgements (thereby, $\eta = 0$). When combined with equation (4), equation (3) is transformed to

$$P = 1 - e^{-\psi\delta_0(1-\eta)^{-\lambda}}\ . \tag{5}$$

δ_0 will differ depending on the intelligence (particularly fluid intelligence) of each person, but it may be also affected by the accessibility of information.

Suppose for simplicity that all persons equally have equation (5) in their minds with the same values of ψ, δ_0, and λ, and they equally initially guess that $\eta = \eta_0$. Harashima (2024) showed that all persons eventually reach a state that satisfies

$$\eta = P = 1 - e^{-\psi\delta_0(1-\eta)^{-\lambda}}\ . \tag{6}$$

2.5 Mutual trust and stable state

A larger value of λ means that a person believes and accepts other people's opinions with less doubt and believes that it is less likely they are being deceived or intentionally told disinformation. After repeatedly experiencing states that satisfy equation (6) for many pieces of disinformation, most people in an economy (society) will be able to roughly correctly know, estimate, guess, or feel the value of λ (i.e., the level of mutual trust) in the economy. This means that the level of mutual trust of an economy can be represented by λ. If an economy's mean value of λ is larger, its level of mutual trust is higher and vice versa.

2.6 Mutual trust and economic efficiency

Harashima (2024) showed that a high level of mutual trust greatly restrains disinformation from spreading in an economy and consequently will considerably increase various

aspects of economic efficiency. That is, mutual trust is important for achieving high levels of economic, social, and perhaps political activities.

3 ECONOMIC MODEL OF CONSCIENCE

3.1 The benefit of group membership

If people work together cooperatively, reciprocally, or altruistically, they can achieve what they cannot if each of them works alone uncooperatively and selfishly, and they can complete very large-scale projects. These achievements greatly benefit all members of the group and exist apart from personal gains that can be obtained and enjoyed individually and separately by each member. It seems likely that establishing and maintaining groups to collectively obtain and enjoy this benefit will have been one of the top priorities for humans to survive and prosper.

Nevertheless, receiving the benefit of group membership requires members to restrain themselves from behaving selfishly, and sometimes to sacrifice themselves to some extent for the sake of the group. The phrase "there is no I in team" is common in team sports, and it means victory (success) requires the self-sacrifice of team members. This nature is common not only to team sports but most activities that require group cooperation.

Note that a deeper reason or mechanism for cooperative, reciprocal, or altruistic behaviors may be explained by models of altruism in the theory of evolution, but this is beyond the scope of this paper and I do not refer to this literature. Note also that the concept of the benefit of group membership will be closely related to the ability-to-pay principle in the theory of tax.

3.2 Utilities from the benefit of group membership and personal gain

The existence of these often-observed cooperative, reciprocal, or altruistic behaviors strongly suggests that people must obtain utilities not only from personal gains that are individually and separately obtained but also from the benefit of group membership that is obtained collectively because people would otherwise never prioritize the benefit of a group.

As mentioned above, however, a member's cooperative, reciprocal, or altruistic behavior for the sake of the benefit of group membership requires the member to relinquish some individual personal gains; that is, self-sacrifice is required to prioritize the benefit of group membership over personal gain. Conversely, if a member behaves selfishly and pursues personal gain instead of behaving for the best interests of the group,

the member who does not behave cooperatively, reciprocally, or altruistically can alone enjoy the personal gain in addition to the benefit of group membership, although the benefit of group membership may become smaller because of the actions of this member. Thus, there is a trade-off relationship between the benefit of group membership and personal gain.

Because of this trade-off relationship, each group member has to decide which should be given greater importance in each instance, in other words, which the member prefers. A decision made by preference can be expressed by a utility function, in this case, a utility function with regard to the benefit of group membership and personal gain.

3.3 Conscience and guilt

To make a decision on the basis of this utility function, a member first has to know, estimate, guess, or feel (1) the amount of the benefit of the group, (2) the amount of the member's personal gain, and (3) the amount of damage on the benefit of the group done by the member's selfish behavior to obtain the personal gain. Of these amounts, the first two should be relatively easy to intuit, at least roughly, but the third amount may not be because the impact of the member's furtive "sabotage" of the activities of the group may be complex and uncertain, and mostly consist of indirect effects and thus be considerably unpredictable. A little leak may sink a great ship.

If it is not easy to even roughly estimate amount (3), some other ways to subjectively evaluate it are required. One of the likely alternative ways is to behave according to one's conscience, guilt, empathy, sympathy, and compassion without directly estimating or guessing that amount. Before a member behaves selfishly and pursues personal gain, that member will imagine and feel other members' disadvantage or disutility that would be suffered by the member's selfish behavior. In many cases, the member will feel some degree of remorse or guilt (i.e., conscience). If the strength of the remorse and guilt is strong, the member will stop behaving selfishly. In this sense, a member's strength of remorse and guilt will be roughly proportionate to the member's subjective evaluation of decrease in the benefit of group membership due to the member's selfish behavior. This means that conscience can be used as a surrogate variable of the subjective evaluation of the damage to the benefit of group membership done by the member's selfish behavior to obtain the personal gain.

3.4 Utility function

One of the simplest utility functions that express the preference between the benefit of group membership and the personal gain of group member i (= 1, 2, …, N) is

$$U_i = u_P(c_{P,i}) + u_W(c_{W,i}) , \tag{7}$$

where

$$c_{W,i} = C - \chi_i c_{P,i} \tag{8}$$

and U_i is the utility of member i, N is the number of members, $u_P(\cdot)$ and $u_W(\cdot)$ are utility functions with regard to the personal gain and the benefit of group membership, respectively, $c_{P,i}$ and $c_{W,i}$ are respectively the personal gain and the benefit of group membership felt by member i when that member obtained $c_{P,i}$, C is the benefit of the group per member when $c_{P,i} = 0$ for any i (i.e., no member behaves selfishly), and $\chi_i (> 0)$ is a parameter.

In addition, it is assumed that

$$\frac{du_P(c_{P,i})}{dc_{P,i}} > 0 \tag{9}$$

and

$$\frac{du_W(c_{W,i})}{dc_{W,i}} > 0 \, . \tag{10}$$

By equation (8) and inequality (10),

$$\frac{du_W(c_{W,i})}{dc_{W,i}} = \frac{dc_{P,i}}{dc_{W,i}} \frac{du_W(c_{W,i})}{dc_{P,i}} = -\frac{1}{\chi_i} \frac{du_W(c_{W,i})}{dc_{P,i}} > 0 \, ;$$

thus,

$$\frac{du_W(c_{W,i})}{dc_{P,i}} < 0 \, . \tag{11}$$

In the utility function (i.e., equations (7) and (8) and inequalities (9), (10), and (11)), as the personal gain c_P increases, utilities from it increase. However, at the same time, the utilities from the benefit of group membership decrease because the term $\chi_i c_{P,i}$ is subtracted from C due to the selfish behavior. χ_i represents member i's strength of remorse and guilt, and that strength differs across members depending on the value of parameter χ_i in equation (8). As the value of χ_i increases, utilities from the benefit of group membership decrease. That is, χ_i is the essential parameter that

determines a member's cooperative, reciprocal, and altruistic behaviors urged by the member's remorse and guilt. Note that similar expressions of disutility (i.e., including a negative term in utility function) due to selfish behavior are often used in the utility functions used in behavioral game theory (e.g., Miettinen, 2011).

More specifically, I assume conventional-type utility functions (i.e., constant relative risk aversion utility functions) such that

$$u_P(c_{P,i}) = \frac{c_{P,i}^{1-\tau}}{1-\tau} \tag{12}$$

and

$$u_W(C - \chi_i c_{P,i}) = \frac{(C - \chi_i c_{P,i})^{1-\tau}}{1-\tau} \ , \tag{13}$$

where τ is a positive constant. Finally, if $c_{P,i} = 0$ for any *i* (i.e., no member behaves selfishly and pursues personal gain), the total benefit for the group as a whole (*B*) is

$$B = NC \ .$$

Suppose that the damage of the selfish behavior of a member on the benefit of group membership is "objectively"

$$\tilde{\chi} c_{P,i}$$

for any *i*, where

$$\tilde{\chi} = \bar{\chi} + \varepsilon \ ,$$

and $\bar{\chi}$ is a constant and ε is i.i.d. Hence,

$$E(\tilde{\chi}) = \bar{\chi} \ ,$$

where $E(\cdot)$ is the expectation operator (i.e., on average $\tilde{\chi} = \bar{\chi}$). It seems likely that $\bar{\chi} > 1$ because a small action can have a great consequence. If some members behave selfishly and obtain personal gain (i.e., $c_{P,i} > 0$ for some *i*), the total benefit for the group as a whole (*B*) "objectively" decreases on average to

$$E(B) = NC - \bar{\chi}\sum_{i=1}^{N} c_{P,i} \ . \tag{14}$$

An important point is that the value of $\tilde{\chi}$ is unknown for any member, and each member only "subjectively" feels $\chi_i c_{P,i}$ through their own respective conscience and guilt.

Note that equations (12) and (13) imply that not only χ but also τ may reflect and thus be able to represent the preference between the benefit of group membership and personal gain. However, as is well known, τ usually represents the degree of risk aversion in the constant relative risk aversion utility function, and therefore it seems highly likely that the value of τ is determined by a different mechanism from that for this preference.

3.5 Congenital and environmental factors

If a group member notices that many members unapologetically cheat on some occasions, that person may also be tempted to cheat. This means that a member's strength of conscience will be influenced by the surrounding environment, particularly by how other members behave. It is highly likely that, as the ratio of selfish members increases, the temptation for a member to also behave selfishly also becomes stronger.

One of the simplest functional forms that describe the influence of other members' behaviors is

$$c_{W,i} = C - \frac{\chi_i}{\kappa^{\xi l_i}} c_{P,i} \tag{15}$$

and

$$u_W\left(C - \frac{\chi_i}{\kappa^{\xi l_i}} c_{P,i}\right) = \frac{\left(C - \frac{\chi_i}{\kappa^{\xi l_i}} c_{P,i}\right)^{1-\tau}}{1-\tau} \ , \tag{16}$$

where

$$l_i = \frac{\sum_{j=1}^{N} c_{P,j} - c_{P,i}}{NC} \ ,$$

and $\kappa(>1)$ and $\xi(>0)$ are constants and are assumed to be common to all members for simplicity. l_i represents the influence of other members' behaviors, and at equilibrium, the value of l_i is stabilized for any i. Note that the values of κ and ξ

determine how much a member is affected by l_i.

Parameter χ_i in equation (8) and (13) is replaced with $\frac{\chi_i}{\kappa^{\xi l_i}}$ in equations (15) and (16). Although χ_i is constant, $\frac{\chi_i}{\kappa^{\xi l_i}}$ changes if l_i changes; therefore, in equations (15) and (16), not only χ_i but also l_i affect $u_W\left(C - \frac{\chi_i}{\kappa^{\xi l_i}} c_{P,i}\right)$. Nevertheless χ_i and l_i have essentially different natures. It seems highly likely that a member's value of χ_i is given exogenously and intrinsically and probably remains unchanged throughout his or her life (i.e., it is a congenital factor). On the other hand, the value of l_i can change depending on surrounding circumstances (i.e., it is an environmental factor). This means that, even if a member was born with a high χ_i, the member may often behave selfishly if he or she is surrounded by members with low values of χ_i. Nevertheless, it seems highly likely that the average value of χ_i in an economy (society) (χ_{ave}) is negatively correlated with the average value of l_i in the economy (l_{ave}).

The value of l_i will be particularly important in a situation where a group may fall into a state of confusion or disarray (e.g., just after a natural disaster) because δ_0 may approach zero due to lack of information and the rapid spread of disinformation. The rapid spread of disinformation may cause many people to unconsciously and mistakenly behave far more selfishly than usual; thus, l_i rapidly increases for any member, and a vicious circle will ensue, in which increases in l_i make other people behave even more selfishly.

3.6 The "optimal" amount of personal gain

The utility U_i in equation (7) is maximized when $\frac{dU_i}{dc_{P,i}} = 0$; that is, at a state that satisfies

$$\frac{du_P(c_{P,i})}{dc_{P,i}} = \frac{\chi_i}{\kappa^{\xi l_i}} \frac{du_W\left(C - \frac{\chi_i}{\kappa^{\xi l_i}} c_{P,i}\right)}{d\left(C - \frac{\chi_i}{\kappa^{\xi l_i}} c_{P,i}\right)} . \tag{17}$$

By equations (12), (16), and (17), the "optimal" amount of personal gain for member *i* in the sense that U_i is maximized is

$$c_{P,i} = \frac{C}{\left(\frac{\chi_i}{\kappa^{\xi l_i}}\right)^{\frac{1}{\tau}} + \frac{\chi_i}{\kappa^{\xi l_i}}} . \tag{18}$$

Let $\tilde{c}_{P,i}$ be the optimal $c_{P,i}$ that satisfies equation (18).

Equation (18) indicates that, as χ_i increases, $\tilde{c}_{P,i}$ decreases because an increase in χ_i makes a member more strongly feel the remorse of conscience and guilt (by equations (8) and (15)). That is, the member more strongly feels the importance of the benefit of group membership, and the member is less likely to behave selfishly and pursue personal gain (i.e., $\tilde{c}_{P,i}$ decreases). Furthermore, as χ_i increases, the total benefit for the group as a whole $B = NC - \bar{\chi}\sum_{i=1}^{N} c_{P,i}$ increases because $\tilde{c}_{P,i}$ decreases.

In addition, equation (18) indicates that, as l_i increases, $\tilde{c}_{P,i}$ increases because $\kappa^{\xi l_i}$ increases and $\frac{\chi_i}{\kappa^{\xi l_i}}$ decreases. That is, if a member is surrounded by a larger number of selfish members, the member behaves more selfishly. Note that equation (18) also indicates that as τ decreases, $\tilde{c}_{P,i}$ decreases.

4 CONSCIENCE, MUTUAL TRUST, AND ECONOMIC EFFICIENCY

4.1 Conscience and mutual trust

4.1.1 $\tilde{c}_P$ and mutual trust

As shown in Section 2, if an economy's (society's) value of λ is smaller, its level of mutual trust is lower, but what factor makes an economy's (society's) λ small? There will be many factors, but one of the most important factors will be a high ratio of selfish people in the economy (society). As shown in the above sections, selfish people are less likely to care about the benefit of a group or the welfare of other people, which means that they can more often deceive others and intentionally disseminate disinformation. Clearly, a person who often deceives others and disseminates disinformation will be less trusted. Hence, if the ratio of selfish people is high in an economy (society), the level of mutual trust (λ) in the economy (society) will be low. That is, it seems highly likely that the average $\tilde{c}_{P,i}$ in an economy (society) ($\tilde{c}_{P,ave}$) is negatively correlated with its λ; that is,

$$\frac{d\lambda}{d\tilde{c}_{P,ave}} < 0 \tag{19}$$

will hold.

4.1.2 Positive correlation between conscience and mutual trust

On the other hand, by equation (18),

$$\frac{d\tilde{c}_{P,ave}}{d\chi_{ave}} < 0 \,. \tag{20}$$

Therefore, by inequalities (19) and (20),

$$\frac{d\lambda}{d\tilde{c}_{P,ave}}\frac{d\tilde{c}_{P,ave}}{d\chi_{ave}} = \frac{d\lambda}{d\chi_{ave}} > 0 \,.$$

That is, the levels of conscience (χ_{ave}) and mutual trust (λ) are positively correlated. As an economy's (society's) level of conscience is higher, its level of mutual trust is higher.

One of the simplest functional forms of λ that satisfies inequality (19) is

$$\lambda = z c_{P,ave}^{-v} \,, \tag{21}$$

where z (> 0) and $v(>0)$ are constants. By equation (18),

$$\tilde{c}_{P,ave} = \frac{C}{\left(\frac{\chi_{ave}}{\kappa^{\xi} l_{ave}}\right)^{\frac{1}{\tau}} + \frac{\chi_{ave}}{\kappa^{\xi} l_{ave}}} \,, \tag{22}$$

and by equations (21) and (22),

$$\lambda = \frac{z}{C^{v}}\left[\left(\frac{\chi_{ave}}{\kappa^{\xi} l_{ave}}\right)^{\frac{1}{\tau}} + \frac{\chi_{ave}}{\kappa^{\xi} l_{ave}}\right]^{v} \,. \tag{23}$$

If constants z, v, κ, ξ and τ and the average C are similar across economies (societies), equation (23) indicates that an economy's (society's) level of mutual trust (λ) depends on its level of conscience (χ_{ave}) and the environment (l_{ave}). As mentioned in Section 3.5, χ_{ave} and l_{ave} are highly likely negatively correlated; therefore, equation (23) simply indicates that an economy's (society's) level of mutual trust crucially depends on its level of conscience, and λ and χ_{ave} are positively correlated.

4.2 Effects of conscience on the economy

4.2.1 Conscience and economic efficiency

Harashima (2024) showed that as an economy's (society's) level of mutual trust is higher (i.e., the value of λ is larger), its efficiency in economic activities (e.g., TFP) is higher.

On the other hand, as shown in Section 4.1, as an economy's (society's) level of conscience is higher (i.e., χ_{ave} is larger), its level of mutual trust is higher (i.e., the value of λ is larger). Putting these two relationships together, as an economy's (society's) level of conscience is higher (i.e., χ_{ave} is larger), its efficiency in economic activities (e.g., TFP) is higher. In addition, as equations (14) and (18) indicate, a larger χ_{ave} results in a smaller amount of $c_{P,i}$ and thus a larger amount of the benefit of group membership. That is, a high level of conscience (i.e., a large value of χ_{ave}) of an economy (society) increases the chances that an economy (society) will prosper, for example, by successfully completing large projects.

4.2.2 Conscience, economic rents, and economic inequality

Harashima (2020a) showed that economic rents are derived from "mistakes" made in business deals. These economic rents are generated because the abilities of negotiating agents are different in business deals. Because most business deals and contracts will be made between economic agents whose abilities are different, these economic rents will exist ubiquitously and at a large scale across an economy.

Furthermore, the model of Harashima (2020a) indicates that as one negotiation partner is less "honest", larger amounts of economic rents are extracted by the other negotiation partner. In this case, an "honest" agent is one who, upon recognizing that the other agent is making a mistake, informs the other agent of the mistake. Conversely, a "dishonest" agent covertly aims to gain an advantage from any opportunity the other agent provides (i.e., a dishonest agent does not inform the other agent of mistakes).

A person's level of honesty will be positively proportionate to their level of conscience because a person with a high conscience level will strongly feel remorse and guilt if the person behaves dishonestly. Hence, as an economy's (society's) level of conscience is lower (i.e., its χ_{ave} is smaller), the economic rents derived from mistakes in business deals will be larger. Because these economic rents increase economic inequality and can generate extreme economic inequality (Harashima, 2020a, 2020b[1], 2021b), as the average level of conscience in an economy (society) is lower (i.e., its χ_{ave} is smaller), the level of economic inequality will be higher.

5 DISCUSSION

5.1 Anarchy

Because most group members will behave to achieve their respective optimal $c_{P,i}$ (i.e., $\tilde{c}_{P,i}$), then by equation (14),

[1] Harashima (2020b) is also available in Japanese as Harashima (2021c).

$$E(B) = NC - \bar{\chi} \sum_{i=1}^{N} \tilde{c}_{P,i} \tag{24}$$

will hold on average. Nevertheless, if

$$NC < \bar{\chi} \sum_{i=1}^{N} \tilde{c}_{P,i} \ , \tag{25}$$

then by equation (24),

$$E(B) < 0 \ ,$$

i.e., the amount of the benefit of a group becomes negative. What does this negative value mean? This benefit is an essential driving force that makes members in a team, organization, economy, nation, or society cooperate and unite because if there is no benefit, there will be no reason to unite to form a group. Therefore, if $E(B)$ is negative, people will not behave cooperatively and unite. Instead, they will behave totally selfishly and pursue only personal gain without considering the welfare of other people, resulting in a society that falls into disorder, chaos, and anarchy.

By equation (18), if $\chi_i \to 0$, then $\tilde{c}_{P,i} \to \infty$, and therefore inequality (25) is actualized in this case. That is, if χ_{ave} is sufficiently small (i.e., the level of conscience is sufficiently low), then the group will fall into disorder unless somebody forcibly establishes order.

5.2 Conscience, democracy, and dictatorship

If the value of χ_{ave} is very small and thus inequality (25) holds, how can order in a nation be established? The simplest, perhaps only way to escape from disorder, will be to establish a dictatorship. If a dictator can force people into obedience and to stop pursing their personal gains, order can be established (i.e., $E(B) > 0$) by compulsorily making

$$c_{P,i} < \tilde{c}_{P,i}$$

for most *i* by, for example, severely punishing people who are thought to pursue personal gain. As a result, even if disorder can be prevented, most people cannot achieve their "optimal" amounts of personal gain; that is, the achieved order does not imply an "optimal" state.

A nation with such a very small χ_{ave} should have no lack of candidates for a dictator because the levels of the conscience of many people in the nation are very low. Thus, many people will not hesitate to brutally dominate other people. Conversely, it will not be easy to establish a democracy in such a nation.

On the other hand, it seems likely that if a nation has sufficiently developed economically, a democracy can be established. In an underdeveloped economy, many people can access only smaller pieces of information; therefore, δ_0 will be very small for many pieces of disinformation, which functions in the same way as smaller values of λ and χ_{ave} by equations (6) and (23). However, if an economy is well developed, many people can access many necessary pieces of information; thus, δ_0 can be large enough for a democracy to become established.

Nevertheless, being developed may not necessarily guarantee democracy because δ_0, λ, and χ_{ave} are all important. If the level of conscience of people is low (i.e., χ_{ave} and λ are small), it will be difficult to establish a democracy even though a nation may be economically well developed (i.e., δ_0 is large). Conversely, even if a society is economically underdeveloped (i.e., δ_0 is small), a democracy may be established if the level of conscience is high (i.e., λ and χ_{ave} are large).

However, even in the case of a well-developed nation with high values of χ_{ave} and λ, a dictatorship may prevail over democracy during periods of economic or political turmoil (e.g., during the Great Depression in the 1930s and World War II in the 1940s) because δ_0 may approach zero due to the fast and widespread dissemination of disinformation amid the economic and political turmoil. As a result, many people may unconsciously and mistakenly behave more selfishly than usual, and l_{ave} will increase. People will in turn behave even more selfishly, and as a result, l_{ave} may approach unity, and inequality (25) may eventually hold.

5.3 Bounded conscience

If $\chi_i \to \infty$, then $\tilde{c}_{P,i} \to 0$ by equation (18); that is, if the level of conscience is infinitely high, nobody behaves selfishly and pursues personal gain. However, such a state seems to be quite unrealistic. In reality, χ_{ave} will take some finite value; that is, on average, people usually behave selfishly and pursue personal gain to some extent. This means that there is a mechanism that prevents χ_{ave} from becoming too large. Possible mechanisms may include inappropriate punishments, uncertainty about C and $c_{P,i}$ (i.e., people will make incorrect guesses or estimates), economic inequalities that hinder union formation, and the advantage of selfish behavior in emergency situations.

6 CONCLUDING REMARKS

In economics, it is usually assumed that each person behaves non-cooperatively to maximize his or her own expected utility (i.e., people are assumed to be entities who behave totally selfishly). However, reciprocal and altruistic behaviors have been commonly reported in many psychological laboratory experiments (see e.g., Camerer and Thaler, 1995; Nowak et al., 2000). This means that most people do not necessarily behave totally selfishly. Behind reciprocal and altruistic behaviors lies conscience, guilt, empathy, sympathy, and compassion. Reciprocal and altruistic behaviors or equivalently behaviors urged by conscience, guilt, empathy, sympathy, and compassion will be particularly important and essential to actualize the benefit of a group, team, organization, nation, or society as a whole. Nevertheless, the benefit requires self-sacrifice of each member (i.e., reciprocal and altruistic behaviors).

In this paper, I examined the role of conscience or guilt in the benefit of group membership from a macroeconomic point of view on the basis of the models of disinformation and mutual trust developed in Harashima (2023a, 2023b, 2024). An important point is that there is a trade-off relationship between the benefit of group membership and personal gain; that is, to actualize and obtain the benefit, self-sacrifice is required. Each member has to decide which is more important on the basis of his or her own preference with regard to the benefit of group membership and personal gain. I assumed a simple utility function, in which a member's preference is influenced by two factors: one is congenital and the other is environmental. By maximizing utilities on the basis of this utility function, the "optimal" amount of personal gain is determined.

Combining the economic model of conscience and the model of disinformation and mutual trust presented in Harashima (2024), I showed that the level of conscience in an economy (society) is positively correlated with its economic efficiency (e.g., productivity). A high level of conscience of people in an economy (society) is important for the economy (society) to prosper. This result means that, although the "invisible hand" indeed can establish harmony, stability, or equilibrium even in an economy that consists of totally selfish people, it cannot necessarily guarantee prosperity. For an economy (society) to prosper, people who behave according to their high-level conscience are necessary.

The model also indicates that if too many members behave selfishly and pursue their personal gains, the benefit of group membership becomes negative, which will cause a society fall into disorder and chaos unless somebody, generally a dictator, forcibly establishes order. This means that, in a nation of low average conscience, it will not be easy to establish a democracy. Nevertheless, even in a nation of high average conscience, a dictatorship may prevail over democracy in a period of utter economic or political turmoil (e.g., during the Great Depression in the 1930s and World War II in the 1940s).

References

Battigalli, Pierpaolo and Martin Dufwenberg (2007) "Guilt in Games," *American Economic Review*, Vol. 97, No. 2, pp. 170-176.

Bolton, Gary E. and Axel Ockenfels (2000) "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, Vol. 90, No. 1, pp. 166-193.

Camerer, Colin F. and Richard H. Thaler (1995) " Anomalies: Ultimatums, Dictators and Manners," *Journal of Economic Perspectives*, Vol. 9, No. 2, pp. 209-219.

Dana, Jason, Roberto A. Weber and Jason Xi Kuang (2007) "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness," *Economic Theory*, Vol. 33, No. 1, pp. 67-80.

Di Tella, Rafael, Ricardo Perez-Truglia, Andres Babino and Mariano Sigman (2015) "Conveniently Upset: Avoiding Altruism by Distorting Beliefs about Others' Altruism," *American Economic Review*, Vol. 105, No. 11, pp. 3416–42.

Dufwenberg, Martin and Martin A. Dufwenberg (2018) "Lies in Disguise – A Theoretical Analysis of Cheating," *Journal of Economic Theory*, Vol. 175, pp. 248-264.

Dufwenberg, Martin and Georg Kirchsteiger (2004) "A Theory of Sequential Reciprocity," *Games and Economic Behavior*, Vol. 47, No. 2, pp. 268-298.

Exley, Christine L. (2016) "Excusing Selfishness in Charitable Giving: The Role of Risk," *The Review of Economic Studies*, Vol. 83, No. 2, pp. 587-628.

Fehr, Ernst, Simon Gächter and Georg Kirchsteiger (1997) " Reciprocity as a Contract Enforcement Device: Experimental Evidence," *Econometrica*, Vol. 65, No 4, pp. 833-860.

Fehr, Ernst and Bettina Rockenbach (2003) " Detrimental Effects of Sanctions on Human Altruism," *Nature*, Vol. 422, pp. 137-140.

Fehr, Ernst and Klaus M. Schmidt (1999) "A Theory of Fairness, Competition, and Cooperation," *The Quarterly Journal of Economics*, Vol. 114, No. 3, pp. 817-868.

Garcia, Thomas, Sébastien Massoni and Marie Claire Villeval (2020) "Ambiguity and Excuse-driven Behavior in Charitable Giving," *European Economic Review*, Vol. 124.

Harashima, Taiji (2020a) "Economic Inequality Exacerbated by Economic Rents from Exploitative Contracting," *MPRA* (*The Munich Personal RePEc Archive*) *Paper*, No. 104178.

Harashima, Taiji (2020b) "Preventing Widening Inequality: Economic Rents and Sustainable Heterogeneity," *MPRA* (*The Munich Personal RePEc Archive*) *Paper* No. 103377.

Harashima, Taiji (2021a) "Economic Inequality and Heterogeneous Success Rates of Investment," *MPRA* (*The Munich Personal RePEc Archive*) *Paper* No. 110688.

Harashima, Taiji (2021b) "Consequence of Heterogeneous Economic Rents under the MDC-based Procedure," *Journal of Applied Economic Sciences,* Vol. 16, No. 2. pp.

185-190.
Harashima, Taiji (2021c) “Preventing Widening Inequality: Economic Rents and Sustainable Heterogeneity,” in Japanese, *Journal of Kanazawa Seiryo University*, Vol. 55, No. 1, pp. 17–37.（「格差拡大を防ぐ一経済レントと持続可能な非均質性」『金沢星稜大学論集』第55巻第1号17～37頁）
Harashima, Taiji (2022) “Asymmetric Information and Ranked Information Are Equivalent in Making Information Utilization Heterogeneous,” *Journal of Applied Economic Sciences*, Vol. 17, No. 3, pp. 250 – 261.
Harashima, Taiji (2023a) “An Economic Theory of Disinformation,” *Theoretical and Practical Research in Economic Fields*, Vol. 14, No. 1. pp. 16-27.
Harashima, Taiji (2023b) “Economic Rents Extracted in the Process of Consumption,” *MPRA* (*The Munich Personal RePEc Archive*) *Paper* No. 119012.
Harashima, Taiji (2024) “Disinformation and Mutual Trust: An Economic Model,” *MPRA* (*The Munich Personal RePEc Archive*) *Paper* No. 121865
Miettinen, Topi (2013) “Promises and Conventions – An Approach to Pre-play Agreements,” *Games and Economic Behavior*, Vol. 80, pp. 68-84.
Nowak, Martin A., Karen M. Page and Karl Sigmund (2000) “Fairness Versus Reason in the Ultimatum Game,” *Science*, Vol. 289, No. 5485, pp. 1773-1775.