Irreversibility of change in a model of boundedly rational agents exchanging goods and capital

Takuma Tanaka

Graduate School of Data Science, Shiga University, 1-1-1 Banba, Hikone, Shiga 522-8522, Japan

E-mail: tanaka.takuma@gmail.com

January 2020

Abstract. This paper investigates the direction of change in a model of agents exchanging goods and capital. Agents have heterogeneous utility functions and are not perfectly rational but can refuse exchanges that decrease their own utility. In this model, an exchange is shown to always increase a quantity defined by the amount of production and the time preference of the agents.

Keywords: Irreversibility, agent-based model, utility function

Submitted to: *New J. Phys.*

1. Introduction

Predicting the behavior of agents is one of the most important tasks in social sciences and sociophysics. Economics has succeeded in making quantitative predictions by assuming optimal behaviors of perfectly rational agents, which can maximize the expected outcomes [1, 2]. However, experiments have suggested that humans are not perfectly rational in many situations [3]. Humans are instead *boundedly* rational [4]. To describe human behavior more realistically, a formulation of boundedly rational agents assumes that agents cannot necessarily choose the action that maximizes the expected outcome but can choose actions that lead to an outcome exceeding a predefined threshold [4]. Making these assumptions might preclude quantitative predictions, but they should allow qualitative predictions, such as those of the direction of social and economic change. A direction of change can be predicted in various cases. For instance, if economic change is unidirectional (i.e., irreversible), the direction is predictable. In this case, the direction may be predicted by a measurable quantity, such as entropy in thermodynamics. This is because, if the quantity increases for any possible social and economic change, any change that decreases the quantity should not to happen. Whether or not economic change is irreversible, the quantity measuring irreversibility exists, and human behavior is irreversible should be useful in social modeling. These issues have not been investigated intensively [5, 6].

In this paper, I demonstrate the irreversibility of change using a model of heterogeneous boundedly rational agents whose utility is a function of the consumption of goods in each period. The changes taking place in the model are the exchange of goods and capital among agents, which has been studied intensively in sociophysics [7, 8, 9, 10]. I show that a quantity defined by the amount of production never decreases. I also discuss the potential extensions of the model.

2. Results

Denote the set of agents by $\mathcal{A} = \{1, \ldots, A\}$. At the beginning of each period, agents produce goods, such as foods. We assume that goods cannot be saved, and consequently, goods produced in a period must be consumed during the same period. We also assume that there is capital, such as land, which cannot be directly consumed but can be used to produce goods. Although we need no assumption on the functional dependency of the production of goods with respect to capital (i.e., a production function) to support the main result, it is convenient for us to assume that the production of goods is an increasing function of capital. Agents can exchange goods for capital with others. The total amount of capital remains constant. The increase and decrease of capital owned by an agent affect the agent's production of goods over the subsequent periods. Thus, the agents can increase goods production during subsequent periods by decreasing the goods in the present period and vice versa.

For simplicity, we concentrate on a case in which agents exchange goods and capital only once. Specifically, at the beginning of period 1, agents produce goods by using capital. Then, agents exchange goods and capital and consume goods. During subsequent periods, each agent consumes only the goods that are self-produced. Thus, the amount of goods produced by agent a in period 1, y_a, can be different from the amount of goods consumed by it in period 1, c_a, whereas the amount of goods produced by agent a in each of the subsequent periods, Y_a, is equal to the amount of goods consumed by it, C_a. We assume that the exchange conserves the total amount of goods:

$$\sum_{a \in \mathcal{A}} y_a = \sum_{a \in \mathcal{A}} c_a. \tag{1}$$

As a result of the exchange of capital used for production, Y_a can be greater or less than y_a, and that $\sum_{a \in \mathcal{A}} y_a = \sum_{a \in \mathcal{A}} Y_a$ does not necessarily hold.

Let us assume that the utility of an agent is a function of the amount of goods it consumes. As is typical in the literature of economics and sociophysics [11, 9], we assume that an agent's utility is an increasing concave function of the consumption of the goods. This is a natural assumption, because increasing the amount of goods improves our lives, but the degree of improvement is diminished if we already have a large amount of goods. In this case, the utility of agent a is given by

$$u_a(c_a) + \beta_a u_a(C_a), \tag{2}$$

where $u_a(\cdot)$ is a strictly increasing concave function, and $\beta_a > 0$ is the parameter that determines the relative preference of future consumption over present consumption. This can be regarded as an additive, separable, intertemporal utility function, $\sum_{t=1}^{\infty} \gamma_a^{t-1} u_a(c_{a,t})$ [11], where

$$c_{a,t} = \begin{cases} c_a & \text{if } t = 1 \\ C_a & \text{if } t > 1 \end{cases} \tag{3}$$

is agent a's consumption in period t and $\gamma_a = \beta_a/(1+\beta_a)$.

If agent a does not engage in the exchange of goods and capital, its utility is

$$u_a(\tilde{c}_a) + \beta_a u_a(\tilde{c}_a) = (1+\beta_a) u_a(\tilde{c}_a), \tag{4}$$

where $\tilde{c}_a = y_a$ is the consumption and production in the periods before the exchange takes place, because it consumes exactly the same amount of goods as that produced by itself. Here, we assume that, without exchange, the amount of produced goods is kept constant because the amount of capital owned by the agent is constant. Let us assume that agents can refuse an exchange decreasing their own utility. Then,

$$(1+\beta_a)\tilde{u}_a \le u_a + \beta_a U_a \tag{5}$$

holds, where we abbreviate $u_a(\tilde{c}_a)$, $u_a(c_a)$, and $u_a(C_a)$ to $\tilde{u}_a$, u_a, and U_a, respectively. In other words, the predefined threshold that a boundedly rational agent in this model wants to exceed is the utility that the agent would achieve if it did not engage in the exchange.

Now, I show the inequality

$$\sum_{a\in\mathcal{A}} \beta_a y_a \le \sum_{a\in\mathcal{A}} \beta_a Y_a. \tag{6}$$

Remarkably, this result holds for agents with any strictly increasing concave utility functions and any production functions.

From Eq. (5), agent a' with $u_{a'} = \tilde{u}_{a'}$, $u_{a'} = U_{a'}$, or, $\tilde{u}_{a'} = U_{a'}$ satisfies $y_{a'} \le Y_{a'}$, and, consequently, $\mathcal{A} = \{a'\}$ satisfies Eq. (6). Hence, it is sufficient for us to support Eq. (6) in excluding agents that satisfy one of these conditions from $\mathcal{A}$ and replacing Eq. (1) with

$$\sum_{a\in\mathcal{A}} \tilde{c}_a \ge \sum_{a\in\mathcal{A}} c_a \tag{7}$$

because $\tilde{c}_{a'} \le c_{a'}$ holds for each excluded agent, a'. Thus, for $a \in \mathcal{A}$, $\tilde{u}_a$, u_a, and U_a are distinct values, and $\tilde{c}_a$, c_a, and C_a are distinct values.

Because $u_a(\cdot)$ is a strictly increasing concave function, we note that

$$\frac{\tilde{u}_a - U_a}{\tilde{c}_a - C_a} > 0, \tag{8}$$

$$\frac{u_a - \tilde{u}_a}{c_a - \tilde{c}_a} > 0, \tag{9}$$

$$\frac{U_a - u_a}{C_a - c_a} > 0, \tag{10}$$

$$\begin{cases} \dfrac{u_a - \tilde{u}_a}{c_a - \tilde{c}_a} \geq \dfrac{\tilde{u}_a - U_a}{\tilde{c}_a - C_a} & \text{if } c_a < C_a \\ \dfrac{u_a - \tilde{u}_a}{c_a - \tilde{c}_a} \leq \dfrac{\tilde{u}_a - U_a}{\tilde{c}_a - C_a} & \text{if } c_a > C_a \end{cases}. \tag{11}$$

We define

$$d_{1a} = \frac{1}{\tilde{u}_a - U_a}, \tag{12}$$

$$d_{2a} = \frac{1}{u_a - \tilde{u}_a}, \tag{13}$$

$$d_{3a} = \frac{1}{U_a - u_a}, \tag{14}$$

$$z_{1a} = \tilde{c}_a - C_a, \tag{15}$$

$$z_{2a} = c_a - \tilde{c}_a. \tag{16}$$

In addition, we define $2A \times (4A+1)$-dimensional matrix $\boldsymbol{M}$ and $2A$-dimensional vectors $\boldsymbol{b}$ and $\boldsymbol{\zeta}$ as

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{D}_1 & \boldsymbol{O} & \boldsymbol{D}_3 & -\boldsymbol{D}_1\boldsymbol{D}_3 \\ \boldsymbol{1} & \boldsymbol{O} & -\boldsymbol{D}_2 & \boldsymbol{D}_3 & \boldsymbol{D}_2\boldsymbol{D}_3 \end{pmatrix}, \tag{17}$$

$$\boldsymbol{b} = [\beta_1, \ \ldots, \ \beta_A, \ 0, \ \ldots, \ 0], \tag{18}$$

$$\boldsymbol{\zeta} = [z_{11}, \ \ldots, \ z_{1A}, \ z_{21}, \ \ldots, \ z_{2A}]. \tag{19}$$

Here, $\boldsymbol{0}$ and $\boldsymbol{1}$ denote the A-dimensional vectors whose elements are all 0 and 1, respectively, $\boldsymbol{O}$ denotes the $A \times A$ null matrix, and

$$\boldsymbol{D}_i = \begin{pmatrix} d_{i1} & 0 & \cdots & 0 \\ 0 & d_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{iA} \end{pmatrix}. \tag{20}$$

Equations (7)–(11) lead to

$$\boldsymbol{M}^T \boldsymbol{\zeta} \leq \boldsymbol{0}. \tag{21}$$

As we show in the following, there is a $4A+1$-dimensional nonnegative vector, $\boldsymbol{\chi}$, satisfying $\boldsymbol{M}\boldsymbol{\chi} = \boldsymbol{b}$. It follows from Farkas' lemma that there is no $\boldsymbol{\zeta}$ satisfying $\boldsymbol{M}^T\boldsymbol{\zeta} \leq \boldsymbol{0}$ and $\boldsymbol{b}^T\boldsymbol{\zeta} > 0$. Hence, we have $\boldsymbol{b}^T\boldsymbol{\zeta} \leq 0$, which is identical to Eq. (6).

The task is now to find

$$\boldsymbol{\chi} = [x_0, \ \boldsymbol{x}_1, \ \boldsymbol{x}_2, \ \boldsymbol{x}_3, \ \boldsymbol{x}_4] \tag{22}$$

satisfying $\boldsymbol{M}\boldsymbol{\chi} = \boldsymbol{b}$, $x_0 \geq 0$, and $\boldsymbol{x}_i \geq \boldsymbol{0}$ $(1 \leq i \leq 4)$. Because

$$\frac{d_{3a}}{d_{1a}}x_{3a} - d_{3a}x_{4a} - \frac{\beta_a}{d_{1a}} = x_{1a}, \tag{23}$$
$$\frac{1}{d_{2a}}x_0 + \frac{d_{3a}}{d_{2a}}x_{3a} + d_{3a}x_{4a} = x_{2a} \tag{24}$$

hold for $1 \leq a \leq A$, the inequalities $x_{1a} \geq 0$ and $x_{2a} \geq 0$ are satisfied by

$$x_0 = 1, \tag{25}$$
$$x_{3a} = 0, \tag{26}$$
$$x_{4a} = \begin{cases} -\beta/(d_{3a}d_{1a}) & \text{if } d_{1a} > 0 \text{ or } (d_{1a} < 0 \text{ and } d_{2a} < 0) \\ 0 & \text{if } d_{1a} < 0 \text{ and } d_{2a} > 0 \end{cases}. \tag{27}$$

Note that $d_{1a} > 0$ implies $d_{2a} > 0$ and $d_{3a} < 0$ and that $d_{1a} < 0$ and $d_{2a} < 0$ imply $d_{3a} > 0$ because

$$-\frac{1}{d_{2a}} + \beta_a \frac{1}{d_{1a}} \leq 0, \tag{28}$$

which follows from Eq. (5), and

$$\frac{1}{d_{1a}} + \frac{1}{d_{2a}} + \frac{1}{d_{3a}} = 0, \tag{29}$$

which is a direct result of Eqs. (12)–(14). This completes the proof.

Let us state a few consequences of this result. First, most change is irreversible because

$$S = \sum_{a \in \mathcal{A}} \beta_a y_a \tag{30}$$

increases with an exchange, unless the equality holds in Eq. (6). Although the amount of production of agent a, y_a, can increase and decrease, S, calculated from the amount of production of all agents, never decreases. Second, selling capital to agents increases S. To support this, let us consider another agent, α, whose production of the goods is an increasing function of capital i. Let us assume that boundedly rational agent α sells capital i to agents in $\mathcal{A}$ and buys the goods. Because the goods produced by agent α decrease, S must increase in order for

$$S' = \beta_\alpha y_\alpha + S \tag{31}$$

to increase.

3. Conclusion

In this paper, I have shown that change is irreversible if agents can avoid decreasing their utility. The model does not assume perfectly rational agents, perfect information, agent budget constraints, a competitive market, or Pareto-efficient allocations [1]. Additionally, we do not need any assumption on the dependency of the production of goods on capital. The main result was derived without using the assumption that the production functions were increasing functions. It is noteworthy that these weak assumptions are sufficient to derive the irreversibility of change. Although we assumed that the exchange takes place only once, the present model can be applied to multiple exchanges in multiple periods if we assume that agents do not know what happens in the subsequent periods and that they expect no exchange during the subsequent periods. The present model can be a widely applicable toy model of goods-exchanging agents, including perfectly rational agents.

In this model, irreversibility originates from the loss-avoiding behavior of agents [12, 6]. Although irreversibility has not been emphasized in economics [5], it plays a vital role in physics, particularly in thermodynamics [13, 14]. This model suggests that there may be a novel correspondence between physical and economic systems. The present results suggest that S may be a quantity corresponding to entropy. This correspondence might enable us to formulate economic and social models from the viewpoint of irreversibility.

Extending the present model could provide great insight into social and economic dynamics. For example, a study could investigate whether agents trading goods whose amounts are represented by two or more scalar values exhibit irreversibility. Although the present study showed that any changes decreasing S are impossible, whether all changes increasing S are possible or not remains unanswered. For thermodynamic entropy S, the answer is yes. Assuming that agents are at the competitive equilibrium [1] would allow us to define an adiabatic process and to answer this question.

Acknowledgments

I thank Takashi Hayakawa for his comments on the draft of this manuscript.

References

[1] Hal R Varian. *Intermediate microeconomics with calculus: a modern approach.* WW Norton & Company, 2014.

[2] Refet S Gürkaynak and Cédric Tille. *DSGE models in the conduct of policy: Use as intended.* CEPR Press, 2017.

[3] Daniel Kahneman. Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93(5):1449–1475, 2003.

[4] Herbert Alexander Simon and Chester Irving Barnard. *Administrative behavior: a study of decision-making processes in administrative organization.* Macmillan, 2nd edition, 1957.

[5] Paul A. Samuelson. Rigorous observational positivism: Klein's envelope aggregation; thermodynamics and economic isomorphisms. In Lawrence Robert Klein, F. Gerard Adams, and Bert G. Hickman, editors, *Global Econometrics: Essays in Honor of Lawrence R. Klein*, pages 1–38. The MIT Press, Cambridge, Massachusetts, 1983.

[6] Katalin Martinás. About irreversibility in economics. *Open Systems & Information Dynamics*, 7(4):349–364, 2000.

[7] Anirban Chakraborti and Marco Patriarca. Variational principle for the pareto power law. *Physical Review Letters*, 103(22):228701, 2009.

[8] Marco Patriarca and Anirban Chakraborti. Kinetic exchange models: From molecular physics to social science. *American Journal of Physics*, 81(8):618–623, 2013.

[9] Giuseppe Toscani, Carlo Brugna, and Stefano Demichelis. Kinetic models for the trading of goods. *Journal of Statistical Physics*, 151(3-4):549–566, 2013.

[10] Pierre Degond, Jian-Guo Liu, and Christian Ringhofer. Evolution of the distribution of wealth in an economic environment driven by local nash equilibria. *Journal of Statistical Physics*, 154(3):751–780, 2014.

[11] Fabio C Bagliano, Giuseppe Bertola, et al. *Models for Dynamic Macroeconomics.* Oxford University Press on Demand, 2004.

[12] Carl Menger. *Grundsätze der Volkswirtschaftslehre (Principles of Economics).* W. Braumüller, 1871.

[13] Elliott H Lieb and Jakob Yngvason. A guide to entropy and the second law of thermodynamics. In *Statistical Mechanics*, pages 353–363. Springer, 1998.

[14] Hal Tasaki and Glenn Paquette. *Thermodynamics–A Modern Point of View.* 2019. (unpublished).