J Econ Interac Coord (2008) 3:107–118
DOI 10.1007/s11403-008-0038-2

REGULAR ARTICLE

Dynamics of a market with heterogeneous learning agents

Tatsuo Yanagita · Tamotsu Onozaki

Published online: 21 March 2008

Abstract We study a simple market model of heterogeneous interacting agents. There are two types of agents, consumers and firms, who are boundedly rational in the sense that they only have partial information. The consumers have a given amount of income at each time period. They determine from which firm to purchase goods, and then spend all the income to purchase as much as they can. The amount of goods they can obtain depends not only on the price offered by the firm but also the amount produced by the firm and the number of customers of the firm. We employ a statistical description of consumers' behavior. The Bolzmann distribution is used to represent firms' share distribution of consumers, which is characterized by a parameter (called "temperature" in physical systems) describing how greedily the consumers pursue higher utility. The firm does not know the shape of demand function it faces, so it revises production and price so as to raise its profit with the aid of a simple reinforcement learning rule which is applied to the "one-armed bandit" problem. Numerical simulations show that there is an optimal greediness, which maximize the time average of consumers' utility. In the vicinity of the optimal greediness, oligopoly emerges although its membership changes frequently. In an oligopolistic market, the market share distribution of firms follows Zipf's law.

Keywords Multi-agent system · Reinforcement learning · Bounded rationality · Numerical simulation · Oligopoly

T. Yanagita (✉)
Research Institute for Electronic Science, Hokkaido University, Sapporo 060–0812, Japan
e-mail: yanagita@nsc.es.hokudai.ac.jp

T. Onozaki
Faculty of Management and Economics, Aomori Public College, Aomori 030–0196, Japan
e-mail: onozaki@bb.nebuta.ac.jp

JEL Classification D43 · L11 · C15 · C63

1 Introduction

A decentralized market economy is a complex adaptive system which consists of a large number of adaptive agents involved in parallel local interactions. These micro-level local interactions give rise to a certain macro-level spontaneous order, and then, the macro-order plays the role of binding conditions for micro-behavior. Complex dynamical behavior emerges as a consequence of recurrent causal chains among individual behavior and the macro-order. This complex two-way feedback between microstructure and macrostructure has been recognized for a very long time, at least since the time of Adam Smith. Nevertheless, until recently, not only economics but other branches of science have lacked the means to model this feedback structure qualitatively in its full dynamical complexity.

Researchers are now able to model, with the aid of high-performance computers, a wide variety of complex phenomena in a decentralized market economy, such as business cycles, endogenous trade network formation, market share fluctuations, and the open-ended coevolution of individual behavior and economic institutions. One branch of this research direction has come to be known as agent-based computational economics, i.e., computer-aided studies of economy modeled as an evolving system of autonomous interacting agents (see, e.g., Kirman and Zimmermann 2001; Namatame 2006).

In the present paper, we investigate the dynamics of a competitive market consisting of locally interacting, boundedly rational firms and consumers. The main focus is on the market share dynamics in common with Onozaki and Yanagita (2003), where product differentiation exists and consumers' brand loyalty plays an important role for emerging oligopoly. However, in the present paper, it is assumed that a consumer decides from which firm to purchase goods so as to increase his utility, and we employ, as a first step, a statistical physics approach since the number of consumers is large. The aggregate consumer behavior is described by the Bolzmann distribution which is characterized by the "inverse temperature" indicating how greedily the consumer seeks to increase his utility. A firm, on the other hand, revises production decision and the price so as to raise its profit with the aid of reinforcement learning algorithm, i.e., by learning through its experience. We mainly focus on the dynamical phases which emerge as consumers' greediness changes.

2 Model

Our model consists of many consumers and many firms that are boundedly rational in the sense that they attempt to increase their utility and profit subject to informational restrictions. The market to be considered is competitive because there exist many consumers and many firms, but the consumers are boundedly rational, hence there are opportunities for a firm to set a price different from those of other firms. In this sense, the firms have the ability to control their prices, and the market is considered to be monopolistically competitive even if there is no product differentiation. Decisions for purchase and production occur at discrete time periods. Goods are homogeneous and perishable within a unit time period.

2.1 Consumer's behavior

We assume the consumers to be identical in the following sense. Each consumer has the same amount of money income at each time step, selects a firm, and is willing to spend all the money to purchase goods from the selected firm.[1] Each consumer's utility $u_i(t)$ at period t is represented by the identical function

$$u_i(t) = U(x_i(t)),$$

where U is a monotonically increasing function of the amount of goods $x_i(t)$ that a consumer obtains from firm i at period t. U is specified as $U(x) = x^{\alpha}, 0 < \alpha < 1$ with $U' > 0$ and $U'' < 0$. The consumer is willing to expense all the money to purchase as much goods as possible because his utility increases as the amount of goods increases. The amount of goods $x_i(t)$ he can obtain from firm i, however, depends not only on prices set by the firm i but also its output level and the number of customers who select firm i.

Each consumer selects a firm so as to raise his utility as high as possible. He compares the utility to be obtained by purchasing goods from firm i or from a randomly selected firm j, and selects one according to a transition probability, $\rho(i, j) = \min(1, (u_j/u_i)^{\beta_1})$, from firm i to firm j, where β_1 is a positive parameter that represents how greedily the consumer behaves. This rule is interpreted as follows. If $(u_j/u_i)^{\beta_1} \geq 1$ (this implies $u_j \geq u_i$), the consumer purchases goods from firm j at the next period. Furthermore, even if $(u_j/u_i)^{\beta_1} < 1$, the consumer curiously chooses firm j with a probability $(u_j/u_i)^{\beta_1}$. This rule is the so-called "softmax action selection" in the field of reinforcement learning. It implies that a firm from which consumers can purchase more goods has a higher probability of capturing consumers. The reason why probability is taken into consideration is that exploring other firms affords consumers the chance to encounter higher utility. Indeed, the softmax rule is used to depict the exploratory decision of human beings (Daw 2006).

From a statistical point of view, when a consumer moves from firm i to firm j with a transition probability $\rho(i, j)$, the stationary share distribution of consumers can be written as

$$w_i^*(t) = u_i^{\beta_1} \Bigg/ \sum_{j=1}^{M} u_j^{\beta_1}, \tag{1}$$

where the denominator is a normalized constant (Binder and Heermann 1979). Note that we can rewrite this share distribution as the Bolzmann distribution: $w_i^*(t) = \exp(-\beta_1 E_i)/Z$, where $E_i = -\log(u_i)$ and $Z = \sum_i \exp(-\beta_1 E_i)$. As stated above,

[1] For the sake of keeping the model simple, here we assume that money income at a certain period can not be carried over to the next period so as to eliminate the intertemporal allocation problem.

U is specified as $U(x) = x^{\alpha}$, $0 < \alpha < 1$. In our setting, the above market share distribution can be obtained for any value of α as long as β_1 is rescaled. Thus we fix $\alpha = 0.5$ without loss of generality. We note that the parameter β_1 corresponds to the inverse temperature in statistical mechanics (Binder and Heermann 1979) and expresses how greedily the consumer behaves. When $\beta_1 \to 0$, i.e., the temperature goes to infinity, the consumers behave in a purely random manner irrespective of their utility, whereas all consumers select the same firm that maximizes their utility when $\beta_1 \to \infty$, i.e., the temperature goes to zero.

In our model, consumers' utility varies with time through a change in price and quantity. To take into account this effect, we introduce, for simplicity, a linear relaxation dynamics toward the stationary distribution Eq. (1) as follows:

$$w_i(t+1) = w_i(t) - \left(w_i(t) - w_i^*(t)\right)/\tau,$$

where τ is a parameter that determines the relaxation time scale.

2.2 Firm's behavior

Owing to bounded rationality, a firm does not know the demand function it faces nor the prices other firms have set, so that it must decide its price p_i and production q_i based only on the restricted local information, i.e., changes in profits. Profit $\Pi_i(t)$ of firm i at period t is defined as

$$\Pi_i(t) = p_i(t)s_i(t) - c_i(t),$$

where $s_i(t)$ denotes the quantity sold by firm i and $c_i(t) = (q_i(t))^2$ is an identical cost function. The sale $s_i(t)$ is represented as

$$s_i(t) = \min\left(q_i(t), w_i(t)T/p_i(t)\right), \tag{2}$$

where T is the total money income of all consumers. Eq. (2) comes from the fact that total demand for firm i's products is given by $w_i T/p_i$, and, if the demand differs from the production q_i, the sale is determined by the short-side. Thus, the profit of a firm depends upon its decision on price and production, and the demand it faces. Considering Eq. (2), the amount of goods $x_i(t)$ that a consumer can obtain from firm i at period t is written as

$$x_i(t) = s_i(t)/(w_i(t)L) = \min\left(q_i(t)/(w_i(t)L), T/(p_i(t)L)\right),$$

where L is the number of consumers.

We assume that a firm does not directly control the price and production, but, instead, it determines rates of change of the previous price and production. Firm i chooses a pair of rates of change $(\delta p_i, \delta q_i)$ among all possible options so as to get higher profits. Note that $p_i(t+1) = \delta p_i \cdot p_i(t)$ and similarly for q_i. Here the rates of change are given by

$$\begin{cases} \delta p_i = 1 + \Delta p \cos(2\pi n_i/N) \\ \delta q_i = 1 + \Delta q \sin(2\pi n_i/N) \end{cases} \quad \text{for } n_i \in \{0, \dots, N-1\},$$

where n_i is an integer number from 0 to $N-1$ denoting a strategy of firm i, N is the number of possible strategies, and Δp and Δq are given constants. Therefore, the maximum rates of change in price and production are $1 \pm \Delta p$ and $1 \pm \Delta q$.

A firm selects a strategy $n \in \{0, \dots, N-1\}$ in pursuit of higher profit, according to a simple reinforcement learning rule which is applied to the "one-armed bandit" problem (Sutton and Barto 1998) as follows (here the subscript i of strategy n_i is omitted because there is no possibility of confusion). First, a firm evaluates all of its actions $\{0, \dots, N-1\}$ by calculating the normalized quantity

$$\tilde{\Pi}_i^n(t) = \left(\hat{\Pi}_i^n(t) - \min_n(\hat{\Pi}_i^n(t))\right) \Big/ \left(\max_n(\hat{\Pi}_i^n(t)) - \min_n(\hat{\Pi}_i^n(t))\right),$$

where $\hat{\Pi}_i^n(t)$ is firm i's expectation of its profit at period t when it selects a strategy n. Then, firm i selects a strategy n, following the "softmax" algorithm, i.e., with a probability proportional to $\exp(-\beta_2 \tilde{\Pi}_i^n(t))$, where β_2 is the inverse temperature determining how greedily the firm behaves. Finally, firm i adaptively revises its profit expectation according to

$$\hat{\Pi}_i^n(t+1) = (1-k)\hat{\Pi}_i^n(t) + k\Pi_i(t), \tag{3}$$

where k is the learning rate.

3 Simulations

3.1 Premises

The absolute level of profit is qualitatively irrelevant to our analysis since the competition among firms drives the dynamics, i.e., only the relative volume of profit matters. Thus we set the total money income of all consumers T to one, and the maximum profit of each firm is rescaled to be one. Similarly, the population L of consumers is qualitatively irrelevant and is set to one.

For the initial condition of profit expectation $\hat{\Pi}_i^n(0)$, we choose an "optimistic" value so that firms revise all their expectations effectively (Sutton and Barto 1998). We set $\hat{\Pi}_i^n(0) = 100\ \forall i, n$, which is large enough for realizing the maximum profit. The initial prices and productions are $(p_i(0), q_i(0)) = (1+\xi, 1+\xi)$, where ξ is a small random number distributed uniformly in $[-0.01, 0.01]$. The initial market share distribution is the same among firms, i.e., $w_i(0) = 1/M\ \forall i$.

We fix the following parameters throughout simulations: $\alpha = 1.0, k = 0.5, \tau = 0.1$, $\beta_2 = 3.0$, $N = 10$, $\Delta p = \Delta q = 0.01$. We mainly consider the dependence of consumer's behavior on the inverse temperature β_1, that is, on how greedily the consumer behaves.

As for time scale, we use a non-dimensional time t/t^*, where $t^* = N/k$ is a learning time scale estimated from Eq. (3). This time scale comes from the following consideration. Let us regard Eq. (3) as a differential equation

$$\frac{d\hat{\Pi}_i(t)}{dt} = k\left(\Pi_i(t) - \hat{\Pi}_i(t)\right), \tag{4}$$

where the superscript n is omitted for notational simplicity. If $\Pi_i(t)$ is fixed in time as Π_i^*, $\hat{\Pi}_i(t)$ exponentially converges to this stationary value Π_i^*. Eq. (4) can be solved as

$$\hat{\Pi}_i(t) - \Pi_i^* = Ae^{-kt},$$

where A is a constant determined by the initial condition $\hat{\Pi}_i(0)$. Starting from $t = 0$, the difference $\hat{\Pi}_i(t) - \Pi_i^*$ becomes $1/e$ times its initial value A when $t = 1/k$. Thus we consider $1/k$ as a typical unit time scale which characterizes the firm's learning process. Each firm must learn N strategies, thus $t^* = N/k$ is the total time required for a single learning process by firm.

3.2 Single-firm case

In order to verify whether a firm agent can optimize its profit, we consider the behavior of a single-firm system by setting $M = 1$. This assumption leads to a situation where monopoly is always realized artificially. Note that since there is only one firm, consumers' behavior is unique, i.e., each consumer has to purchase goods from the monopolist. Hence, there is no competition, and the best strategy of the monopolist is to raise the price and to reduce the production so as to increase the profit. As a result, the utility of consumer gradually decreases and approaches zero. Indeed, as shown in Fig. 1a, the price and the production gradually increase and decrease respectively in the $M = 1$ case. The corresponding decision histories for three different initial conditions in (p, q)-space are depicted in Fig. 1b. The firm, even if starting from any randomly selected initial conditions, updates its decision through learning process so as to increase its profit. The monopolistic firm seems to behave as a boundedly rational agent seeking higher profit.

3.3 Multi-firm cases

In a single-firm case, consumers' utility gradually decreases with time in an "artificial" monopoly market. In a multi-firm system, however, simultaneously raising the price and reducing production is not always the best strategy. Suppose that the market share of one firm is sufficiently larger than that of the others, the market is virtually monopolistic and the dominant firm tends to adopt the optimal strategy for monopolist, namely raising the price and reducing production. When the dominant firm raises the price, its customers' utility decreases. Decrease in utility causes decrease in the number of customers through Eq. (1). As a result, sooner or later, the firm realizes that its strategy is no longer the best through a reinforcement learning process, i.e., by encountering

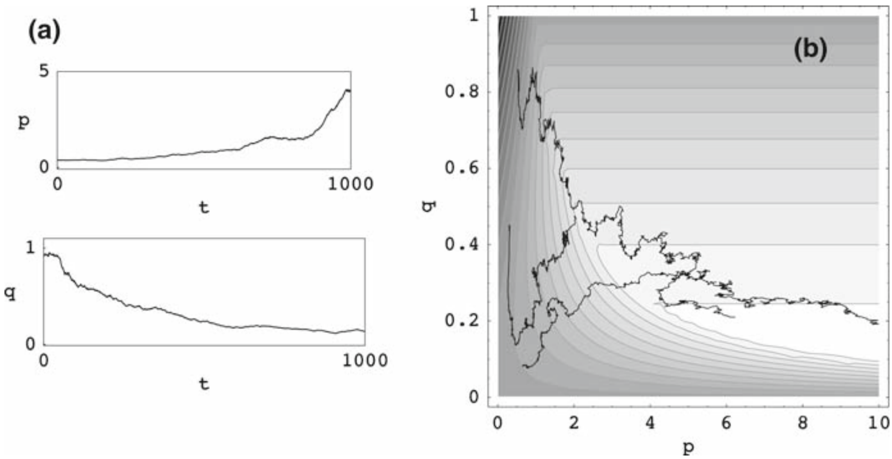

Fig. 1 **a** Time evolutions of price and production are shown in an "artificial" monopoly case. The monopolist raises price and reduces production to get higher profit. Therefore, consumers' utility approaches zero. **b** Typical trajectories of firm's decisions for three different initial conditions are plotted in an iso-profit contour diagram projected on the (p, q)-plane. $t \in [0, 1000]$ and $M = 1$

the fact that the profit is actually decreasing. In other words, the monopolist's strategy of raising the price and reducing production is restrained through the feedback from consumers' behavior. With regard to production, the dominant firm slows down the pace of reducing production; otherwise its profit decreases rapidly. Typical time series of these values for the case of two firms where $\beta_1 = 4.0$ are shown in Fig. 2, and it is easy to confirm the above explanation from observing the dotted area. The fluctuation range of prices is very wide as compared to other variables, but its absolute scale does not matter because it is the difference between consumers' utility that drives the dynamics.

Dynamics of the market crucially depends on the parameter β_1, which represents how greedily the consumers seek higher utility. For lower greediness (smaller β_1), almost all consumers choose firms in a purely random manner, i.e., the choices are irrelevant to consumers' utility. Thus, the market share distribution is almost uniform among firms and stationary as depicted in Fig. 3a. For higher greediness, almost all consumers choose the "best" firm so as to seek higher utility. Therefore, monopoly emerges as a "quasi-stationary" state, which means that strong monopolists drastically change places with time (Fig. 3c). While "quasi-stationary monopoly" is sustained, the time evolutions of the price and production are similar to those of the single-firm system. In Fig. 3b, the time evolution of the market share distribution for intermediate value of β_1 is shown. In this case, oligopoly persists, while severe market-share battles are observed.

In order to characterize a system of larger size, let us utilize the Herfindahl index

$$H(t) = \sum_{i=1}^{M} (w_i(t))^2 \, ,$$

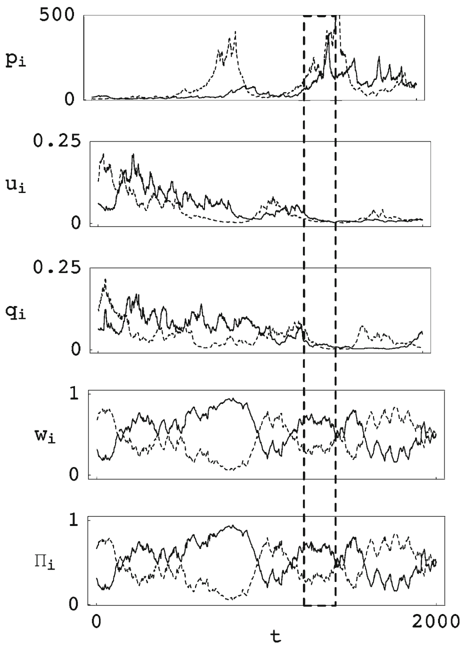

Fig. 2 Typical time series of price p_i, utility u_i, production q_i, share w_i and profit Π_i for the two-firm system. $\beta_1 = 4.0$ and $M = 2$

which measures the non-uniformity of market share distribution w_i. This index was first introduced to indicate the concentration ratio of industries (Herfindahl 1950; Hirschman 1964). As $\sum_{i=1}^{M} w_i(t) = 1.0$, the arithmetic mean $\bar{w}(t)$ of market share distribution is equal to $1/M$. Then, the variance $V(w_i(t))$ is calculated as

$$V(w_i(t)) = (1/M)\sum_{i=1}^{M}(w_i(t) - \bar{w}(t))^2 = H(t)/M - 1/M^2.$$

Solving the above equation with respect to $H(t)$ yields

$$H(t) = 1/M + MV(w_i(t)),$$

which implies that a higher variance due to a higher level of non-uniformity of firms' market shares results in a lower value of $H(t)$. The variance $V(w_i(t))$ takes the lowest value of 0.0 when consumers are equally distributed to each firm, i.e., $w_i(t) = 1/M$ $\forall i$. Then the lowest value of $H(t)$ is equal to $1/M$. When there is a firm, say i, that satisfies $w_i(t) = 1.0$ and $w_j(t) = 0.0$ for $j \neq i$, the variance $V(w_i(t))$ takes the highest value of $1/M(1 - 1/M)$. Then the highest value of $H(t)$ is equal to 1.0. Thus, The index $H(t)$ ranges from $1/M$ to 1.0. The time series of $H(t)$ for $M = 50$ is depicted in Fig. 3d–f. For smaller β_1, $H(t)$ takes the values close to 1/50 all the time as shown in Fig. 3d, meaning that the equi-share oligopoly is realized because each consumer selects a firm without considering his utility. Thus, the dynamics in this case is the same as that in the artificial monopoly case. Firms may set any prices and produce any

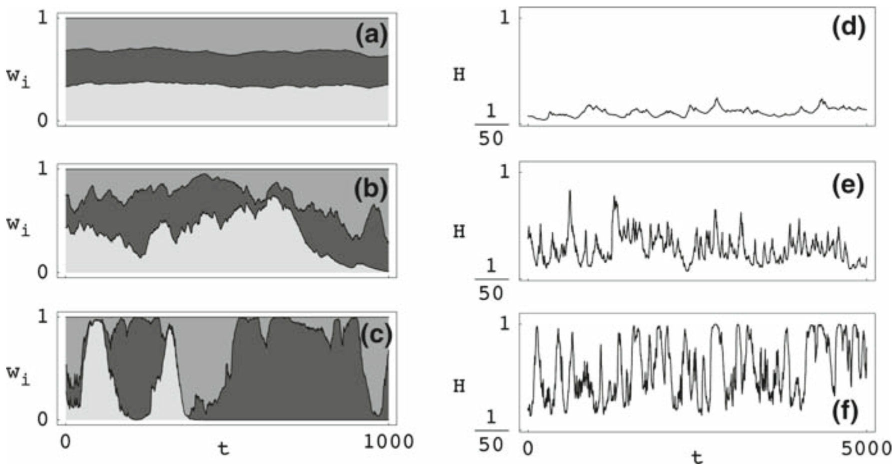

Fig. 3 **a–c**: Time series of market share w_i in the case of three competitive firms are shown for different values of β_1. $M = 3$. **a** $\beta_1 = 0.05$: stationary equi-share oligopoly—market shares are uniform among firms and stationary. **b** $\beta_1 = 0.5$: Market-share battle in oligopoly—market shares change dynamically. **c** $\beta_1 = 2.5$: changing monopoly—monopolistic firms drastically change their role. **d**–**f** Time series of Herfindahl index $H(t)$ in the case of 50 firms are shown for different values of β_1, $M = 50$. **d** $\beta_1 = 0.05$: since market shares are uniform among firms, $H(t)$ approximately takes the minimal value $1/50$ almost always. **e** $\beta_1 = 0.5$: oligopolistic states are formed and $H(t)$ takes some intermediate values between $1/50$ and 1.0. Oligopoly comes from competition among firms through the feedback of consumers' behavior. **f** $\beta_1 = 2.5$: different monopolistic firms appear and $H(t) = 1.0$ frequently

amount because those do not significantly affect customers' behavior. In consequence, price gradually increases and production decreases. For large β_1, consumers are very greedy and over-concentration may occur in one firm that seems to give highest utility. Thus a monopolistic firm change places and $H(t) = 1.0$ frequently as depicted in Fig. 3f. The duration of monopoly gradually increases as β_1 decreases. For intermediate value of β_1, oligopoly comes from the competition among firms through the feedback of consumers' behavior. Therefore, $H(t)$ tends to take some intermediate value between $1/M$ and 1.0 as shown in Fig. 3e.

Since $H(t)$ fluctuates with time, we calculate the time average of the Herfindahl index $\langle H \rangle_t = (1/T) \sum_{t=1}^{T} H(t)$ in order to see its β_1-dependence. $\langle H \rangle_t$ gradually increases with β_1 as a consequence of the appearance of non-uniform share distribution, and approaches 1.0 corresponding to the emergence of an eternal monopolist (see Fig. 4).

Next, we investigate the β_1-dependence of utility and profit. For this purpose, we calculate the time average of the ensemble mean of consumers' utility, i.e., $\langle E(u) \rangle_t = (1/T) \sum_{t=1}^{T} E(u(t))$ where $E(u(t)) = \sum_{i=1}^{M} w_i(t) u_i(t)$. We also calculate the time average of firms' profit per capita, i.e., $\langle \bar{\Pi} \rangle_t = (1/T) \sum_{t=1}^{T} \bar{\Pi}(t)$ where $\bar{\Pi}(t) = (1/M) \sum_{i=1}^{M} \Pi_i(t)$. In Fig. 5, $\langle E(u) \rangle_t$ and $\langle \bar{\Pi} \rangle_t$ are depicted as functions of β_1. It is clearly seen that there is an optimal greediness $\beta_1 \sim 1.5$, at which the time-averaged consumers' utility is maximized. For smaller β_1, the time-averaged utility is very small because each consumer selects a firm in a purely random manner (note that firm's best decision in this case is raising the price and reducing production).

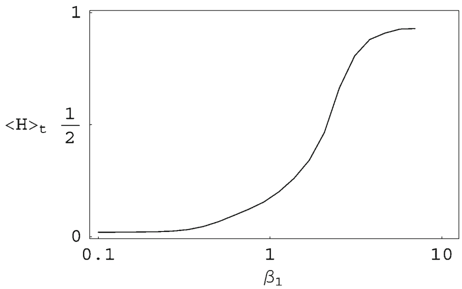

Fig. 4 The time-averaged $H(t)$ versus β_1 for $M = 50$. For smaller β_1, $\langle H \rangle_t \sim 1/M$, i.e., the market share w_i is uniformly distributed among firms. Beyond critical value of $\beta_1^* \sim 0.45$, oligopolistic firms emerge and the transition of states is clearly observed. $\langle H \rangle_t$ gradually increases with β_1 and approaches 1.0

With the increase of β_1, the market-share battle in oligopoly starts to emerge while the time-averaged utility gradually increases and takes an optimal value that gives the maximum utility. Beyond the optimal value, the time-averaged utility gradually decreases because each consumer is too eager to choose the best firm by seeking higher utility, thereby causing the formation of a monopolistic market. On the other hand, the time-averaged profit per capita $\langle \bar{\Pi} \rangle_t$ gradually decreases with increasing β_1 and reaches the minimal value at $\beta_1 \sim 2.4$. Beyond the minimum value, the time-averaged profit per capita gradually increases since monopoly starts to emerge.

Finally, the transition among a uniform-share, an oligopolistic, and a monopolistic market is statistically characterized by means of the probability distribution of market share. For smaller β_1, the market share is uniform and the distribution clearly has a peak at $1/M$ as shown in Fig. 6a. As β_1 increases, the distribution comes to exhibit a long-tail (or fat-tail) and to follow Zipf's law. In fact, as shown in Fig. 6b, c, where oligopoly and monopoly naturally emerge, the distribution seems to obey the power-law (or Pareto-law) and its exponent is approximately 1. It is well known that Zipf's law is ubiquitous and is also observed in firm size distribution (Axtell 2001; Fujiwara 2004). Furthermore, it should be noted that the simulation results above support the claim that the situation where the exponent of the power-law takes approximately 1 is the transition point between the oligopoly phase and the pseudo-equality phase (Aoyama 2004; Aoyama 2007; Bouchaud and Mèzard 2000; Fujiwara 2004).

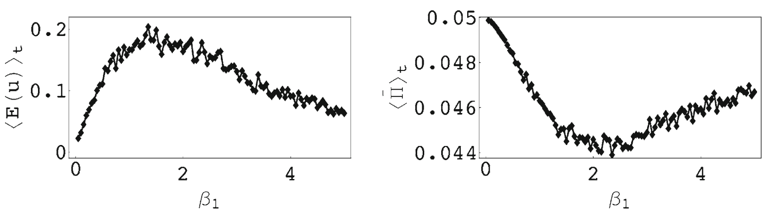

Fig. 5 The time average of the mean of consumers' utility $\langle E(u) \rangle_t$ and the time average of firms' profit per capita $\langle \bar{\Pi} \rangle_t$ versus β_1 for $M = 20$. The utility and profit are averaged over $T = 5000$ time steps and are sampled over 20 different initial conditions after initial transients. There is an optimal greediness $\beta_1 \sim 1.5$ which maximizes the time-averaged utility

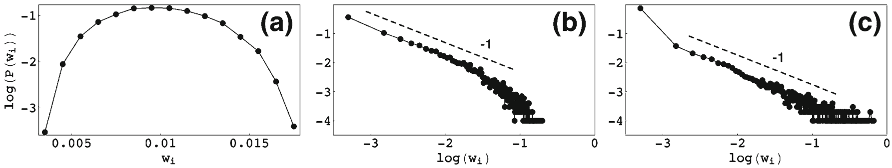

Fig. 6 The probability distributions of market share w_i for $M = 100$. **a** $\beta_1 = 0.1$: For such smaller β_1, the market share is uniform and the distribution has a peak at $1/M$. **b** $\beta_1 = 1.0$: For this larger β_1, the distribution exhibits a long-tail and follows Zipf's law. **c** $\beta_1 = 5.0$: Same as in (**b**)

4 Summary

We have investigated the dynamics of a competitive market consisting of locally interacting, boundedly rational firms and consumers. The behavior of consumers is described by the market share distribution, i.e., the stationary distribution of a large number of consumers who employ softmax strategy. The market share distribution is characterized by a single parameter β_1 that represents how greedily the consumers behave. Firms revise their production decisions and prices so as to raise their profit with the aid of a simple reinforcement learning rule which is applied to the "one-armed bandit" problem.

Numerical simulations show that there is an optimal greediness at which each consumer maximizes his time-averaged utility. For smaller greediness, the consumer chooses a firm in a purely random manner, i.e., the consumer does not consider his utility when selecting a firm. The dynamics in this case is the same as that in a monopolistic market case since there is no competition among firms and a firm's best decision is raising the price and reducing production. As a result, the time-averaged utility decreases for lower greediness. For larger greediness, on the other hand, the consumer is too eager to choose the best firm by seeking higher utility, thereby causing monopoly. Monopoly is realized as a "quasi-stationary" state, in which strong monopolists frequently change places with time, and oligopolistic situations also appear in transition periods. The time evolutions of price and production are the same as those for lower greediness case when a "quasi-monopoly" state is achieved. Thus, the time-averaged utility is also smaller for higher greediness. For intermediate greediness, oligopoly persists while the market share of oligopolists changes dynamically. In this case, the time-averaged utility takes larger values. It is also shown that in an oligopolistic and monopolistic market, the distribution of market share exhibits a long-tail and follows Zipf's law.

Acknowledgments We thank an anonymous referee for constructive suggestions. This research was supported in part by the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C), 19530155, 2007. We are grateful for their support.

References

Aoyama H (2004) Scaling laws of income and firm-size. Presentation at workshop on economics with heterogeneous interacting agents (WEHIA)
Aoyama H et al (2007) Pareto firms. Nippon Keizai Hyoron-sha (in Japanese)

Axtell R (2001) Zipf distribution of US firm sizes. Science 293:1818–1820
Binder K, Heermann DW (1979) Monte Carlo methods in statistical physics. Springer, Berlin
Bouchaud J-P, Mèzard M (2000) Wealth condensation in a simple model of economy. Physica A 282: 536–545
Daw ND et al (2006) Cortical substrates for exploratory decisions in humans. Nature 441:876–579
Fujiwara Y et al (2004) Do Pareto-Zipf and Gibrat laws hold true. Physica A 335:197–216
Herfindahl OC (1950) Concentration in the steel industry. Unpublished PhD Dissertation, Columbia University
Hirschman AO (1964) The paternity of an index. Am Econ Rev 54:761–762
Kirman A, Zimmermann J-B (eds) (2001) Economics with heterogeneous interacting agents. Springer, Berlin
Namatame A (ed) et al (2006) The complex networks of economic interactions: essays in agent-based economics and econophysics. Springer, Berlin
Onozaki T, Yanagita T (2003) Monopoly, oligopoly and the invisible hand. Chaos Solit Fract 18:537–547
Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. MIT Press, Cambridge