Heterogeneity, Price Discovery and Inequality in an Agent-Based Scarf Economy

Computing in Economics and Finance

CEF 2015: Complex Systems Modeling and Simulation in Economics and Finance pp 113-139 | Cite as

- Shu-Heng Chen (1)
- Bin-Tzong Chie (2)
- Ying-Fang Kao (1)
- Wolfgang Magerl (1) (3)
- Ragupathy Venkatachalam (4)

1. AI-ECON Research Center, Department of Economics, National Chengchi University, , Taipei, Taiwan
2. Tamkang University, , Tamsui, Taipei, Taiwan
3. Vienna University of Technology, , Vienna, Austria
4. Institute of Management Studies, Goldsmiths, University of London, , London, UK

Conference paper
First Online: 21 November 2018

Part of the Springer Proceedings in Complexity book series (SPCOM)

Abstract

In this chapter, we develop an agent-based Scarf economy with heterogeneous agents, who have private prices and adaptively learn from their own experiences and those of others through a meta-learning model. We study the factors affecting the efficacy of price discovery and coordination to the Walrasian Equilibrium. We also find that payoff inequality emerges endogenously over time among the agents and this is traced back to intensity of choice (a behavioural parameter) and the associated strategy choices. Agents with high intensities of choice suffer lower payoffs if they do not explore and learn from other agents.

Keywords

Non-tâtonnement processes Coordination Learning Agent-based modeling Walrasian general equilibrium Heterogeneous agents

1 Introduction

How do market economies achieve coordination, even if only imperfectly, among millions of actors performing a multitude of actions without centralization? Research in General Equilibrium theory over many decades has provided many interesting insights. The notion of *coherence* is in this literature interpreted in terms of an *equilibrium* phenomenon. Furthermore, the discovery of equilibrium prices that balance the demands and supplies of various actors in the aggregate is assumed to be mediated through a fictional, centralized authority. This authority, known as the *Walrasian auctioneer* , supposedly achieves this discovery through a process of trial and error (tâtonnement). In this framework, trading happens, if it at all happens, only in equilibrium.

What happens when agents are out of this equilibrium configuration? Whether and *how* these agents manage to coordinate back to equilibrium over time through decentralized exchanges is an important question. The tâtonnement narrative is highly divorced from how agents go about achieving coordination by searching and learning to discover efficient prices in reality. Therefore, it becomes necessary to go beyond the tâtonnement process and find plausible descriptions of behaviour that respect the cognitive and informational limitations that human agents face. Research on non-tâtonnement processes [8, 10, 25] that permit disequilibrium trading have been developed over the years. Agent-based models provide a conducive setup to understand decentralized, disequilibrium dynamics of market economies.

Many aggregate economic outcomes such as growth and fluctuations are often explained as being driven by the expectations of agents. There is often a diversity of expectations among agents and in the face of this heterogeneity, the aggregate outcomes that result from the interaction among agents are complex to understand. But how these expectations come into being and evolve over time has not yet been completely understood. Similarly, their psychological underpinnings and relation to agent-level characteristics have not been sufficiently clarified. The mechanics of expectation formation are intimately linked to how agents learn from their experiences and the environment. Hence, it may not be sufficient to just point to the existence of a diversity of expectations, but it may be necessary to go deeper into understanding their origins and dynamics.

Even in a simple decentralized economy with barter, there is no a priori reason to believe that heterogeneous agents can successfully learn to align their expectations regarding future prices. There is no guarantee that they will eventually discover the equilibrium prices, where all mutually beneficial exchanges will be exhausted. Even if we suppose that they do, how the gains from these exchanges will be distributed among these agents and their relation to the structure of expectations is not obvious.

In this chapter, we investigate these issues through an agent-based, Scarf economy, where the agents are characterized by Leontief payoff functions . Agents attempt to maximize their payoffs, and engage in bilateral trade with others based on their initial endowment and a subjective perception of "fair" prices. Trade failures, either in the form of unsatisfied demand or involuntary inventory, signal misperceptions in the economy. No one person has complete knowledge of the entire system and the economy is best viewed as a complex adaptive process that dynamically evolves. In order to successfully coordinate in such a system, agents may need to learn both from their own experience (individual learning) and from each other (social learning). In our model, agents *consciously* decide when to engage in social and individual. This choice is formulated as meta-learning and we apply reinforcement learning to model this meta-learning. Heterogeneity among agents is in terms of differences in their learning process, more specifically, the intensity of choice parameter of the meta-learning model. Within this set-up, we ask the following questions:

- Can heterogeneous agents successfully discover the (unique) equilibrium prices of the model?
- How do current and accumulated payoffs vary among agents who are identical in all respects except for their intensity of choice?
- If there is a substantial variation in payoffs, what are the factors that drive this inequality?

The rest of this chapter is organized as follows: Sect. 2 develops the agent-based Scarf economy, and Sect. 3 describes the individual, social and meta-learning mechanisms that the agents in the model employ. The simulation design is explained in Sects. 4 and 5 presents the results. Section 6 provides a discussion of the results.

2 The Scarf Economy and the Non-Tâtonnement Process

Herbert Scarf [20] demonstrated an economy in which prices may not converge to the Walrasian general equilibrium under the tâtonnement process, even when the equilibrium is unique. This paper has since led to a lot of research on the issue of the stability of the general equilibrium. Recently, [9] developed an agent-based model and showed that when agents hold private prices and imitate other agents who have been successful, prices converge to the unique general equilibrium of the Scarf-like economy. The Scarf economy has a simple structure that is amenable to investigation and has been widely studied. Previous studies in this area have not tackled the issue of heterogeneity among agents in a Scarf-like set-up. In the remainder of the section, we will develop a heterogeneous, agent-based model of the Scarf economy .

2.1 An Agent-Based Model of the Scarf Economy

Following [20], we consider a pure exchange economy composed of N agents. This economy has three goods, denoted by j = 1, 2, 3, and correspondingly, N agents are grouped into three different 'types', $\tau_j, j = 1, 2, 3$, where $\tau_1 \equiv \{1, ..., N_1\}; \tau_2 \equiv \{N_1 + 1, ..., N_1 + N_2\}; \tau_3 \equiv \{N_1 + N_2, ..., N\}$. Agents who belong to τ_j (type-j) are initially endowed with w_j units of good j, and zero units of the other two goods. Let $\mathbf{W_i}$ be the endowment vector of agent i:

$$\mathbf{W_i} = \begin{cases} (w_1, 0, 0), & i \in \tau_1, \\ (0, w_2, 0), & i \in \tau_2, \\ (0, 0, w_3), & i \in \tau_3. \end{cases}$$

(1)

All agents are assumed to have a Leontief-type payoff function .

$$U_i(x_1, x_2, x_3) = \begin{cases} \min\{\frac{x2}{w2}, \frac{x3}{w3}\}, & i \in \tau_1, \\ \min\{\frac{x1}{w1}, \frac{x3}{w3}\}, & i \in \tau_2, \\ \min\{\frac{x1}{w1}, \frac{x2}{w2}\}, & i \in \tau_3. \end{cases}$$

(2)

Given the complementarity feature of this payoff function, we populate equal numbers of agents in each type. This ensures that the economy is balanced and that no free goods appear.

We assume that agents have their own subjective expectations of the prices of different goods,

$$\mathbf{P_i^e(t)} = (P^e_{i,1}(t), P^e_{i,2}(t), P^e_{i,3}(t)), \;\; i = 1, \dots, N. \tag{3}$$

where $P^e_{i,j}(t)$ is agent *i*'s price expectation of good *j* at time *t*.[1]

Given a vector of the subjective prices (private prices), the optimal demand vector, $\mathbf{X}^* = \Psi^*(\mathbf{W})$ can be derived by maximizing payoffs with respect to the budget constraint.

$$\mathbf{X_i^*} = (x^*_{i,1}, x^*_{i,2}, x^*_{i,3}) = \begin{cases} (0, \psi_i^* w_2, \psi_i^* w_3), & i = 1, \dots, N_1, \\ (\psi_i^* w_1, 0, \psi_i^* w_3), & i = N_1 + 1, \dots, N_1 + N_2, \\ (\psi_i^* w_1, \psi_i^* w_2, 0), & i = N_1 + N_2 + 1, \dots, N. \end{cases} \tag{4}$$

where the multiplier

$$\psi_i^* = \begin{cases} (P^e_{i,1} w_1)/(\sum_{j=2,3} P^e_{i,j} w_j), & i = 1, \dots, N_1 \\ (P^e_{i,2} w_2)/(\sum_{j=1,3} P^e_{i,j} w_j), & i = N_1 + 1, \dots, N_1 + N_2, \\ (P^e_{i,3} w_3)/(\sum_{j=1,2} P^e_{i,j} w_j), & i = N_1 + N_2 + 1, \dots, N. \end{cases} \tag{5}$$

Note that the prices in the budget constraint are 'private' prices. Rather than restricting the prices within a certain neighbourhood (for instance, a unit sphere in [20]), we follow [2] and set one of the prices (P_3) as the numéraire. The Walrasian Equilibrium (WE) for this system is $P_1^* = P_2^* = P_3^* = 1$, when the endowments for each *type* are equal and symmetric.[2] However, in this model agents may have own price expectations that may be very different from this competitive equilibrium price, and may base their consumption decision on their private price expectations $P^e_{i,j}$. To facilitate exchange, we randomly match agents in the model with each other and they are allowed to trade amongst each other if they find it to be beneficial. For this, we need to specify a precise bilateral trading protocol and the procedures concerning how agents dynamically revise their subjective prices.

2.2 Trading Protocol

We randomly match a pair of agents, say, *i* and *i′*. Let *i* be the *proposer*, and *i′* be the *responder*. Agent *i* will initiate the trade and set the price, and agent *i′* can accept or decline the offer. We check for the double coincidence of wants, i.e., whether they belong to the same type. If they do not, they will be rematched. If not, we will then check whether the agents have a positive amount of endowment in order for them to engage in trade. Let m_i be the commodity that *i* is endowed with (m_i = 1, 2, 3). Agent *i*, based on his subjective price expectations, proposes an exchange to agent *i′*.

$$x^*_{i,m_{i'}} = \frac{P^e_{i,m_i} x_{i',m_i}}{P^e_{i,m_{i'}}}, \tag{6}$$

Here, the proposer (*i*) makes an offer to satisfy the need of agent *i′* up to x_{i',m_i} in exchange for his own need $x^*_{i,m_{i'}}$.

Agent *i′* will evaluate the 'fairness' of the proposal using his private, subjective expectations and will consider the proposal to be interesting provided that

$$P^e_{i',m_i} x_{i',m_i} \geq P^e_{i',m_{i'}} x^*_{i,m_{i'}}; \tag{7}$$

otherwise, he will decline the offer. Since the offer is in the form of *take-it-or-leave-it* (no bargaining), this will mark the end of trade.

Agent i' will accept the proposal if the above inequality (7) is satisfied, and if $x_{i',mi} \leq x^*_{i',m_i}$. This *saturation condition* ensures that he has enough goods to trade with *i* and meet his demand. However, only if the saturation condition is not satisfied, will the proposal still be accepted, but the trading volume will be adjusted downward to $x_{i,m_{i'}} < x^*_{i,m_{i'}}$. Agents update their (individual) excess demand and as long as they have goods to sell, they can continue to trade with other agents. The agents are rematched many times to ensure that the opportunities to trade are exhausted. Once the bilateral exchange is completed, the economy enters the consumption stage and the payoff of each agent $U_i(\mathbf{X_i(t)})$, $i = 1, 2, \ldots,$ N, is determined. Note that $\mathbf{X_i(t)}$, the realized amount after the trading process may not be the same as the planned level $\mathbf{X^*_i(t)}$. This may be due to misperceived private prices and a sequence of 'bad luck', such as running out of search time (number of trials) before making a deal, etc. Based on the difference between $\mathbf{X_i(t)}$ and $\mathbf{X^*_i(t)}$, each agent *i* in our model will adaptively revise his or her private prices through a process of learning.

3 Learning

Agents have to *learn* to coordinate, especially in dynamic, disequilibrium environments and non-convergence to the most efficient configuration of the economy can be interpreted as a coordination failure . In this section, we introduce two modes of learning that are frequently used in agent-based computational economics, namely, individual learning and social learning . In the individual learning mode, agents learn from their own past experiences. In social learning , agents learn from the experiences of other agents with whom they interact. These two modes of learning have been characterized in different forms in the literature and have been extensively analysed (see: [3, 14, 21, 22]). The impact of individual and social learning on evolutionary dynamics has been analysed in [18, 26] and more recently in [5]. We describe the learning procedures in more detail below.

3.1 Individual Learning

Under individual learning , an agent typically learns from analysing his own experience concerning the past outcomes and strategies. The mechanism that we employ can be thought of as a modified agent-level version of the Walrasian price adjustment equation. Let the optimal and actual consumption bundles be denoted by the vectors $\mathbf{X^*_i(t)}$ and $\mathbf{X_i(t)}$, respectively. Agent *i* can check his excess supply and excess demand by comparing these two vectors component-wise. Agents then reflect on how well their strategy (private price expectations) has performed in the previous trading round and adjust their strategies based on their own experience of excess demand and supply. We employ a gradient descent approach to characterize individual learning and in a generic form it can be written as:

$$\mathbf{P^e_i(t+1)} = \mathbf{P^e_i(t)} + \underbrace{\Delta\mathbf{P_i(X_i(t), X^*_i(t))}}_{\text{gradient descent}}, \quad i = 1, 2, \ldots, N$$

(8)

We shall detail its specific operation as follows.

Let $\mathbf{m}^y_i$ and $\mathbf{m}^c_i$ denote the production set and consumption set of agent *i*. In the Scarf economy , $\mathbf{m}^y_i \cap \mathbf{m}^c_i = \emptyset$ and in this specific 3-good Scarf economy , $\mathbf{m}^y_i = \{m_i\}$. At the end of each market period, agent *i* will review his expectations for all commodities, $P^e_{i,j}(t), \forall j$. For the good that the agent 'produces',[3] $j \in \mathbf{m}^y_i$), the price expectations $P^e_{i,j}(t)$ will be adjusted downward if m_i is not completely sold out (i.e., when there is excess supply). Nonetheless, even if m_i has been completely sold out, it does not mean that the original price expectation will be sustained. In fact, under these circumstances, there is still a *probability* that $P_{i,j}^e(t)$ may be adjusted *upward*. This is to allow agent *i* to explore or experiment with whether his produced commodity might deserve a better price. However, so as not to make our agents over-sensitive to zero-inventory, we assume that such a tendency for them to be changing their prices declines with the passage of time. That is to say, when agents constantly learn from their experiences, they gradually gain confidence in their price expectations associated with zero-inventory. Specifically, the time-decay function applied in our model is exponential, which means that this kind of exploitation quickly disappears with time. For those goods in the vector $\mathbf{X}_i$ that are a part of the consumption set of the agent, i.e., $j \in \mathbf{m}_i^c$, the mechanism will operate in exactly the opposite manner by increasing the price expectations if there is excess demand. The individual learning protocol is summarized below.

3.1.1 Protocol: Individual Learning

1. 1.

   At the end of the trading day, agent *i* examines the extent to which his planned demand has been satisfied. Let

   $$\Delta x_{i,j}(t) = \begin{cases} x_{i,j}^*(t) - x_{i,j}(t), & \text{if } \{j \in \mathbf{m}_i^c\}, \\ 0 - x_{i,j}(t), & \text{if } \{j \in \mathbf{m}_i^y\} \end{cases}$$

   (9)

2. 2.

   The subjective prices $P_{i,j}^e$ of all three goods will be adjusted depending on $|\Delta x_{i,j}(t)|$.

3. 3.

   If $|\Delta x_{i,j}(t)| > 0$ (i.e., $|\Delta x_{i,j}(t)| \neq 0$),

   $$P_{i,j}^e(t+1) = \begin{cases} (1+\alpha(|\ \Delta x_{i,j}(t)\ |))P_{i,j}^e(t), & \text{if } \{j \in \mathbf{m}_i^c.\} \\ (1-\alpha(|\ \Delta x_{i,j}(t)\ |))P_{i,j}^e(t), & \text{if } \{j \in \mathbf{m}_i^y.\} \end{cases}$$

   (10)

   where $\alpha(.)$ is a *hyperbolic tangent function* , given by:

   $$\alpha(|\ \Delta x_{i,j}(t)\ |) = \tanh(\varphi \mid \Delta x_{i,j}(t)\ |) = \frac{e^{(\varphi|\Delta xi,jt)|)} - e^{(-\varphi|\Delta xi,jt)|)}}{e^{(\varphi|\Delta xi,jt)|)} + e^{(-\varphi|\Delta xi,jt)|)}}$$

   (11)

4. 4.

   If $|\Delta x_{i,j}(t)| = 0$,

   $$P_{i,j}^e(t+1) = \begin{cases} (1-\beta(t))P_{i,j}^e(t), & \text{if } \{j \in \mathbf{m}_i^c.\} \\ (1+\beta(t))P_{i,j}^e(t), & \text{if } \{j \in \mathbf{m}_i^y.\} \end{cases}$$

   (12)

   where β is a random variable, and is a function of time.

   $$\beta = \theta_1 \exp\frac{-t}{\theta_2}, \quad \theta_1 \sim U[0, 0.1],$$

   (13)

   where θ_2 is a time scaling constant.

3.2 Social Learning

Social learning broadly involves observing the actions of others (peers, strangers or even members of different species) and acting upon this observation. This process of acquiring relevant information from other agents can be decisive for effective adaptation and evolutionary survival. Imitation is one of the most commonly invoked, simplest forms of social learning . Players imitate for a variety of reasons and the advantages can be in the form of lower information-gathering costs and information-processing costs, and it may also act as a coordination device in games [1]. Although the idea of imitation is fairly intuitive, there are many different forms in which agents and organisms can exhibit this behaviour ([11], section G, 115–120.).

We adopt a fairly basic version of imitation behaviour, where agents exchange their experiences regarding payoffs with other agents, with whom they are randomly matched. This can be thought of as a conversation that takes place with friends in a café or a pub at the end of a trading day, where they share their experiences regarding their amount of and pleasure associated with consumption as well as their price expectations. An agent with a lower payoff can, based on observing others, replace his own price expectations with those of an agent with a higher payoff. This is assumed to be done in the hope that the other agent's strategy can perform better. If they both have the same payoffs, then the agent chooses between them randomly. If the agent ends up meeting someone who has performed worse than him, he does not imitate and retains his original price expectations.

3.2.1 Protocol: Social Learning

1. 1.

   At the end of each day, each agent consumes the bundle of goods that he has obtained after trading, and derives pleasure from his consumption $U_i(t)$ ($i = 1, \ldots, N$).

2. 2.

   Agents are matched randomly, either with other agents of the same type or with agents who are of different types. This is achieved by randomly picking up a pair of agents (i, i') *without replacement* and they are given a chance to interact.

3. 3.

   Their payoffs are ranked, and the price expectations are modified as follows:

$$\mathbf{P}_i^e(t+1) = \begin{cases} \mathbf{P}_{i'}^e(t), & \text{if } \{U_i(t) < U_{i'}(t),\} \\ \mathbf{P}_i^e(t), & \text{if } \{U_i(t) > U_{i'}(t),\} \\ \text{Random}(\mathbf{P}_i^e(t), \mathbf{P}_{i'}^e(t)), & \text{if } \{U_i(t) = U_{i'}(t).\} \end{cases}$$

(14)

The protocol makes it easier for the agents to meet locally and enables them to exchange information. An agent who has performed well can influence someone who hasn't performed as well by modifying his perception of the economy (i.e., price expectations).

3.3 Meta Learning

In our model, an agent can consciously choose between social learning or individual learning during each period. In such a setting, it is necessary to specify the basis on which—i.e., how and when—such a choice is made. Agents may choose between these strategies either randomly, or based on the relative expected payoffs by employing these strategies. In our model, adaptive agents choose individual or social learning modes based on the past performance of these modes. In other words, the focus is on *learning how to learn*. This meta-learning framework is not restricted to two modes alone, but we begin with the simplest setting. We formulate this choice between different learning modes as a two-armed bandit problem.

3.3.1 Two-Armed Bandit Problem

In our market environment, an agent repeatedly chooses between two learning modes, individual learning and social learning . Denote the action space (feasible set of choice) by Γ, $\Gamma = \{a_{il}, a_{sl}\}$, where a_{il} and a_{sl} refer to the actions of individual learning and social learning , respectively. Each action chosen at time *t* by agent *i* yields a payoff $\pi(a_{k,t})(k = il, sl)$. This payoff is uncertain, but the agent can observe this payoff ex-post and this information is used to guide future choices. This setting is analogous to the familiar *two-armed bandit* problem. In the literature, reinforcement learning has been taken as a standard behavioural model for this type of choice problem [4] and we follow the same approach.

3.3.2 Reinforcement Learning

Reinforcement learning has been widely investigated both in artificial intelligence [23, 27] and economics [7, 19]. The intuition behind this learning scheme is that better performing choices are reinforced over time and those that lead to unfavourable or negative outcomes are not, or, alternatively, the better the experience with a particular choice, the higher is the disposition or the propensity towards choosing it in the future. This combination of *association* and *selection*, or, equivalently, *search* and *memory*, or the so-called the *Law of Effect* , is an important aspect of reinforcement learning [23, chapter 2].

In our model, each agent reinforces only two choices, i.e., individual learning and social learning ($\Gamma = \{a_{il}, a_{sl}\}$). In terms of reinforcement learning, the probability of a mode being chosen depends on the (normalized) propensity accumulated over time. Specifically, the mapping between the propensity and the choice probability is represented by the following Gibbs-Boltzmann distribution :

$$\mathrm{Prob}_{i,k}(t+1) = \frac{e^{\lambda_i \cdot q_{i,k} t)}}{e^{\lambda_i \cdot q_{i,a_{il}} t)} + e^{\lambda_i \cdot q_{i,a_{sl}} t)}}, \quad k \in \{a_{il}, a_{sl}\},$$

(15)

where $\mathrm{Prob}_{i,k}(t)$ is the choice probability for learning mode *k* ($k = a_{il}, a_{sl}$); we index this choice probability by *i* (the agent) and *t* (time) considering that different agents may have different experiences with respect to the same learning mode, and that, even for the same agent, experience may vary over time. The notation $q_{i,k}(t)$ ($k = a_{il}, a_{sl}$) denotes the propensity of the learning mode *k* for agent *i* at time *t*. Again, it is indexed by *t* because the propensity is revised from time to time based on the accumulated payoff. The propensity updating scheme applied here is the one-parameter version of [19].

$$q_{i,k}(t+1) = \begin{cases} (1-\phi) q_{i,k}(t) + U_i(t), & \text{if } \{k\} \text{ is chosen.} \\ (1-\phi) q_{i,k}(t), & \text{otherwise.} \end{cases}$$

(16)

where $U_i(t) \equiv U_i(\mathbf{X_i(t)})$, $k \in \{a_{il}, a_{sl}\}$, and ϕ is the so-called recency parameter , which can be interpreted as a memory-decaying factor.[4] The notation λ is known as the intensity of choice. With higher λs, the agent's choice is less random and is heavily biased toward the better-performing behavioural mode; in other words, the degree of exploration that the agent engages in is reduced. In the limit as $\lambda \rightarrow \infty$, the agent's choice is degenerated to the greedy algorithm which is only interested in the "optimal choice" that is conditional on the most recent updated experience; in this case, the agent no longer explores.

3.4 Reference Points

We further augment the standard reinforcement learning model with a reference-point mechanism to decide when Eq. (15) will be triggered. Reference dependence in decision making has been made popular by *prospect theory* [13], where gains and losses are defined relative to a reference point . This draws attention to the notion of *position* concerning the stimuli and the role it plays in cognitive coding that describes the agent's perception of the stimuli.

We augment our meta-learning model with reference points , where agents are assumed to question the appropriateness of their 'incumbent' learning mode *only* when their payoffs fall short of the reference point (along the lines of [6], p. 152–153). Let $U_i(t)(\equiv U_i(\mathbf{X_i(t)}))$ be the payoff of an agent i at time t. Let his reference point at time t be $R_i(t)$. The agent will consider a mode switch *only* when his realized payoff $U_i(t)$ is lower than his reference point $R_i(t)$.

The reference points indexed by t, $R_i(t)$, imply that they need not be static or exogenously given; instead, they can endogenously evolve over time with the experiences of the agents. Based on the current period payoffs, the reference point can be revised up or down. This revision can be symmetric, for example, a simple average of the current period payoffs and the reference point . A richer case as shown in Eq. (18) indicates that this revision can be asymmetric; in Eq. (18), the downward revisions are more pronounced than the upward revisions.

$$R_i(t+1) = \begin{cases} R_i(t) + \alpha^{+}\left(U_i(t) - R_i(t)\right), & if \;\; U_i(t) \geq R_i(t) \geq 0, \\ R_i(t) - \alpha^{-}\left(R_i(t) - U_i(t)\right), & if \;\; R_i(t) > U_i(t) \geq 0. \end{cases}$$

(18)

In the above equation, α^- and α^+ are revision parameters and $\alpha^-, \alpha^+ \in [0, 1]$. The case with $\alpha^- > \alpha^+$ would indicate that the agents are more sensitive to negative payoff deviations from their reference points.[5] Note that this is similar to the idea of loss aversion in prospect theory [24], where the slopes of the values function on either side of the reference point are different. For the rest of the simulations in this paper, we have utilized the asymmetric revision case.

4 Simulation Design

4.1 A Summary of the Model

4.1.1 Scale Parameters

Table 1 presents a summary of the agent-based Scarf model which we describe in Sects. 2 and 3. It also provides the values of the control parameters to be used in the rest of this paper. The Walrasian dynamics of the Scarf economy run in this model has the following fixed scale: a total of 270 agents (N = 270), three goods (M = 3), and hence 90 agents for each type of agent $N_1 = N_2 = N_3 = 90$. The initial endowment for each agent is 10 units (w_i = 10), which also serves as the Leontief coefficient for the utility function. Our simulation routine proceeds along the following sequence for each market day (period): production process (endowment replenishment), demand generation, trading, consumption and payoff realization, learning and expectations updating, propensity and reference point updating, and learning mode updating. Each market day is composed of 10,000 random matches (S = 10, 000), to ensure that the number of matches is large enough to exhaust the possibilities of trade. At the beginning of each market day, inventories perish (δ = 1) and the agents receive fresh endowments (10 units). Each single run of the simulation lasts for 2500 days (T = 2500) and each simulation series is repeated 50 times.[6] These scale parameters will be used throughout all simulations in Sect. 5.

Table 1

Table of control parameters

Parameter	Description	Value/Range
N	Number of agents	270
M	Number of types	3
N_1, N_2, N_3	Numbers of agents per type (4), (5)	90, 90, 90
w_i (i = 1, 2, 3)	Endowment (1)	10 units
	Leontief coefficients (2)	10
δ	Discount rate (Perishing rate)	1
S	Number of matches (one market day)	10,000
T	Number of market days	2500
$P_i^e(0)$ (i = 1, 2, 3)	Initial price expectations (3)	~ Uniform[0.5, 1.5]
φ	Parameter of price adjustment (11)	0.002
θ_1	Parameter of price adjustment (13)	~ Uniform[0,0.1]
θ_2	Parameter of price adjustment (13)	1
K	Number of arms (Sect. 3.3.1)	2
λ	Intensity of choice (15)	~ Uniform(0,8), ~Uniform(0,15),
		~Normal(4,1), ~Normal(8,1)
$\mathrm{POP}_{a_{il}}(0)$	Initial population of a_{il}	1/2
$\mathrm{POP}_{a_{sl}}(0)$	Initial population of a_{sl}	1/2
ϕ	Recency effect (16)	0
α^+	Degree of upward revision (18)	0.1

Parameter	Description	Value/Range
α^-	Degree of downward revision (18)	0.2

The numbers inside the parentheses in the second column refer to the number of the equations in which the respective parameter lies

4.1.2 Behavioural Parameters

The second part of the control parameters is related to the behavioural settings of the model, which begins with initial price expectations and price expectation adjustment, followed by the parameters related to the meta-learning models. As noted earlier, we set good 3 as a numéraire good, whose price is fixed as one. Most of these parameters are held constant throughout all simulations, as indicated in Table 1, and specific assumptions concerning distributions are specified in places where necessary.

We can systematically vary the values of the different parameters. The focus of this paper is on the intensity of choice (λ) (Table 1). The initial price vector, $\mathbf{P}_\mathbf{i}^\mathbf{e}(\mathbf{0})$, for each agent is randomly generated where the prices lie within the range [0.5,1.5], i.e., having the WE price as the centre. Except in the cases involving experimental variations of initial conditions, agents' learning modes, namely, individual learning (innovation) and social learning (imitation) are initially uniformly distributed ($\mathrm{POP}_{ail}(0) = \mathrm{POP}_{asl}(0) = 1/2$).

4.2 Implementation

All simulations and analysis were performed using NetLogo 5.2.0 and Matlab R2015a. To comply with the current community norm, we have made the computer program available at the OPEN ABM.[7] We classify the figure into five blocks. The first block (the left-most blocks) is a list of control bars for uses to supply the values of the control parameters to run the economy. The parameters are detailed in Table 1, and include N, M, S, T, φ, θ_1, θ_2, K, λ, POP_{RE}, and $\mathrm{POP}_{ail}(0)$. In addition to these variables, other control bars are for the selection of the running model, and include individual learning (only), social learning (only), the exogenously given market fraction, and the meta-learning model. For the exogenously given market fraction, $\mathrm{POP}_{ail}(0)$ needs to be given in addition.

5 Results

5.1 Heterogeneity and Scaling Up

Unlike the representative agent models in mainstream economic theory, agent-based models are capable of incorporating heterogeneity in terms of the varying characteristics of the agents. The importance of heterogeneity in agent-based models is further underscored because of its potential to break down linear aggregation due to interaction effect s. In this context, we can ask whether the quantitative phenomena observed in a heterogeneous economy are simply a linear combination of the respective homogeneous economies. If the aggregation of the effects is linear, the behaviour of the heterogeneous model can be simply deduced from the homogeneous economies. If not, studying a heterogeneous version of the economy becomes a worthwhile exercise to gain additional insights.

In our model, the only non-trivial source of heterogeneity comes from the intensity of choice (λ) that the agents possess. For simplicity, we simulate a heterogeneous economy with only two (equally populated) groups of agents with λ_i values of 1 and 7, respectively. To compare these, first, we need to determine the variable which will form the basis of comparison between the homogeneous and heterogeneous economies. Second, we need a method for determining the values of the above chosen variable across different homogeneous economies, which correspond to the heterogeneous economy in question. For the first, we consider the price deviations from the WE, measured in terms of the Mean Absolute Percentage Error (MAPE) of good 1, calculated based on the last 500 rounds of the simulation. The MAPE of the homogeneous economies, each with distinct values of λ is shown in the left panel of Fig. 1.

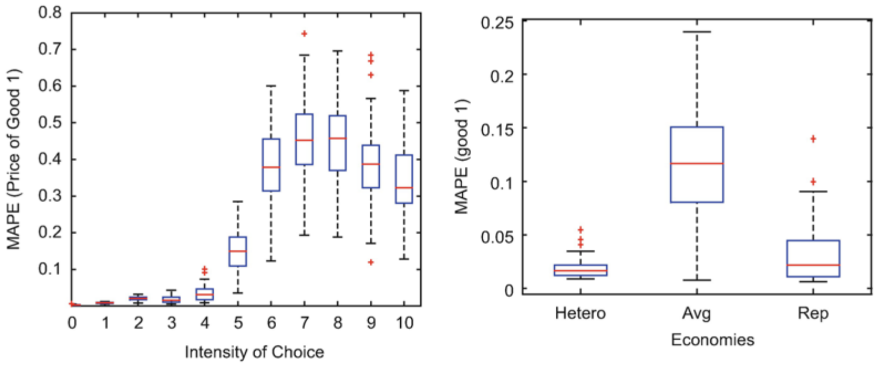

Fig. 1

Homogeneity vs. heterogeneity and aggregation: the above figures compare the values of the Mean Absolute Percentage Error (MAPE) of good 1 across different economies. The MAPE is calculated based on the last 500 periods of each run and it is defined as $\mathrm{MAPE}(F_j) = \frac{1}{500}\sum_{t=2001}^{2500} | F_j(t) - P_j^*|$ (j = 1, 2), where $P_j^* = 1$. The left panel indicates the price deviation for homogeneous economies according to λ. The right panel shows the comparisons between the heterogeneous and relevant homogeneous economies. The three economies that are compared are: (1) an economy consisting of agents with heterogeneous intensities of choice (**Hetero**), (2) an economy constructed by averaging the MAPE values of two homogeneous economies with λ = 1 and 7(**Avg**), and (3) a homogeneous economy with agents holding an 'average' value of intensity of choice (i.e., λ = 4, which is the average of 1 and 7.).(**Rep**). The boxplot shows the variability of the MAPE across 50 different runs for each economy

For the second, there are at least two ways to compare a heterogeneous economy (HE) (where 50% of the agents have λ = 1 and the other half have λ = 7) with homogeneous economies. Let M^{HE} represent the MAPE value in the heterogeneous economy. Let M_i^{HO} be the MAPE in a homogeneous economy, where all the agents have the same intensity of choice , i and Ω is the weighting constant. We can then consider an average of the MAPE values between the homogeneous economies with λ = 1 and 7, respectively. i.e.,

$$\bar{M}_{1,7}^{HO} = \Omega M_1^{HO} + (1 - \Omega) M_7^{HO}$$

(19)

For the case of a simple average that weights the MAPE values of the two economies equally, we have:

$$\bar{M}_{1,7}^{HO} = \frac{M_1^{HO} + M_7^{HO}}{2}$$

(20)

Another option would be to choose a single, homogeneous economy (**Rep**) that is *representative*, given the λ values in the heterogeneous counterpart. Instead of averaging the values of the two economies as shown above, we can choose MAPE values corresponding to the homogeneous economy with λ = 4, which is the midpoint of 1 and 7. We now have three different 'economies' to compare: (1) An economy with heterogeneous intensities of choice—λ_i = 1, 7(**Hetero**), (2) an economy representing the *average* MAPE values of two homogeneous economies with λ = 1 and 7(**Avg**), and (3) a *representative* homogeneous economy with agents holding an 'average' value of intensity of choice (i.e., λ = 4, the average of 1 and 7) (**Rep**). The boxplot in the right panel of Fig. 1 compares the MAPE values across these three versions. The whiskerplot shows the variability of MAPE across 50 different runs for each economy.

We find that the values of MAPE corresponding to the heterogeneous version are remarkably different from the other two versions in terms of the median values and its dispersion across runs. In particular, MAPE values of the **Avg** economy are much higher compared to the **Hetero** economy, indicating the absence of linear scaling up. From the right panel of Fig. 1, we observe that the MAPE across homogeneous economies is nonlinear with respect to λ. This partly explains the inadequacy of the analysis that solely relies on the information from homogeneous economies to understand the macro level properties. This nonlinearity, combined with potential interaction effects , indicates the breakdown of a linear scale-up and merits an independent investigation of the heterogeneous Scarf economy .

5.2 Price Convergence

In this section, we explore the ability to coordinate and steer the economy to the WE, starting from out-of-equilibrium configurations. To do so, we simulate the agent-based Scarf economy outlined in the previous section, where agents adaptively learn from their experiences using a meta-learning model. Given that the heterogeneity amongst agents is characterized in terms of their varying intensities of choice, we explore the influence of this parameter in detail. We simulate an economy with parameters indicated in Table 1. The agents differ in terms of the intensity of choice parameter (the exploration parameter, λ) and the intensities are uniformly distributed across agents[8] as $\lambda_i \in U(0, 8)$. The bounds of the said uniform distribution are chosen based on the initial insights gained from simulations with homogeneous agents (in terms of λ), where the behaviour of the economy is qualitatively similar for higher values of the parameter. We simulate the economy for 2500 periods for each run and conduct 50 such repetitions. The data on prices from these repetitions are pooled together and analysed in order to obtain robust results.

Figure 2 shows the relationship between the prices of goods 1 and 2 among different agents and their corresponding intensities of choice. The prices of goods 1 and 2 reported here are the mean market prices for each agent, which are calculated by considering the last 500 periods of each run, i.e., $P_{ij} = \frac{1}{500}\sum_{t=2001}^{2500} P_{ij}(t), (j = 1, 2)$. The agent-level data on mean prices across 50 repetitions are pooled together and plotted. We find that these mean prices resulting from the interaction and adaptive learning reveal some distinctive patterns. In particular, Fig. 2 shows a range of values of intensity of choice , roughly between 0 and 3, for which the agents remain in the neighbourhood of the WE. Beyond this range, the agents with higher values of λ do not necessarily have their mean prices confined to the neighbourhood of the WE. Even though some agents stay in the neighbourhood of the WE for $\lambda_i \in (3, 8)$, we observe a greater dispersion of mean prices on both sides of the WE. This is illustrated better in terms of the 3-D plot (Fig. 3) where the frequency of agents corresponding to different prices is shown on the z-axis. Although there is a dispersion on both of its sides, the WE still remains the predominant neighbourhood around which the mean prices gravitate. We also observe that there is a sizeable number of agents (more than 100) who hover close to zero as their mean price for the values of intensity of choice approaches 8.

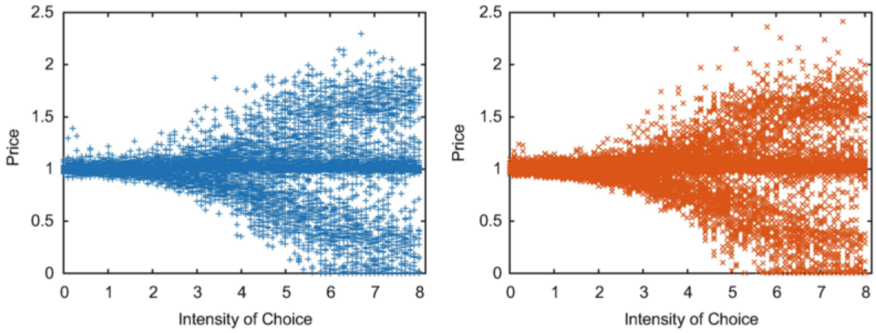

Fig. 2

Price convergence and heterogeneity: the above figure indicates the association between the mean prices for good 1 among different agents and their associated intensities of choice. The intensity of choice ranges between 0 and 8 and is distributed uniformly across agents in the economy, i.e., $\lambda_i \in U(0, 8)$. The prices of goods 1 and 2 reported here are the mean market prices for each agent which are calculated based on the last 500 periods of each run, i.e., $F_{ij} = \frac{1}{500}\sum_{t=2001}^{2500} F_{ij}(t), (j = 1, 2)$. The mean prices of some agents in the economy diverge away from the neighbourhood of the WE prices ($P_{ij}^* = 1$, for $j = 1, 2$) for values of intensity of choice greater than 3

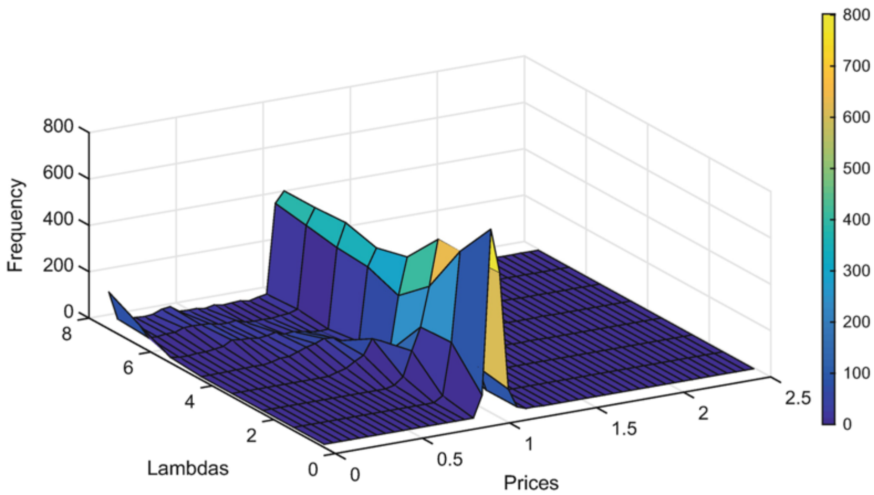

Fig. 3

Intensity of choice and price convergence for $\lambda_i \in U(0, 8)$: the above figure shows both the prices of good 1, the frequency of agents with those prices and the associated intensities of choice of each agent, where $\lambda_i \in U(0, 8)$. The prices of good 1 reported here are the mean market prices for each agent, calculated based on the last 500 periods of each run, i.e., $F_{i1} = \frac{1}{500}\sum_{t=2001}^{2500} F_{i1}(t)$. We carry out 50 repetitions of the simulation and pool together the data from all repetitions on the mean prices for all agents (270 × 50 = 13, 500) for the above plot. The agents are grouped into eight equally spaced bins based on their intensity of choice and the mean prices are grouped into 50 equally spaced bins

Notice that the distribution of the intensity of choice among agents is uniform and that the distribution is split between two regions, one with less dispersion in terms of mean prices with the WE as the centre and the other with higher dispersion ($\lambda \in (4, 8)$). We test whether the WE remains as the predominant neighbourhood with which heterogeneous agents coordinate by altering the proportion of agents between high and low dispersion regions in Fig. 2. To do so, we alter the bounds of the uniform distribution according to which agents are assigned different intensities of choice in the economy. We increase the upper bound of this distribution to 15, $\lambda_i \in U(0, 15)$ and examine how the mean prices vary according to λ. Figure 4 illustrates this relationship. We find that the results are qualitatively similar and that the WE continues to be the attraction point for the agents with an increasing dispersion with increasing λ. As in the previous case, a group of agents seems to cluster close to zero, which seems to be the dominant attraction point other than the WE for the agents in the economy.

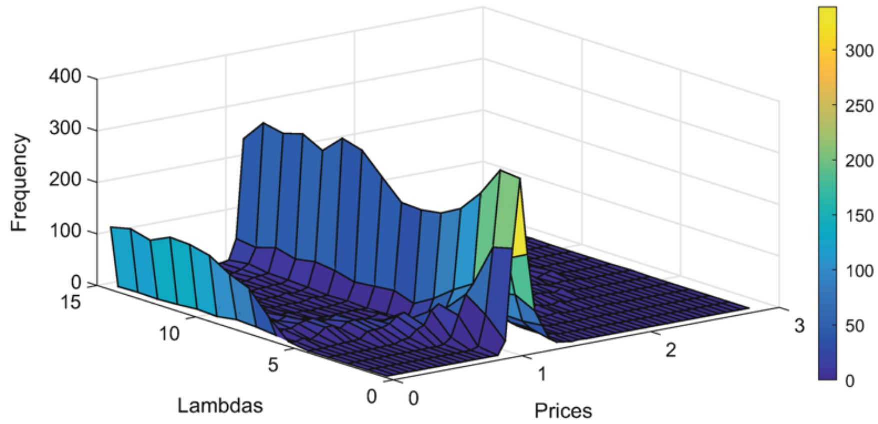

Fig. 4

Intensity of choice and price convergence for $\lambda_i \in U(0, 15)$: the above figure shows both the prices of good 1, the frequency of agents with those prices and the associated intensities of choice of each agent, where $\lambda_i \in U(0, 15)$. The prices of good 1 reported here are the mean market prices for each agent, calculated based on the last 500 periods of each run, i.e., $F_{i1} = \frac{1}{500}\sum_{t=2001}^{2500} F_{i1}(t)$. We carry out 50 repetitions of the simulation and pool together the data from all repetitions on the mean prices for all agents (270 × 50 = 13, 500) for the above plot

In addition to the normal distribution, we investigate the behaviour of the prices in economies with alternative distributional assumptions concerning the intensity of choice. Table 2 summarizes the prices of goods 1 and 2 at the aggregate level and at the level of agent types for 2 variations of normal and uniform distributions. For the case where $\lambda_i \in U(0, 8)$, the aggregate prices do stay within the 6% neighbourhood around the WE. They remain in a slightly larger neighbourhood (8%) for $\lambda_i \in U(0, 15)$, albeit with relatively higher variations among different runs. In the case of a normal distribution where $\lambda_i \in N(4, 1)$, the mean intensity of choice among agents lies where the increasing dispersion of prices takes off. The balance of stable and unstable tendencies explains the extremely high proximity of aggregate prices to the WE in this case. However, when we place the mean of λ_i in the normal distribution at 8, which is in the unstable region (i.e., $\lambda_i > 4$), the aggregate prices can be as much as 51% more than the WE.

Table 2

Prices of Goods 1 and 2 are shown at the aggregate level and according to the types of agents

	Aggregate		Type 1		Type 2		Type 3	
Prices/Distrib.	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
U [0,8]	0.943	0.995	0.889	0.922	0.873	0.931	1.068	1.132
	(0.004)	(0.014)	(0.152)	(0.133)	(0.173)	(0.148)	(0.008)	(0.026)
U [0,15]	1.080	1.044	0.924	0.888	0.948	0.925	1.370	1.320
	(0.014)	(0.030)	(0.324)	(0.280)	(0.280)	(0.299)	(0.086)	(0.072)
N (4,1)	1.000	1.001	0.932	0.949	0.939	0.928	1.130	1.127
	(0.004)	(0.005)	(0.169)	(0.152)	(0.166)	(0.174)	(0.016)	(0.008)
N(8,1)	1.407	1.510	1.197	1.296	1.109	1.270	1.914	1.964
	(0.089)	(0.008)	(0.404)	(0.573)	(0.571)	(0.528)	(0.434)	(0.019)

The prices reported are the mean market prices calculated based on the last 500 periods of each run, which are averaged over all agents. For the aggregate case, $F_{ij} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{500}\sum_{t=2001}^{2500}P_{ij}(t)$, where j = 1, 2 and N = 270. For the type wise prices, the mean prices are averaged over 90 agents. We carry out 50 repetitions of the simulation and the standard deviations across 50 runs are indicated in the parenthesis

Remember that agents with excess supply have to adjust the price expectations of production goods downward and adjust the price expectations of consumption goods upward (Eq. (10)). The upward bias in the divergence of prices stems from the fact that the agents endowed with the numerairé good who are dissatisfied can only adjust the price expectations of consumption goods upward and, unlike the agents endowed with other two goods, they cannot adjust their prices of their endowed goods downward. Those other agents with unsatisfied demand naturally bid up their prices for the goods in question. Therefore, all commodities receive more upward-adjusted potentials than downward-adjusted potential (2/3 vs 1/3 of the market participants), except for commodity 3, which serves as a numéraire. Hence, there is a net pulling force for the prices of commodities 1 and 2, leading them to spiral up.

Although we observe a high variation in mean prices among agents for larger values of λ_i, we still observe that a huge proportion of agents continue to remain in the neighbourhood of the WE for these high values of λ_i (see Figs. 3 and 4). This motivates us to search for other discriminating factors that might shed light on the dispersion of prices. In order to see whether the learning strategies chosen by the agents, viz., innovation or imitation, hold any advantage in enhancing price coordination, we classify the agents based on their strategy choices during the course of the simulation. We define *Normalized Imitation Frequency* (NIF), which indicates the ratio of the number of periods in which an agent has been an imitator over the entire course (2500 periods) of the simulation. If the agent has almost always been an innovator throughout the simulation, his NIF will be close to zero. On the contrary, if the agent has been an imitator throughout, the NIF will be equal to 1 and the intermediate cases lie between 0 and 1. The mean prices of good 1 held by the agents are plotted against the NIF in Fig. 5. The left and right panels in the figure correspond to the distributions $\lambda_i \in U(0, 8)$ and $\lambda_i \in U(0, 15)$, respectively. The entries in red correspond to agents with $\lambda_i > 4$ and those in blue denote agents with $\lambda_i < 4$.

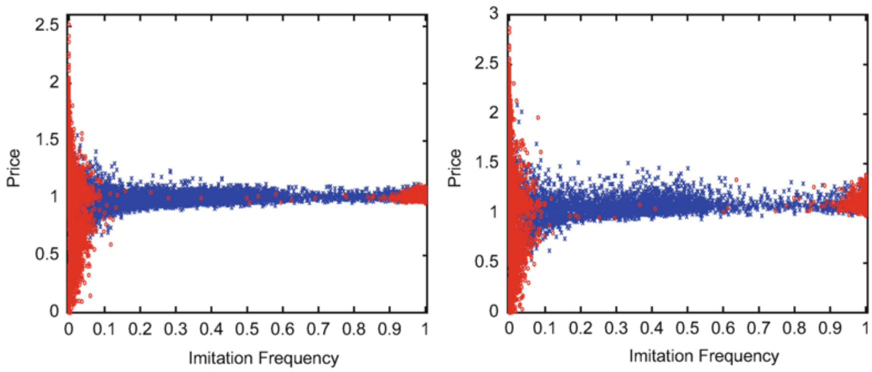

Fig. 5

Mean price of Good 1 and imitation frequency (normalized): the figure indicates the price distribution among all agents and their associated imitation frequencies. The intensity of choice (λ) is distributed uniformly across agents in the economy and the left and the right panel denote economies with distributions $\lambda_i \in U(0, 8)$ and $\lambda_i \in U(0, 15)$. Agents with $\lambda_i \leq 4$ are denoted by blue and those in red denote agents with $\lambda_i > 4$. The prices of good 1 reported here are the mean market prices for each agent, calculated based on the last 500 periods of each run, i.e., $F_{i1} = \frac{1}{500}\sum_{t=2001}^{2500} F_{i1}(t)$. We carry out 50 repetitions of the simulation and pool together the data from all repetitions on mean prices for all agents ($270 \times 50 = 13,500$) for the above plot. The imitation frequency (normalized) denotes the ratio of the number of periods in which an agent has been an imitator over the entire course (2500 periods) of each repetition of the simulation

We observe that the there are two distinct clusters for agents in red ($\lambda_i > 4$). Agents with high intensities of choice who happen to be 'almost always innovators' (with an NIF close to 1) can be seen to remain in the neighbourhood of WE, while the high dispersion in prices seems to be coming from innovating agents with a high λ_i. In our meta-learning model characterized by the reinforcement learning mechanism, agents reduce their tendency to explore for higher values of λ_i and get locked into one of the two learning strategies. These 'immobile agents' populate either corner of the NIF scale. Figure 5 indicates that the price strategy to which agents get locked in while cutting down their exploration seems to have a decisive impact on whether or not agents can coordinate themselves in the neighbourhood of the WE.

To summarize, in this section, we find that the ability of the agents to steer themselves toward the WE on aggregate depends on their tendency to explore that is captured by the intensity of choice. The distributional assumptions concerning the intensity of choice among agents influence whether or not the aggregate prices are closer to the WE. More specifically, the relative proportion of agents with λ_i corresponding to less and more volatile regions, roughly on either side of $\lambda_i = 4$ is crucial and the larger the proportion of the latter, the higher is the deviation from the WE in the aggregate. Despite the high variability in prices among agents with a high λ, a sizeable number still remain in the neighbourhood of the WE. Agents who do not explore and remain predominantly as innovators for all periods exhibit a higher dispersion in their mean prices.

5.3 Current and Accumulated Payoffs

In this section, we investigate whether there are any significant differences amongst agents in terms of their current and the accumulated payoffs.[9]

Figure 6 shows the relationship between the accumulated payoffs (or 'lifetime wealth') of agents and their corresponding intensity of choice (λ_i). While there is a noticeable variability in accumulated payoffs for agents with higher intensities of choice ($\lambda_i > 4$), however, λ alone does not seem to sufficiently explain the cause. For instance, by increasing λ_i beyond 4, some (but not all) agents seem to be increasingly worse off in terms of their accumulated payoffs that keep falling. However, some other agents with the same values of λ_i enjoy much higher payoffs, which are comparable to those with $\lambda_i <$. In other words, there is a wealth inequality that emerges among agents over time. We need to identify the characteristics of agents that are responsible for these payoff differences.

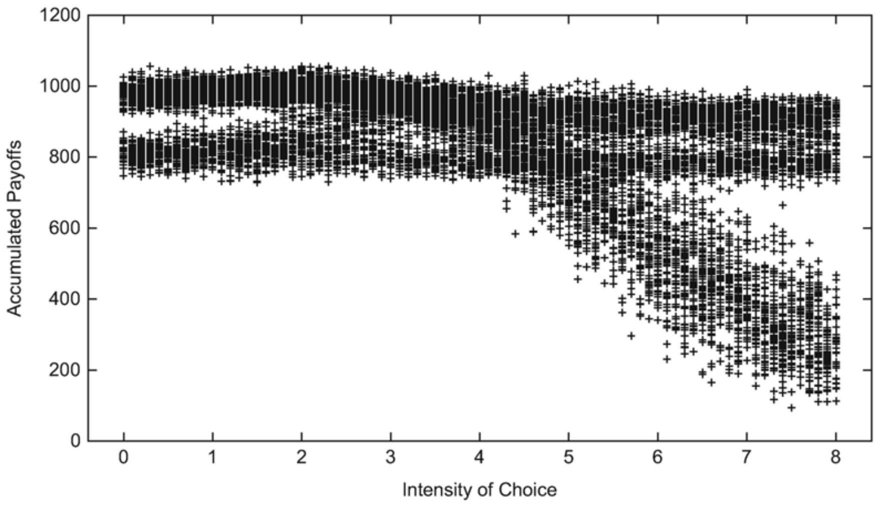

Fig. 6

Accumulated payoffs: the figure shows the accumulated payoffs for all agents (i), where $\lambda_i \in U(0, 8)$. The variability in accumulated payoffs is not entirely explained by variations in their intensity of choice alone. Notice that there are remarkable deviations between agents in terms of their accumulated payoffs for $\lambda_i > 4$

Since our agents are similar in terms of all relevant characteristics other than λ, potential indirect channels of influence stemming from λ need to be examined. Strategy choices by agents—to innovate or imitate—over the period constitute one such channel since they are governed by the λ parameter in the meta-learning (reinforcement) model. In the previous section, we learnt that imitating agents managed to coordinate to the WE on average, despite having a high λ. Along the same lines, we examine whether imitation and innovation also impact the payoffs among agents. Figure 7 shows how the current payoffs of agents vary according to their NIF for $\lambda_i \in U(0, 8)$ and $\lambda_i \in U(0, 15)$ in the left and right panels, respectively. Agents with $\lambda_i \leq 4$ are shown in blue and those with $\lambda_i > 4$ are shown in red. This indicates that agents who are predominantly innovators (for a high λ) seem to have high variability among themselves in terms of their current payoffs and agents who are predominantly imitators seem to enjoy relatively higher payoffs with considerably less variability.

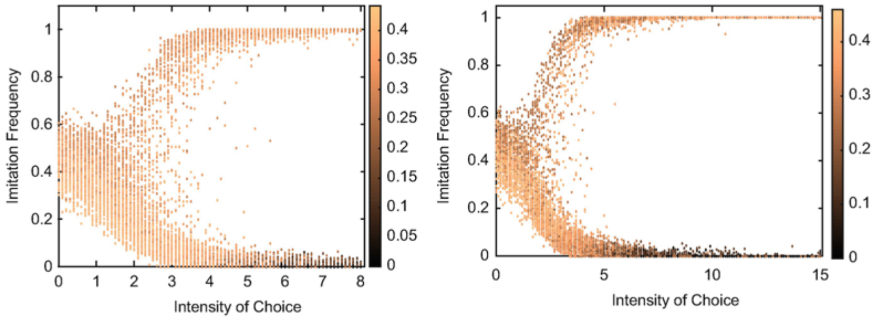

Fig. 7

Intensity of choice , imitation frequency and current payoffs: the figures in the left and right panels denote the relationship between λ and the normalized mean value imitation frequency (NIF) for all agents based on 50 repetitions, corresponding to $\lambda_i \in U(0, 8)$ and $\lambda_i \in U(0, 15)$, respectively. NIF values of 0 and 1 indicate that the agent is an innovator and imitator for 100% of the time, respectively. The heat map denotes the associated mean current payoffs, which are calculated based on the last 500 periods

In Fig. 7 we infer a clearer idea of the underlying phenomena. The left and right panels denote the relationship between intensity of choice and normalized frequency of imitation for $\lambda_i \in U(0, 8)$ and $\lambda_i \in U(0, 15)$, respectively. As the intensity of choice increases, there is a polarization in terms of NIF : for lower values of λ (roughly, $0<\lambda < 3$), NIF values span a wide range and they are not clustered close to the end of the interval (0,1). These intermediate values indicate that agents switch between their strategies (innovation and imitation) and are 'mobile', thus balancing between exploration and exploitation . By contrast, for λ values of 4 and beyond, we observe that there is a drastic fall in the number of agents and the agents are either predominantly innovators or imitators. This phenomenon can be explained by the *power law of practice* that is well recognized in psychology.[10] The heat map in Fig. 7 indicates that the distribution of current payoffs varies among the agents. Mobile agents (with intermediate values of NIF) and agents who are predominantly imitators enjoy higher current payoffs compared to agents who are predominantly innovators (close to the X-axis). This pattern persists despite changes in assumptions regarding the distributions indicated above.

5.3.1 Accumulated Payoffs

We now turn to the differences among agents in terms of their accumulated payoffs that we observed in Fig. 6. We group agents into different groups based on their accumulated payoffs and analyse specific characteristics associated with agents in each of these groups that could explain the differences in their accumulated payoffs.

Figure 8 shows the agents clustered into four different groups (Very High, High, Medium, Low) based on their accumulated payoffs for $\lambda_i \in U(0, 8)$ and $U(0, 15)$ in the left and right panels, respectively. We use the K-means technique (Lloyd's algorithm) to group the agents, which allows us to group agents into K mutually exclusive clusters based on the distance from the K different centroids. We use $K = 4$ based on visual heuristics and it provides the best classification for the intended purpose. Higher values of K do not show improvements in terms of their silhouette values that measure the degree of cohesion with other points in the same cluster compared to other clusters.

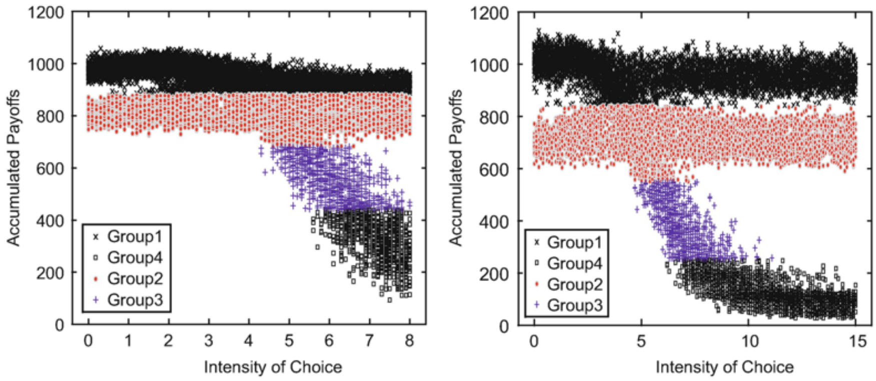

Fig. 8

Accumulated payoffs—clustered: the figures on the left and right show the accumulated payoffs for all agents (i), where the intensities are distributed uniformly as $\lambda_i \in U(0, 8)$ and $\lambda_i \in U(0, 15)$, respectively. Based on their accumulated payoffs, the agents are clustered in four distinct groups using the *K*-means clustering technique

We showed earlier that eventual strategy choices associated with a high λ could explain variations in current payoffs. We examine whether agents (with high intensity of choice) who are predominantly imitators end up with significantly higher accumulated payoffs compared to innovating agents. Table 3 shows the four different groups consisting of agents pooled from 50 different runs of the simulation. The table also indicates the number of agents in each group and the group-wise averages of accumulated payoffs, intensity of choice, and imitation frequency. Agents with relatively better average accumulated payoffs for the two high groups differ significantly in terms of their (AIF) compared to those for the medium and low groups. Average Imitation Frequency measures the average number of periods during which agents have been imitators in the 2500 periods of the simulation.

Table 3

Accumulated payoffs of heterogeneous agents: table compares four different clusters of agents, who are pooled from 50 different runs of the simulation

	U(0,8)				**U(0,15)**			
	N	***Π***	**AIF**	$\bar{\lambda}$	***N***	***Π***	**AIF**	$\bar{\lambda}$
Low	891	320.29	24.797	7.175	2006	123.71	9.485	11.266
Medium	902	558.42	53.396	6.038	725	383.83	40.662	6.906
High	4329	805.11	1215.298	4.127	4911	714.08	1243.968	7.322
Very High	7378	957.85	965.453	3.311	5858	968.42	1607.920	6.293

These are the number of agents in each group (N) and the group-wise averages of the accumulated payoffs (*Π*), intensity of choice ($\bar{\lambda}$), and imitation frequency (AIF)

From Fig. 1, we see that agents who have $\lambda > 4$ are those that constitute the medium and low payoff clusters. From Table 3, by comparing low and high clusters in U(0,8), we observe that agents with higher AIF (predominantly imitators) have relatively higher accumulated payoffs. However, the significance of AIF comes into play largely for agents with $\lambda > 4$. For lower values of λ the strategy choices are not tilted predominantly towards innovation or imitation . In this range, both innovating and imitating agents have a chance to obtain medium or high accumulated payoffs.

The strength of AIF as a discriminating factor becomes evident when comparing different clusters in U(0,15), where the number of agents with a high λ are in relative abundance. This relative abundance helps for a robust understanding of payoff inequality among agents with a high λ. Table 3 denotes the monotonic and increasing relationship between accumulated payoffs and AIF between clusters. The average accumulated payoff for agents from the 'Very High' cluster is 968.4, compared to 123.7 for those in the 'Low' cluster. Their corresponding AIF are 9.5 and 1607.9. Agents with abysmal accumulated payoffs for comparable λ values are explained by their strategy choice—reluctance to imitate.

6 Discussion

We have examined the possibility of coordination among agents in a stylized version of a decentralized market economy. We have analysed whether these agents can reach the equilibrium (on average) in a Scarf-like economy starting from disequilibrium states. In the presence of heterogeneity, Sect. 5.1 demonstrated that it may not be adequate to look only at corresponding economies with homogeneous agents and extrapolate patterns concerning price deviations. A straightforward linear aggregation breaks down in our model due to potential interaction effects and more importantly, due to the presence of a non-linear relationship between intensity of choice —the sole source of heterogeneity in our model—and the MAPE . The consequences of diversity among agents in shaping macroeconomic outcomes can be readily investigated using agent-based models compared to equation-based models.

Agents can and do succeed in coordinating themselves without a centralized mechanism, purely through local interaction and learning . They learn to revise private beliefs and successfully align their price expectations to those of the WE. However, this is not guaranteed in all cases and the coordination to the WE critically depends on a single behavioural parameter, viz., intensity of choice (λ). The level of intensity indirectly captures the degree or the frequency with which agents engage in exploration by trying different strategies. Lower levels of λ are associated with a higher degree of exploration and vice versa. When the intensity of choice is zero, it is equivalent to the case where agents choose between innovation and imitation with equal probability. As the λ values rise, agents increasingly stick to the 'greedy' choice and explore less and less. As a consequence, we have fewer agents who are switching between their strategies at higher λ values. They remain immobile and the strategy choices are thus polarized with AIF being either 0 or 1.

Since agents differ only in terms of λ and coordination success is linked to λ, the *extent* and nature of diversity among agents in an economy is crucial. It is hard to speak about the possibility of success in coordination independently of the distributional assumptions regarding λ. As pointed out in Sect. 5.2, if a relatively large proportion of agents have high λ values, the coordination to equilibrium is less likely to be successful, when compared to an economy with most agents holding lower intensities of choice. In sum, we find that (a) the balance between exploration and exploitation is crucial for each agent to reach the neighbourhood of equilibrium prices, and (b) the level of heterogeneity (distributions of λ)

among agents determines the proportion of agents who explore little. This, in turn, determines the extent of the average price deviation from the WE for the economy as a whole (see Table 2).

The influence of a high λ (or 'strong tastes') of agents, although important, is not the whole story. Similarly, sole reliance on meso-level explanations that look at the fixed proportion of agents holding different strategies (market fraction) is also inadequate in explaining the success or failure of coordination. From Fig. 5, we find that individuals can reach the neighbourhood of the WE prices for a large intermediate range ($0 < NIF < 1$) where the agents are mobile. We also find that having a strong affinity to a particular strategy that performs well during the early periods is not detrimental per se. Instead, the *nature* of the strategy to which an agent is committed matters more. Notice from the figure that agents who are pure imitators ($NIF \approx 1$) also reach the neighbourhood of the WE, just as mobile agents do. Thus, even with high intensity of choice , agents who explore (social learning) are more effective in correcting their misperceptions unlike those who learn only from their own experiences.

We observe that inequality emerges among agents in their accumulated payoffs. There have been different explanations concerning the emergence of income and wealth inequality among individuals and nations, such as social arrangements that deny access to resources and opportunities to some groups vis-à-vis others, differences in initial conditions like endowments or productivity [12], and financial market globalization [15], to name a few. Other scholars see wealth and income inequality as a by-product of the capitalistic system and due to differential rewards to various factors of production [17].[11] In addition to these explanations, we point to another possible source of inequality over time: differences in learning behaviour. Our analysis shows that payoff inequality can potentially stem from diversity in expectations, which is traced back to a micro-level parameter that governs the learning behaviour of agents. Even though agents have the same amount of endowments, some become relatively worse off over time due to their choice of learning strategy. Agents who do not explore enough and learn from others end up receiving drastically lower payoffs and thus remain poor.

7 Conclusion

We have examined the possibility of price discovery in a decentralized, heterogeneous agent-based Scarf economy and whether or not agents can coordinate themselves to the WE starting from disequilibrium. We find that the coordination success is intimately tied to learning mechanisms employed by agents and, more precisely, how they find a balance between exploration and exploitation . A reduction in or an absence of exploration has an adverse effect on accumulated payoffs and coordination possibilities, endogenously giving rise to inequality. Social learning has often been considered to be inferior and there has been a relatively bigger focus on individual learning . It turns out that there is plenty that we can learn from others after all. Neither of these learning modes is in itself sufficient to ensure efficient outcomes and hence it may be necessary to balance them both. Although our model is highly stylized, with the results lacking the desired generality, it brings us a step closer towards comprehending the role of heterogeneity, and learning dynamically shapes various aggregate outcomes.

Footnotes

1. 1.

More specifically, t refers to the whole market day t, i.e., the interval [t, t – 1).

2. 2.

Given this symmetry, the type-wise grouping of agents does not qualify as a source of real heterogeneity in terms of characteristics. This, as we will see in the later sections, will be characterized in terms of the agents' learning mechanisms.

3. 3.

More precisely, the good that he is endowed with.

4. 4.

Following [4], a normalization scheme is also applied to normalize the propensities $q_{i,k}(t+1)$ as follows:

$$q_{i,k}(t+1) \leftarrow \frac{q_{i,k}(t+1)}{q_{i,a_{il}}(t+1) + q_{i,a_{sl}}(t+1)}. \quad (17)$$

5. 5.

The results do not vary qualitatively for perturbations of these parameters.

6. 6.

We have examined the system by simulating the same treatments for much longer periods and we find that the results are robust to this.

7. 7.

https://www.openabm.org/model/4897/
(https://www.openabm.org/model/4897/)

8. 8.

The granularity of λ values is chosen in increments of 0.1, and hence there are 81 distinct values of λ_i corresponding to U(0,8).

9. 9.

In our model, there is no material wealth or savings since all goods are consumed entirely within each period. However, it is possible to interpret the accumulated payoffs of each agent over the simulation period as a proxy for their 'quality of life' or 'lifetime wealth'.

10. 10.

In the psychology literature, the *power law of practice* states that the subjects' early learning experiences have a dominating effect on their limiting behaviour. It is characterized by initially steep but then flatter learning curves.

11. 11.

See also: [16].

Notes

Acknowledgements

Shu-Heng Chen and Ragupathy Venkatachalam are grateful for the research support provided in the form of the Ministry of Science and Technology (MOST) grants, MOST 103-2410-H-004-009-MY3 and MOST 104-2811-H-004-003, respectively.

References

1. Alós-Ferrer, C., & Schlag, K. H. (2009). Imitation and learning. In P. Anand, P. Pattanaik and C. Puppe (Eds.), The handbook of rational and social choice. New York: Oxford University Press.
Google Scholar (http://scholar.google.com/scholar_lookup?title=Imitation%20and%20learning&author=C.%20Al%C3%B3s-Ferrer&author=K%20H.%20Schlag&publication_year=2009)
2. Anderson, C., Plott, C., Shimomura, K., & Granat, S. (2004). Global instability in experimental general equilibrium: The Scarf example. *Journal of Economic Theory, 115*(2), 209–249.
CrossRef (https://doi.org/10.1016/S0022-0531(03)00185-6)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Global%20instability%20in%20experimental%20general%20equilibrium%3A%20The%20Scarf%20example&author=C.%20Anderson&author=C.%20Plott&author=K.%20Shimomura&author=S.%20Granat&journal=Journal%20of%20Economic%20Theory&volume=115&issue=2&pages=209-249&publication_year=2004)
3. Arifovic, J., & Ledyard, J. (2004). Scaling up learning models in public good games. *Journal of Public Economic Theory, 6*, 203–238.
CrossRef (https://doi.org/10.1111/j.1467-9779.2004.00165.x)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Scaling%20up%20learning%20models%20in%20public%20good%20games&author=J.%20Arifovic&author=J.%20Ledyard&journal=Journal%20of%20Public%20Economic%20Theory&volume=6&pages=203-238&publication_year=2004)
4. Arthur, B. (1993) On designing economic agents that behave like human agents. *Journal of Evolutionary Economics 3*(1), 1–22.
CrossRef (https://doi.org/10.1007/BF01199986)
Google Scholar (http://scholar.google.com/scholar_lookup?title=On%20designing%20economic%20agents%20that%20behave%20like%20human%20agents&author=B.%20Arthur&journal=Journal%20of%20Evolutionary%20Economics&volume=3&issue=1&pages=1-22&publication_year=1993)
5. Bossan, B., Jann, O., & Hammerstein, P. (2015). The evolution of social learning and its economic consequences. *Journal of Economic Behavior & Organization, 112*, 266–288.
CrossRef (https://doi.org/10.1016/j.jebo.2015.01.010)
Google Scholar (http://scholar.google.com/scholar_lookup?title=The%20evolution%20of%20social%20learning%20and%20its%20economic%20consequences&author=B.%20Bossan&author=O.%20Jann&author=P.%20Hammerstein&journal=Journal%20of%20Economic%20Behavior%20%26%20Organization&volume=112&pages=266-288&publication_year=2015)

6. Erev, I., & Rapoport, A. (1998). Coordination, “magic,” and reinforcement learning in a market entry game. *Games and Economic Behavior, 23*, 146–175.
CrossRef (https://doi.org/10.1006/game.1997.0619)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Coordination%2C%20%E2%80%9Cmagic%E2%80%9D%2C%20and%20reinforcement%20learning%20in%20a%20market%20entry%20game&author=I.%20Erev&author=A.%20Rapoport&journal=Games%20and%20Economic%20Behavior&volume=23&pages=146-175&publication_year=1998)

7. Erev, I., & Roth, A. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review, 88*(4), 848–881.
Google Scholar (http://scholar.google.com/scholar_lookup?title=Predicting%20how%20people%20play%20games%3A%20Reinforcement%20learning%20in%20experimental%20games%20with%20unique%2C%20mixed%20strategy%20equilibria&author=I.%20Erev&author=A.%20Roth&journal=American%20Economic%20Review&volume=88&issue=4&pages=848-881&publication_year=1998)

8. Fisher, F. M. (1983). *Disequilibrium foundation of equilibrium economics*. Cambridge, UK: Cambridge University Press.
CrossRef (https://doi.org/10.1017/CBO9781139052108)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Disequilibrium%20foundation%20of%20equilibrium%20economics&author=F%20M.%20Fisher&publication_year=1983)

9. Gintis, H. (2007). The dynamics of general equilibrium. *Economic Journal, 117*(523), 1280–1309.
CrossRef (https://doi.org/10.1111/j.1468-0297.2007.02083.x)
Google Scholar (http://scholar.google.com/scholar_lookup?title=The%20dynamics%20of%20general%20equilibrium&author=H.%20Gintis&journal=Economic%20Journal&volume=117&issue=523&pages=1280-1309&publication_year=2007)

10. Hahn, F. H., & Negishi, T. (1962). A theorem on non-tâtonnement stability. *Econometrica, 30*(3), 463–469.
CrossRef (https://doi.org/10.2307/1909889)
Google Scholar (http://scholar.google.com/scholar_lookup?title=A%20theorem%20on%20non-t%C3%A2tonnement%20stability&author=F%20H.%20Hahn&author=T.%20Negishi&journal=Econometrica&volume=30&issue=3&pages=463-469&publication_year=1962)

11. Hoppitt, W., & Laland, K. N. (2008). Social processes influencing learning in animals: A review of the evidence. *Advances in the Study of Behavior, 38*, 105–165.
CrossRef (https://doi.org/10.1016/S0065-3454(08)00003-X)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Social%20processes%20influencing%20learning%20in%20animals%3A%20A%20review%20of%20the%20evidence&author=W.%20Hoppitt&author=K%20N.%20Laland&journal=Advances%20in%20the%20Study%20of%20Behavior&volume=38&pages=105-165&publication_year=2008)

12. Hopkins, E., & Kornienko, T. (2010). Which inequality? The inequality of endowments versus the inequality of rewards. *American Economic Journal: Microeconomics, 2*(3), 106–37.
Google Scholar (http://scholar.google.com/scholar_lookup?title=Which%20inequality%3F%20The%20inequality%20of%20endowments%20versus%20the%20inequality%20of%20rewards&author=E.%20Hopkins&author=T.%20Kornienko&journal=American%20Economic%20Journal%3A%20Microeconomics&volume=2&issue=3&pages=106-137&publication_year=2010)

13. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.
CrossRef (https://doi.org/10.2307/1914185)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Prospect%20theory%3A%20An%20analysis%20of%20decision%20under%20risk&author=D.%20Kahneman&author=A.%20Tversky&journal=Econometrica&volume=47&pages=263-291&publication_year=1979)

14. Laland, K. (2004). Social learning strategies. *Animal Learning & Behavior, 32*(1), 4–14.
CrossRef (https://doi.org/10.3758/BF03196002)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Social%20learning%20strategies&author=K.%20Laland&journal=Animal%20Learning%20%26%20Behavior&volume=32&issue=1&pages=4-14&publication_year=2004)

15. Matsuyama, K. (2004). Financial market globalization, symmetry-breaking, and endogenous inequality of nations. *Econometrica*, 72(3), 853–884.
CrossRef (https://doi.org/10.1111/j.1468-0262.2004.00514.x)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Financial%20market%20globalization%2C%20symmetry-breaking%2C%20and%20endogenous%20inequality%20of%20nations&author=K.%20Matsuyama&journal=Econometrica&volume=72&issue=3&pages=853-884&publication_year=2004)

16. Piketty, T. (2000), Theories of persistent inequality and intergenerational mobility. Chapter 8, *Handbook of Income Distribution* (vol. 1, pp. 429–476). Amsterdam: Elsevier.
Google Scholar (https://scholar.google.com/scholar?q=Piketty%2C%20T.%20%282000%29%2C%20Theories%20of%20persistent%20inequality%20and%20intergenerational%20mobility.%20Chapter%208%2C%20Handbook%20of%20Income%20Distribution%20%28vol.%201%2C%20pp.%20429%E2%80%93476%29.%20Amsterdam%3A%20Elsevier.)

17. Piketty, T. (2014), *Capital in the Twenty-First Century*. Cambridge: Harvard University Press.
CrossRef (https://doi.org/10.4159/9780674369542)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Capital%20in%20the%20Twenty-First%20Century&author=T.%20Piketty&publication_year=2014)

18. Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., et al. (2010). Why copy others? Insights from the social learning strategies tournament. *Science, 328*(5975), 208–213.
CrossRef (https://doi.org/10.1126/science.1184719)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Why%20copy%20others%3F%20Insights%20from%20the%20social%20learning%20strategies%20tournament&author=L.%20Rendell&author=R.%20Boyd&author=D.%20Cownden&author=M.%20Enquist&author=K.%20Eriksson&author=M%20W.%20Feldman&journal=Science&volume=328&issue=5975&pages=208-213&publication_year=2010)

19. Roth, A., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behaviour, 8*, 164–212.
CrossRef (https://doi.org/10.1016/S0899-8256(05)80020-X)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Learning%20in%20extensive-form%20games%3A%20Experimental%20data%20and%20simple%20dynamic%20models%20in%20the%20intermediate%20term&author=A.%20Roth&author=I.%2

0Erev&journal=Games%20and%20Economic%20Behaviour&volume=8&pages=164-212&publication_year=1995)

20. Scarf, H. (1960). Some examples of global instability of the competitive economy. *International Economic Review, 1*(3), 157–172.
CrossRef (https://doi.org/10.2307/2556215)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Some%20examples%20of%20global%20instability%20of%20the%20competitive%20economy&author=H.%20Scarf&journal=International%20Economic%20Review&volume=1&issue=3&pages=157-172&publication_year=1960)

21. Schlag, K. (1998). Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory, 78*(1), 130–156.
CrossRef (https://doi.org/10.1006/jeth.1997.2347)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Why%20imitate%2C%20and%20if%20so%2C%20how%3F%20A%20boundedly%20rational%20approach%20to%20multi-armed%20bandits&author=K.%20Schlag&journal=Journal%20of%20Economic%20Theory&volume=78&issue=1&pages=130-156&publication_year=1998)

22. Schlag, K. (1999). Which one should I imitate? *Journal of Mathematical Economics, 31*(4), 493–522.
CrossRef (https://doi.org/10.1016/S0304-4068(97)00068-2)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Which%20one%20should%20I%20imitate%3F&author=K.%20Schlag&journal=Journal%20of%20Mathematical%20Economics&volume=31&issue=4&pages=493-522&publication_year=1999)

23. Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction*, Cambridge, MA: MIT Press.
Google Scholar (http://scholar.google.com/scholar_lookup?title=Reinforcement%20learning%3A%20An%20introduction&author=R.%20Sutton&author=A.%20Barto&publication_year=1998)

24. Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics, 106*, 1039–1061.
CrossRef (https://doi.org/10.2307/2937956)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Loss%20aversion%20in%20riskless%20choice%3A%20A%20reference-dependent%20model&author=A.%20Tversky&author=D.%20Kahneman&journal=Quarterly%20Journal%20of%20Economics&volume=106&pages=1039-1061&publication_year=1991)

25. Uzawa, H. (1960). Walras' tâtonnement in the theory of exchange. *The Review of Economic Studies, 27*(3), 182–194.
CrossRef (https://doi.org/10.2307/2296080)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Walras%E2%80%99%20t%C3%A2tonnement%20in%20the%20theory%20of%20exchange&author=H.%20Uzawa&journal=The%20Review%20of%20Economic%20Studies&volume=27&issue=3&pages=182-194&publication_year=1960)

26. Vriend, N. (2000). An illustration of the essential difference between individual and social learning, and its consequences for computational analyses. *Journal of Economic Dynamics & Control, 24*(1), 1–19.
CrossRef (https://doi.org/10.1016/S0165-1889(98)00068-2)
Google Scholar (http://scholar.google.com/scholar_lookup?title=An%20illustration%20of%20the%20essential%20difference%20between%20individual%20and%20social%20learning%2C%20and%20its%20consequences%20for%20computational%20analyses&author=N.%20Vriend&journal=Journal%20of%20Economic%20Dynamics%20%26%20Control&volume=24&issue=1&pages=1-19&publication_year=2000)

27. Wiering, M., & van Otterlo, M. (2012). *Reinforcement learning: State of the art*. Heidelberg: Springer.
CrossRef (https://doi.org/10.1007/978-3-642-27645-3)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Reinforcement%20learning%3A%20State%20of%20the%20art&author=M.%20Wiering&author=M.%20Otterlo&publication_year=2012)

Copyright information

About this paper

Cite this paper as:

Chen SH., Chie BT., Kao YF., Magerl W., Venkatachalam R. (2018) Heterogeneity, Price Discovery and Inequality in an Agent-Based Scarf Economy. In: Chen SH., Kao YF., Venkatachalam R., Du YR. (eds) Complex Systems Modeling and Simulation in Economics and Finance. CEF 2015. Springer Proceedings in Complexity. Springer, Cham

- First Online 21 November 2018
- DOI https://doi.org/10.1007/978-3-319-99624-0_6
- Publisher Name Springer, Cham
- Print ISBN 978-3-319-99622-6
- Online ISBN 978-3-319-99624-0
- eBook Packages Economics and Finance