University of Warsaw
Faculty of Economic Sciences

Working Papers
No. 31/2026 (525)

Co-adaptation: From MARL Non-Stationarity to the Dynamics of Economic Adjustment

Jan Lisicki
Tomasz Kopczewski

Warsaw 2026

ISSN 2957-0506

UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
WORKING PAPERS

Co-adaptation: From MARL Non-stationarity to the Dynamics of Economic Adjustment

Jan Lisicki[1], Tomasz Kopczewski[1]*

[1] *University of Warsaw, Faculty of Economic Sciences*
[*] *Corresponding author: jt.lisicki@uw.edu.pl*

Abstract: Economic agents adapt their behavior to experience generated by interactions with other market participants. When several agents update their behavioral rules simultaneously, these adaptation processes become interdependent. We refer to this mechanism as co-adaptation. The paper develops a formal representation that separates the continuation of an agent's own rule updates from cross-agent effects on subsequent updates. In a two-agent system, mutual influence can increase the cumulative effect of an initial change and prolong the resulting adjustment process. A simple model of price adjustment between two firms shows how these effects depend on adjustment speed and the response to a competitor's price change. With more agents, interdependencies between adaptation processes form a directed network. This perspective links economic adjustment to multi-agent reinforcement learning, where learning by one agent changes the environment experienced by others, and clarifies when such non-stationarity is part of economic dynamics rather than only a computational difficulty.

Keywords: economic adjustment, multi-agent reinforcement learning, adaptive learning, agent-based computational economics, co-adaptation

JEL codes: B41, C61, C63, C73, D83

1. Introduction

The economy is dynamic, and its participants adjust their strategies and decisions to a changing environment. These changes not only affect current market conditions but also trigger responses from other market participants, who in turn adjust their actions to the decisions of competitors, consumers, or business partners. Such sequences of adjustments in strategies and decision rules can have a significant effect on market dynamics and on the structure of entire industries. A useful example is American Airlines' introduction of its *Value Pricing* strategy in 1992, which triggered a broad price war and successive rounds of responses by airlines to changes in their competitors' fares, with substantial effects on the industry as a whole (Morrison & Winston, 1996). In this paper, we use the term *co-adaptation* to describe this interdependence between changing strategies and adaptation processes (Savit et al., 2013).

A large part of economic theory analyzes interactions between agents in terms of equilibrium, focusing on states in which their actions, strategies, or expectations are mutually consistent (Arthur, 2006; Farmer & Geanakoplos, 2009). This approach makes it more difficult to capture the adjustment process itself when agents change their rules of behavior in response to others and their subsequent responses alter the conditions for future decisions (Gallegati & Kirman, 2012; A. Kirman, 1989, 1992). Adaptation processes and interactions between agents are, of course, present in many areas of economics, from game theory and macroeconomics with adaptive learning to complexity economics and agent-based economics (Arthur, 2021; Evans & Honkapohja, 2001; Fudenberg & Levine, 1998; Marcet & Sargent, 1989; Tesfatsion, 2003). However, interdependencies between the adaptation processes of different agents are less often treated as a central subject of economic analysis at a general level.

By contrast, a closely related form of interdependence arises naturally in multi-agent reinforcement learning (MARL), where multiple agents update their policies simultaneously. A change in one agent's decision rule alters the observations, rewards, or states experienced by other agents, while their subsequent updates change the environment again, generating the non-stationarity characteristic of MARL (Foerster et al., 2017; Hernandez-Leal et al., 2017). Coupled learning processes can lead not only to convergence but also to cycles, oscillations, and more complex trajectories (Barfuss, 2022; Sato & Crutchfield, 2003). What is often treated as a difficulty of the learning process from the perspective of MARL may, from an economic perspective, be part of the endogenous dynamics of the economy itself.

In this paper, we focus on interdependencies between the adaptation processes of different economic agents. The behavior of one agent can change the experience used by another agent to update its future decision rule. We propose a formal representation that separates the continuation of an agent's own rule updates from the effects that one agent's update has on subsequent updates by others. This allows us to analyze how the structure of these interdependencies affects the way an update's effects propagate across agents, feed back to the originating agent, accumulate over time, and shape the duration of the adjustment process. We show that interdependencies between adaptation processes can change the speed and course of economic adjustment. A change in one agent's rule can affect the subsequent decisions of other agents, whose responses can then feed back to the agent that initiated the change. As a result, the effects of the initial change can persist longer, spread across agents, and, in more complex systems, affect the market as a whole. We illustrate these results with a model of price competition between two firms and indicate how MARL can be used to study richer settings with many heterogeneous and adapting agents.

Section 2 develops the concept of co-adaptation, relates it to MARL, and places it in the economic literature. Section 3 introduces a formal representation of interdependencies between adaptation processes, and Section 4 analyzes its consequences for two mutually interacting agents. Section 5 presents a simple model illustrating this phenomenon, while Section 6 discusses the implications of co-adaptation for modeling an economy with many adapting agents, the role of MARL, and issues of attribution, validation, and inference.

2. Co-adaptation in Economics: Representation and Modeling

2.1. Interaction, Adaptation, and Co-adaptation

Firms and consumers operate in a constantly changing environment and must adjust to market conditions, technological change, regulation, and demand and supply shocks. Managers make numerous day-to-day decisions that can affect the economic environment in which their firms operate. From time to time, however, external or internal conditions require strategic decisions that shape firms over the longer run. These decisions can also have important effects on other market participants—competitors, business partners, employees, or customers—who may respond to substantial changes in a firm's activity by adjusting their own behavior. The strategies and decision rules of different market participants are connected through a continuously evolving network of interdependencies. Such sequences of response and adjustment are common in markets, but they are also difficult to model economically because they do not fit neatly into an equilibrium-based framework.

The differences between types of decisions are not always clear, and it can be difficult to draw a strict boundary between a standard response to anticipated circumstances and a strategic change of direction. For the purposes of this paper, we define several categories of such decisions in order to conduct a more precise methodological analysis. An *interaction* occurs when the action of one market participant affects another. It is the broadest concept and includes any transaction, price adjustment, or agreement between agents. *Adaptation* means a change in a belief, forecasting rule, strategy, or another rule that affects an agent's future behavior as a result of experience. *Co-adaptation* occurs when the behavior of one agent affects the experience on the basis of which another agent changes the rules governing its future behavior. In computational models, such a process can occur both during the trajectory of the modeled economy itself and during agent training, if experience generated by the behavior of some agents affects how other agents update their rules. Co-adaptation therefore requires two elements: economic interaction, through which one agent's behavior affects another agent's experience, and a change in the latter agent's decision rule that alters future behavior.

This dependence between agents can take different forms. In the simplest case it is one-way: the behavior of agent 1 affects the adaptation of agent 2, without an analogous effect in the opposite direction. Dependencies can also form chains, in which a change in one agent affects the adaptation of another and then further agents. With more agents, a directed network of links between adaptation processes can emerge. A special case is mutual dependence, in which the effect can return, after successive adjustments, to the agent from which the change originated, for example $1 \to 2 \to 1$.

Importantly, the channel of co-adaptation does not require direct observation of other agents' behavioral rules. A firm's action can affect another firm indirectly by changing price, demand, sales, profit, resource availability, or market information, and these variables can then affect the beliefs or rules governing the future behavior of other firms. A similar dependence can operate through asset prices, aggregate variables, observed decisions of other people, or other elements of the shared economic environment. The co-adaptation channel can therefore operate directly or indirectly through variables that are themselves endogenous outcomes of market participants' actions.

2.2. Co-adaptation in Economics

Mechanisms consistent with this definition have been studied in many areas of economics. They differ mainly in the object being adapted and in the channel through which the actions of other participants affect that adaptation. In microeconomics, these include beliefs and strategies updated through strategic interaction, experimentation, or observation of other participants. In macroeconomics, forecasting rules are updated on the basis of aggregate variables jointly generated by the decisions of economic agents. In financial markets, decision-making strategies change, while endogenous prices serve as the transmission channel. The common feature of the cases considered here is that the experience used by one agent to change its future behavior depends on the behavior of other participants in the system. We briefly review how this type of interdependence appears in different areas of economics.

One of the more natural settings for such a mechanism is game theory and behavioral economics, more specifically models of learning in games. In each period, players choose strategies on the basis of changing beliefs and use the history of previous interactions to do so. Reinforcement learning is one tool for adaptive behavior: participants learn which decisions are most beneficial through trial and error. Other models, such as experience-weighted attraction, update the values, or *attractions*, assigned to individual strategies, which then affect the probabilities of future choices (Camerer & Ho, 1999; Fudenberg & Levine, 1998). Because a player's payoffs and observations depend on the actions of opponents, the opponents' behavior affects the experience from which the player learns and, consequently, the player's subsequent choices. The adaptation of each participant is therefore embedded in the history of the game.

A similar structure appears in the literature on strategic experimentation. A participant's action matters not only because of its current payoff but also because it generates additional information about the environment. If this information is available to other participants, one agent's choices change the information available to others and thereby affect their beliefs and future decisions. Bolton and Harris (1999) study agents who learn from the outcomes of experimentation in a multi-agent version of the two-armed bandit problem, where decisions to acquire information are interdependent because agents can use information generated by others. From the perspective of co-adaptation, the relevant channel is simple: the action of one agent changes the information available to others, and updating beliefs on the basis of that information changes their subsequent behavior.

In industrial organization, the link between action, information, and subsequent behavior appears in models of dynamic competition with learning. Aghion et al. (1993) analyze a duopoly in which firms learn about demand from observed sales and prices affect the informativeness of market outcomes. Bergemann and Välimäki (1996) study a market in which a consumer learns the value of products offered by successive sellers by trying them. Sellers compete in prices, and price affects the consumer's willingness to experiment: a lower price can encourage the consumer to buy a product, generate new information about its value, and thereby change subsequent choices. In Bergemann and Välimäki (1997), learning takes place on both sides of the market. Buyers and sellers acquire information about the value of a new product through successive transactions, while prices affect sales and therefore the rate at which information is accumulated. The resulting knowledge changes later consumer decisions, prices, and market shares. In such models, current market decisions therefore help generate the experience on which participants base later decisions. Price affects not only current demand but also what information is generated in a given period and how that information changes the subsequent behavior of market participants.

The structure of relationships between participants is even more explicit in the literature on social learning. Individuals update beliefs on the basis of observed decisions, opinions, or outcomes of others, while the structure of contacts determines the direction and reach of information flows. Models of observational learning and information cascades show that earlier decisions can affect later choices through the information they reveal about participants' private signals (Bikhchandani et al., 1998). In network models, beliefs are updated using information from neighbors, so an agent's position in the network affects the process of information aggregation (Bala & Goyal, 1998; Golub & Jackson, 2010). This mechanism has also been tested empirically. Conley and Udry (2010) examine the spread of technology among pineapple farmers in Ghana and show that the experience of some farmers had a significant effect on the adjustment and subsequent decisions of others.

In macroeconomics, feedback between adaptation and endogenously generated experience often operates through aggregate variables. Marcet and Sargent (1989) study self-referential models in which agents do not know the true law of motion of the economy in advance and estimate its parameters from observed data using least squares. They use the estimated rule to form expectations, and decisions based on these expectations affect the subsequent dynamics of economic variables. This creates a closed loop: past data shape agents' beliefs, beliefs affect their decisions, decisions help generate subsequent outcomes of the economy, and these outcomes become new observations used to update beliefs. Marcet and Sargent study the conditions under which the law of motion perceived by agents and the actual law of motion of the economy can become consistent. Evans and Honkapohja (2001) develop this logic within a broader theory of adaptive learning. Households and firms form forecasts from observed data and then modify their forecasting rules in response to emerging errors. Because expectations affect consumption, investment, prices, and wages, among other variables, changes in forecasting rules also affect the data on which subsequent learning is based. In such a system, the channel linking the adaptation processes of individual participants can operate through inflation, output, prices, or other aggregate variables jointly generated by their decisions. A particularly clear example of dependence between different economic agents is provided by Evans and Honkapohja (2003) in their analysis of monetary policy with adaptive expectations in the private sector. They show that a monetary policy rule based only on fundamentals, which would be optimal under rational expectations, can lead to instability when households and firms learn adaptively. Stability can be restored when the central bank observes private-sector expectations and incorporates them directly into its policy rule. The authors extend this result to the case in which both the private sector and the policymaker learn at the same time. This creates a direct interdependence between two adaptation processes: the expectations of households and firms affect central-bank decisions, monetary policy changes the path of inflation and output, and new realizations of these variables affect subsequent updates of private-sector expectations.

In financial markets, models of heterogeneous and adaptive beliefs provide a useful example. In the model of Brock and Hommes (1998), participants can use different forecasting rules, their decisions affect the asset price, and the resulting price determines the subsequent performance of the rules and future decisions. Price is therefore both an outcome of participants' actions and part of the environment that shapes further adaptation. The Santa Fe Artificial Stock Market develops a similar logic: individual expectation models are adapted to the outcomes of a market jointly generated by agents using those models (Arthur et al., 1997). The dependence between adaptation processes is transmitted indirectly through the common market price.

Agent-based computational economics (ACE) is a natural approach to modeling interdependent and adapting economic processes. It allows these mechanisms to be represented within a single model containing many heterogeneous agents. In this approach, the economy is a system of autonomous agents whose actions jointly generate successive states of the system, while the state of each agent can change as a result of previous interactions with other participants and the environment (Tesfatsion, 2002, 2003). Agents' behavioral rules can take different forms: they can be specified directly by the researcher, respond to other agents, or be updated on the basis of past experience. In A. P. Kirman and Vriend (2001), buyers and sellers learn from the outcomes of previous interactions, and the market relationships that emerge affect their subsequent experience and choices. In the artificial stock market of Arthur et al. (1997), an analogous dependence operates through the endogenously determined price. ACE therefore provides a direct representation of an economy in which each participant adapts in an environment jointly shaped by the actions of others.

Different strands of the literature describe these dependencies using concepts specific to the mechanism under study: learning in games, strategic experimentation, social learning, adaptive learning, adaptive beliefs, imitation, or co-evolution. The term co-adaptation has also been used previously for systems of adapting agents, including by Klos (1999) and Savit et al. (2013). Savit, Riolo, and Riolo further show that dependencies between agents can be transmitted entirely indirectly through a shared environment. In this paper, the concept of co-adaptation is used to isolate a common feature of the mechanisms described above: a dependence in which the behavior or rule change of one participant affects how another participant updates the rules governing its future behavior. ACE allows such dependencies to be represented directly at the level of individual agents. A central question remains, however: how should the adaptation process itself be specified—that is, the rule by which experience changes a participant's future behavior?

2.3. Agent-Based Computational Economics and Reinforcement Learning

In an agent-based model, the choice of adaptation mechanism is crucial because it determines how a participant's experience translates into subsequent decisions. A different updating rule can produce a different response to the same market outcomes and, when many participants adapt simultaneously, different dynamics for the system as a whole. In simple models, this mechanism can be specified directly, for example through a fixed rule for updating beliefs or by changing selected strategy parameters. In more complex environments, it is harder to specify in advance how an agent should adjust its behavior across a large number of states, actions, and possible experiences. One way to model such adaptation is reinforcement learning (RL), which allows agents to learn from experience. An agent repeatedly makes decisions, observes their consequences, and gradually changes how it chooses subsequent actions. The modeler does not have to specify in advance which decision is appropriate in every possible state. Instead, the agent learns from observed outcomes which actions are beneficial under given conditions. A rule that maps observed states into action probabilities or choices is called a policy. Learning therefore consists of changing the policy on the basis of successive experiences.

In ACE, reinforcement learning can therefore serve as the adaptation rule for an individual participant. Policy parameters can then play a role analogous to beliefs, forecasting rules, or heuristics in the models discussed above: they determine how the agent will behave in the future and can change with experience. When many participants in the modeled economy use RL to make decisions and update their policies during a single simulation run, their learning processes become interdependent. Such a system is naturally described as multi-agent reinforcement learning (MARL). In a multi-agent environment, part of each agent's experience is generated by the actions of other participants. A change in agent 1's policy changes its actions and, through the shared environment, can change the rewards or states experienced by agent 2. The data used by agent 2 to update its policy therefore depend on the current behavioral rule of agent 1. If agent 2 is also learning, its changed policy in turn affects the environment in which other participants continue to learn. From the perspective of an individual agent, this means that its environment changes even when the external rules of the model remain fixed. If other participants update their policies and adjust their behavior, this can change the dynamics of the environment faced by the agent. In the MARL literature, this is described as environmental non-stationarity induced by the learning of other agents (Foerster et al., 2017; Hernandez-Leal et al., 2017). From the perspective of the system as a whole, the learning processes are interdependent: an update by one participant changes the experience on the basis of which others update. Sato and Crutchfield (2003) show that such dependence can generate joint learning dynamics even when agents do not observe one another's internal states or rules and interact only through a shared environment.

For the economic interpretation of this process, the design of the learning simulation also matters. If experience accumulated up to period t changes the policy under which an agent makes decisions in subsequent periods, learning is part of the dynamics of the modeled economy and agents can continue to respond to a changing environment. If RL is used only to determine strategies beforehand and agents' policies remain fixed along the trajectory being analyzed, co-adaptation may occur during joint training, but training then serves as a solution procedure rather than as part of the modeled economic dynamics. The analysis below focuses on the first case, because only then can subsequent economic experience change participants' behavioral rules and, through them, affect the further evolution of the system. Learning of this kind already appears in economic applications of RL. Tesauro and Kephart (2002) study competing sellers that use Q-learning to adjust pricing decisions to the outcomes of successive market interactions. Calvano et al. (2020) use Q-learning in repeated price competition, where each firm learns from the profits generated by prices chosen simultaneously by its competitor. In the AI Economist, both economic agents and the planner update their strategies using deep reinforcement learning, and the authors refer to their parallel learning process as co-adaptation (Zheng et al., 2022). These models show that RL can represent adaptation even when a behavioral rule depends on many state variables and cannot be conveniently written as a simple analytical rule. With many learning participants, this leads directly to the question of how interdependencies between their adaptation processes change economic dynamics.

2.4. Economic Significance of Co-adaptation

Interdependence between adaptation processes can make the effects of a single event persist much longer than the immediate response to that event. If a competitor lowers its price, a firm may respond with a one-time change in its own price and return to its previous behavior in the next period. If, however, the experience generated by the price cut changes the firm's beliefs, target markup, or pricing rule, the competitor's influence becomes embedded in the way the firm makes future decisions. The changed rule continues to affect the market in subsequent periods, and the decisions it generates can in turn affect the adaptation processes of other participants. In such an economy, variables governing agents' future behavior become part of the endogenous dynamic state: their evolution affects the paths of prices, output, investment, and other economic variables. Co-adaptation adds dependencies between changes in the rules of different participants.

The direction of these dependencies determines how changes are transmitted. In a one-way system, the adjustment of one agent can change the future behavior of others without substantial feedback. In a chain, the change passes through successive participants, while in a network its reach depends on the pattern of links between them. If there is a cycle, the effect can return, after successive updates, to the agent from which the change originated. The structure of dependencies between rules therefore becomes a separate element of system dynamics. In the remainder of the paper, we introduce a formal representation that separates the dynamics of each participant's own rule from the effect of changes in other participants' rules on its subsequent adjustments. We then analyze the simplest case of two mutually connected agents. This setting is sufficient to show that interdependence between adaptation processes can alter the persistence of an initial disturbance, propagate its effects across participants, and change both the path and the cumulative magnitude of the resulting adjustments. Section 3 introduces the formal representation of these dependencies.

3. Formalization

Section 2 defined co-adaptation. To analyze it further, we need a formal framework that separates the elements affecting an agent's decision process. This will allow us to determine how an initial rule change propagates through interdependent adaptation processes. Our main interest is the path of a change once it has been initiated: how it propagates across agents, how large a cumulative adjustment it generates, and how quickly the induced sequence of updates dies out. We begin by introducing the basic elements—a rule, an action, and an adjustment mechanism. We then distinguish between the way an earlier rule change can lead to another change by the same agent and the way it can trigger adjustments by other participants. Finally, we show how the structure of change across agents can be represented in this setting.

3.1. Decision Rule, Action, and Adjustment

Consider an economy with N agents. Let s_t denote the state of the economy in period t. For agent i, the variable

$$\theta_{i,t}$$

represents an element of the decision rule that can be updated by the agent. For a firm, this may be a parameter determining how strongly it adjusts its price to changes in costs or demand. In an RL model, $\theta_{i,t}$ may represent a parameter of the policy used by the agent to choose subsequent actions.

Let $a_{i,t}$ denote the action taken by agent i in period t. For a firm, this may be the price set in that period. The choice of a particular action depends on the rule $\theta_{i,t}$ and on the information available to the agent:

$$a_{i,t} = g_i(o_{i,t}; \theta_{i,t}),$$

where $o_{i,t}$ denotes the information available to agent i when the decision is made. The same decision rule can lead to different actions when the environment and the information to which the agent responds change. A firm may change its price because demand or costs have changed while the pricing rule itself remains unchanged. A change in the current action alone therefore does not constitute adaptation.

The actions of all agents in period t change the state of the economy:

$$s_{t+1} = F(s_t, a_t), \qquad a_t = (a_{1,t}, \ldots, a_{N,t}).$$

After this transition, each agent receives feedback from the environment, denoted by

$$z_{i,t} = Z_i(s_t, a_t, s_{t+1}).$$

For a firm, this may be sales, profit, or an error in an earlier forecast. If the agent uses this information to change its decision rule, we write

$$\theta_{i,t+1} = L_i(\theta_{i,t}, z_{i,t}).$$

The function L_i specifies the updating mechanism. It may be a simple adjustment rule or a complex learning algorithm. Adaptation occurs when feedback from the environment changes the rule used by the agent in subsequent periods. Co-adaptation occurs when the adaptation of one agent depends on the behavior of another participant—that is, when that behavior affects the feedback used to update the rule. Suppose, for example, that firm 1 changes its pricing rule and consequently sets a different price. This price affects the outcomes of firm 2. If firm 2 uses this information to update its own pricing rule, its adaptation depends on the behavior of firm 1. The mechanism can be represented as

$$\theta_{1,t} \longrightarrow a_{1,t} \longrightarrow z_{2,t} \longrightarrow \theta_{2,t+1}.$$

Firm 2 does not need to observe firm 1's rule directly. The effect can operate through changes in prices, demand, or other economic variables that firm 2 observes as feedback. What matters is that the behavior of one participant changes the environment in which another updates its future rule.

3.2. Adaptation over Time

To analyze the adjustment process over time, we distinguish between the level of a decision rule and its change in a given period. Define

$$\Delta\theta_{i,t} = \theta_{i,t} - \theta_{i,t-1}.$$

Thus, $\Delta\theta_{i,t}$ measures how much agent i changed its rule between periods $t-1$ and t. We are interested in the implications of such a change for the system. We therefore isolate the subsequent path of a single update that has already occurred. This does not mean that all future rule changes in the economy must result solely from earlier $\Delta\theta$. New information, shocks, and changes in the state of the economy may also arise. Here, however, we focus on one specific mechanism: how an earlier update carries over into subsequent periods.

To analyze this mechanism, we assume a linear relationship between successive rule updates and abstract from new independent impulses arising in later periods. First consider how an earlier rule update affects the agent's next update. If an earlier change in agent i's rule affects its next change, we denote the strength of this relationship by D_i. Temporarily leaving aside the effect of other participants, we write

$$\Delta\theta_{i,t+1} = D_i\Delta\theta_{i,t}.$$

In this linear representation, D_i is a coefficient describing how strongly the current rule change affects the next one. Equivalently,

$$D_i = \frac{\partial\, \Delta\theta_{i,t+1}}{\partial\, \Delta\theta_{i,t}}.$$

D_i therefore describes how a change in agent i's rule in one round affects its next update.

If $0 < D_i < 1$, successive changes have the same direction but become progressively smaller. If $D_i = 0$, an earlier update does not by itself lead to another one. A negative D_i means that the next correction moves in the opposite direction.

If successive positive updates become progressively smaller, the agent may gradually converge to a new stable level.

A change in one agent's rule can also trigger an update of another participant's rule. For $i \neq j$, we denote the strength of this relationship by C_{ij}. In the linear representation, the term

$$C_{ij}\Delta\theta_{j,t}$$

describes the effect of a change in agent j's rule on the next change in agent i's rule. Equivalently,

$$C_{ij} = \frac{\partial\, \Delta\theta_{i,t+1}}{\partial\, \Delta\theta_{j,t}}, \qquad i \neq j.$$

C_{ij} describes how a change in agent j's rule affects the next change in agent i's rule. The first index identifies the agent whose rule is updated, and the second identifies the agent from which the effect originates. For example, C_{21} denotes the direction $1 \longrightarrow 2$. The analysis below focuses on a specific channel of co-adaptation: the propagation of the effects of earlier rule updates across participants. The coefficient C_{ij} measures the strength with which an earlier update of agent j's rule affects the next update of agent i's rule.

The coefficient C_{ij} does not assume a direct effect of one rule on another. A change in agent j's rule may first change its action, then the outcome observed by agent i, and only then affect the next update of agent i's rule:

$$\Delta\theta_{j,t} \longrightarrow a_{j,t} \longrightarrow z_{i,t} \longrightarrow \Delta\theta_{i,t+1}.$$

If only agent 1 changed its rule in period t,

$$\Delta\theta_{1,t} \neq 0, \qquad \Delta\theta_{2,t} = 0,$$

then agent 1's earlier update can lead, after one round, both to another change in its own rule,

$$D_1 \Delta\theta_{1,t},$$

and to a change in agent 2's rule,

$$C_{21} \Delta\theta_{1,t}.$$

The first effect describes the continuation of the agent's own adjustment process. The second is the channel through which an earlier update by one participant triggers adaptation by another.

3.3. Dynamics with Multiple Agents

In the general case, the next change in agent i's rule can result from both its own earlier update and earlier changes in the rules of other participants. Under the assumed linear relationship, we write

$$\Delta\theta_{i,t+1} = D_i \Delta\theta_{i,t} + \sum_{j \neq i} C_{ij} \Delta\theta_{j,t}.$$

The first term describes the continuation of agent i's own adjustment process. The remaining terms describe the effect of changes in other agents' rules on its adjustment.

For N agents, the own-update relationships can be collected in the matrix

$$D = \begin{pmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_N \end{pmatrix},$$

and the cross-agent relationships in the matrix

$$C = \begin{pmatrix} 0 & C_{12} & \cdots & C_{1N} \\ C_{21} & 0 & \cdots & C_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{N1} & C_{N2} & \cdots & 0 \end{pmatrix}.$$

Collecting the updates of all rules in the vector

$$\Delta\Theta_t = \begin{pmatrix} \Delta\theta_{1,t} \\ \vdots \\ \Delta\theta_{N,t} \end{pmatrix},$$

we obtain

$$\boxed{\Delta\Theta_{t+1} = (D + C)\Delta\Theta_t.}$$

This equation describes the subsequent path of updates in the system as a whole. D determines how earlier changes by individual agents affect their own subsequent changes. C determines how earlier changes by some participants trigger subsequent changes in others.

The rules themselves evolve by accumulating these individual changes:

$$\Theta_{t+1} = \Theta_t + \Delta\Theta_{t+1}.$$

The linear representation is used to analyze the propagation of a change that has already been initiated. In a more general model, the strength of the relationships between successive updates may depend on the state of the economy and vary over time. The analysis below focuses on the simpler case in which the coefficients D_i and C_{ij} can be treated as constant.

The equation

$$\Delta\Theta_{t+1} = (D + C)\Delta\Theta_t$$

describes the isolated mechanism of update propagation. If other variables independently affect future updates, they should be included in a complete description of the dynamics. The $D + C$ representation is used here to isolate the dependence between successive rule changes. It is deliberately simplified: we trace how earlier rule updates generate subsequent updates and leave other sources of economic dynamics outside this representation.

3.4. Propagation Structure

The fact that rule changes can affect other participants does not by itself determine the exact path of this effect. Dependencies can be one-way or mutual, and with more agents they can form longer paths. The structure of the matrix C makes it possible to distinguish these cases and identify the paths through which an earlier update by one participant can lead to subsequent updates in the system.

If

$$C_{21} \neq 0, \qquad C_{12} = 0,$$

a change in agent 1's rule can trigger a subsequent change in agent 2's rule, with no corresponding effect in the opposite direction. The dependence then has the direction

$$1 \rightarrow 2.$$

With more participants, the effect can pass through successive adaptation processes. If

$$C_{21} \neq 0, \qquad C_{32} \neq 0,$$

a change in agent 1's rule can first affect the update of agent 2, and the change in agent 2's rule can then affect the update of agent 3:

$$1 \rightarrow 2 \rightarrow 3.$$

A mutual dependence is a special case. If

$$C_{21} \neq 0, \qquad C_{12} \neq 0,$$

a change initiated by one agent can return to that agent after successive rounds through the adaptation process of the other:

$$1 \rightarrow 2 \rightarrow 1.$$

The mutual case is particularly interesting because an earlier change by one participant can indirectly become a source of its own subsequent update.

To separate cross-agent effects from the continuation of agents' own updates, we also consider the case

$$C = 0.$$

Then

$$\Delta\Theta_{t+1} = D\Delta\Theta_t.$$

In this case, an earlier update can still lead to subsequent changes in the same agent's rule, but it no longer triggers further updates by other participants through the channels represented by C. Setting $C = 0$ leaves changes that have already occurred in Θ in place and removes only the subsequent cross-agent effects of those changes. Comparing the dynamics under $D + C$ with those under D therefore makes it possible to determine which features of the dynamics in the system under analysis arise from interdependence between adaptation processes.

In the analysis below, we assume a synchronous sequence of events. Agents begin period t with rules Θ_t and use them to choose actions a_t. These actions affect the state of the economy and agents' experience, after which the agents update their rules:

$$\Theta_t \to a_t \to s_{t+1}, z_t \to \Theta_{t+1}.$$

The index t denotes successive periods of the modeled economy.

This representation allows us to track separately subsequent rule updates and their cumulative effect on agents' behavior. The next section uses the case of two agents that influence one another to show how such cross-agent adjustments jointly determine the duration of the process, the path of change, and the possibility that the effect returns to its source.

4. Effect of Co-adaptation on Dynamics

4.1. A Coupled System of Two Agents

Section 3 introduced a formal representation that allows us to analyze the consequences of co-adaptation. We now consider the simplest case: two agents that affect each other to the same degree. We analyze how this mutual influence changes their subsequent rule-update dynamics.

We assume

$$\Delta\theta_{1,t+1} = d\Delta\theta_{1,t} + c\Delta\theta_{2,t},$$

$$\Delta\theta_{2,t+1} = c\Delta\theta_{1,t} + d\Delta\theta_{2,t}.$$

The parameter d describes the effect of an agent's earlier update on its own rule, while c describes the effect of the other agent's earlier update on the agent's next rule update. We assume the same values of d and c for both agents.

As a benchmark, we consider the case

$$c = 0.$$

Then

$$\Delta\theta^0_{i,t+1} = d\Delta\theta^0_{i,t}.$$

Rule changes then no longer induce updates in the other agent, and each agent continues only its own adjustment process. Comparing the coupled system with this case shows what the influence of the other agent adds to the subsequent dynamics.

4.2. Transmission and Return of the Effect

We begin with the first few iterations of the system after an initial change of size ξ affecting only the decision rule of agent 1:

$$\Delta\Theta_0 = \begin{pmatrix} \xi \\ 0 \end{pmatrix}.$$

After one round,

$$\Delta\Theta_1 = \begin{pmatrix} d\xi \\ c\xi \end{pmatrix}.$$

Agent 1's rule changes further by $d\xi$, as a result of its own earlier update. Agent 2's rule changes by $c\xi$, triggered by the earlier update of agent 1.

After another round,

$$\Delta\Theta_2 = \begin{pmatrix} \left(d^2 + c^2\right)\xi \\ 2dc\xi \end{pmatrix}.$$

The next update of agent 1 now contains two components:

$$d^2\xi + c^2\xi.$$

The first comes from its own earlier change. The second has passed through agent 2:

$$1 \to 2 \to 1.$$

The term $c^2\xi$ therefore identifies the part of agent 1's later change that returns to it through the response of the other agent.

This is easy to see in the example of two firms. Firm 1 changes its pricing rule. This affects firm 2, which changes its own rule. The new behavior of firm 2 then affects the next change by firm 1. Part of this later change by firm 1 therefore originates in its own earlier rule change but returns through the response of its competitor.

A particularly simple case is

$$d = 0.$$

Then

$$(\xi, 0) \to (0, c\xi) \to \left(c^2\xi, 0\right) \to \left(0, c^3\xi\right) \to \cdots .$$

Neither agent carries its own previous update directly into the next period. The process can nevertheless continue because each agent's update induces a subsequent update by the other. Further adaptation can therefore be sustained entirely through mutual influence.

The first return to the initial agent is equal to $c^2\xi$. If the process continues, there may be many such returns. This leads to the next question: what is the total effect of all subsequent updates triggered by the initial change ξ?

4.3. Cumulative Effect of Mutual Adaptation

Let

$$A = \begin{pmatrix} d & c \\ c & d \end{pmatrix}.$$

Successive updates are

$$\Delta\Theta_0, \quad A\Delta\Theta_0, \quad A^2\Delta\Theta_0, \quad \ldots .$$

We assume

$$d \geq 0, \qquad c \geq 0, \qquad d + c < 1.$$

Under these conditions, successive updates eventually die out. We can therefore sum all changes generated by the initial ξ:

$$S = \sum_{k=0}^{\infty} A^k \Delta\Theta_0 = (I - A)^{-1}\Delta\Theta_0.$$

Thus, S gives the total change in the rules of both agents over the entire sequence of updates initiated by ξ.

For

$$\Delta\Theta_0 = \begin{pmatrix} \xi \\ 0 \end{pmatrix}$$

we obtain

$$S_1 = \frac{1-d}{(1-d)^2 - c^2}\xi,$$

$$S_2 = \frac{c}{(1-d)^2 - c^2}\xi.$$

Here, S_1 is the total change in agent 1's rule and S_2 is the total change in agent 2's rule. Their ratio is

$$\frac{S_2}{S_1} = \frac{c}{1-d}.$$

This ratio shows the total change in agent 2's rule relative to the total change in agent 1's rule and depends jointly on two factors. It increases with c, but it is also larger when earlier updates die out more slowly, that is, when d is higher. The longer the adjustment process lasts, the more rounds there are in which the effect can pass between agents.

Even more importantly, consider what happens to agent 1. When $c = 0$, its total change is

$$S_1^0 = \frac{\xi}{1-d}.$$

With mutual influence, part of the change passes to agent 2 and then returns to agent 1. The fraction of agent 1's cumulative adjustment generated through this cross-agent feedback channel is

$$\boxed{\frac{S_1 - S_1^0}{S_1} = \left(\frac{c}{1-d}\right)^2}.$$

This expression therefore isolates the share of the cumulative adjustment at the source that depends on feedback through the second agent. It links the simple path $1 \to 2 \to 1$ to the full sequence of subsequent returns.

We can also calculate the total change of both agents:

$$\boxed{S_1 + S_2 = \frac{\xi}{1-d-c}}.$$

When $c = 0$, this is

$$\frac{\xi}{1-d}.$$

The ratio of these two quantities is

$$\boxed{M = \frac{1-d}{1-d-c}}.$$

M measures the factor by which mutual influence increases the total change in decision rules generated by the same initial disturbance. For positive c, successive responses of the agents make the effect of the initial ξ larger than when each agent continues only its own updates.

Consider a numerical example with

$$d = 0.5, \qquad c = 0.4, \qquad \xi = 1.$$

Without cross-agent influence on rule updates,

$$S_1^0 = 2.$$

With mutual influence,

$$S_1 \approx 5.56, \qquad S_2 \approx 4.44.$$

Thus,

$$S_1 + S_2 = 10.$$

The same initial update therefore leads to a total change of 2 when $c = 0$ and 10 when $c = 0.4$. Although the initial ξ is the same in both cases, when the agents are interdependent and adjust their decision rules in response to one another, the total effect of the disturbance is five times larger.

In the same example,

$$\frac{S_2}{S_1} = 0.8.$$

The total change in agent 2 therefore reaches 80% of the total change in agent 1, and

$$\frac{S_1 - S_1^0}{S_1} = 0.64.$$

Thus, when the cross-agent adjustment channel is removed, the cumulative change in agent 1 falls by 64%. In this simple system, this quantifies how much of agent 1's cumulative adjustment is generated by responses that its own earlier change triggers in agent 2 and that then feed back to agent 1.

4.4. Duration of the Adjustment Process

The cumulative effect tells us the total size of the change. A separate question concerns time: how long the initial change continues to generate subsequent rule updates. Because the previous subsections followed a single initial disturbance, we can now examine what fraction of the updates it generated remains in the system after each successive round. To do this, we sum the updates of both agents:

$$R_t = \Delta\theta_{1,t} + \Delta\theta_{2,t}.$$

Summing the two update equations gives

$$R_{t+1} = (d + c)R_t.$$

Here, d captures the continuation of earlier changes by the agents themselves, while c captures their mutual influence. Thus, $d + c$ determines what fraction of the total update carries over into the next round.

For the initial disturbance from Section 4.2,

$$R_0 = \xi,$$

and after k rounds,

$$R_k = (d + c)^k \xi.$$

When $c = 0$, changes are not transmitted between agents, so

$$R_k^0 = d^k \xi.$$

Mutual influence therefore changes the rate at which the induced updates die out. For positive

c, a larger share of the previous update carries over into the next round, so the process can last longer than when each agent continues only its own adjustment.

We can directly calculate the number of periods needed for the effect to fall below a given fraction ε of its initial value. For $0 < d + c < 1$, we must have

$$k \geq \frac{\ln \varepsilon}{\ln(d + c)}.$$

Let T_ε denote the smallest integer satisfying this inequality. For the case without mutual influence, we analogously obtain

$$k \geq \frac{\ln \varepsilon}{\ln d}, \qquad 0 < d < 1,$$

and denote the corresponding number of periods by T_ε^0. For example, $\varepsilon = 0.1$ corresponds to the point at which less than 10% of the initial value remains.

Returning to the earlier example,

$$d = 0.5, \qquad c = 0.4.$$

Then

$$d + c = 0.9.$$

For a 10% threshold,

$$T_{0.1} = 22$$

periods. When $c = 0$, the coefficient is $d = 0.5$, so the same threshold is reached after

$$T_{0.1}^0 = 4$$

periods.

Mutual influence on the agents' decision rules therefore changes the duration of the induced adjustment process substantially. In this example, the effect of the initial change falls below 10% after about 4 rounds without mutual influence and only after about 22 rounds when $c = 0.4$. Mutual adjustment therefore not only increases the cumulative effect of the disturbance shown in Section 4.3 but also extends the period during which it continues to generate further updates.

4.5. Conclusions from the Two-Agent Case

In summary, co-adaptation can have a substantial effect on adjustment dynamics, both on the cumulative magnitude of adjustment and on its persistence. The analysis considers a very simple case of two agents that influence one another, isolating the effects of a single initial change while abstracting from other sources of dynamics in the environment. Even this simple case shows that interdependence between adaptation processes can substantially amplify an initial disturbance that would otherwise decay quickly when the cross-agent adjustment channel is absent. Section 5 examines the same problem in a model of two competing firms in which d and c follow from a specific rule of price adjustment.

5. Co-adaptation in Dynamic Price Adjustments of Two Firms

5.1. A Simple Model of Price Adjustment

Section 4 showed how interdependence between the updates of two agents can affect the propagation of an initial disturbance's effects, their cumulative magnitude, and the duration of the adjustment process. We now show how such a structure can arise in a simple model of two competing firms.

We consider a deliberately minimal mechanism of dynamic price adjustment between two symmetric firms. Its purpose is to show how interdependencies between rule updates can arise from a simple, economically interpretable response to a change in a competitor's price.

Let $\bar{p}$ denote a constant reference level for price, which can be interpreted as unit cost. The parameter $\theta_{i,t}$ is part of firm i's pricing rule; it is updated over time and determines the firm's target markup over this level, so that

$$p_{i,t} = \bar{p} + \theta_{i,t}.$$

In this minimal example, a change in the rule parameter under analysis leads directly to an equal change in price:

$$p_{i,t} - p_{i,t-1} = \Delta\theta_{i,t}.$$

This allows us to trace directly the channel from an update of one participant's rule to an action observed by the other.

We assume that a firm responds to the most recent change in its competitor's price. We write the target value of its next rule update as

$$\Delta\theta^*_{i,t+1} = \beta\left(p_{j,t} - p_{j,t-1}\right) = \beta\Delta\theta_{j,t}, \qquad 0 < \beta < 1, \qquad j \neq i.$$

The parameter β determines the strength of the response to a change in the competitor's price. A higher value of β means that a larger share of the other firm's earlier price change is reflected in the target value of the firm's next rule update. A positive β means that a price change by one firm induces an adjustment of the competitor's pricing rule in the same direction.

The firm gradually adjusts the size of its next update toward this target value:

$$\Delta\theta_{i,t+1} = \Delta\theta_{i,t} + \alpha\left(\Delta\theta^*_{i,t+1} - \Delta\theta_{i,t}\right), \qquad 0 < \alpha \leq 1.$$

The parameter α determines the speed of adjustment in the updating process. With a higher α, the next rule change moves more strongly toward the response implied by the most recent change in the competitor's price. Substituting the target value gives

$$\boxed{\Delta\theta_{i,t+1} = (1-\alpha)\Delta\theta_{i,t} + \alpha\beta\Delta\theta_{j,t}.}$$

The first term describes the continuation of the firm's own earlier update. The second carries the effect of the competitor's earlier price change into the firm's next update. A change in firm 1's pricing rule changes its price; this change affects the next update of firm 2; and the changed rule of firm 2 then affects its future price and the next update of firm 1. This generates exactly the mutual adjustment channel analyzed in Section 4.

5.2. Connection with the Formalization of Co-adaptation

The dynamics implied by the price-adjustment rule take exactly the form analyzed in Section 4:

$$\boxed{\Delta\theta_{i,t+1} = (1-\alpha)\Delta\theta_{i,t} + \alpha\beta\Delta\theta_{j,t}.}$$

Comparing coefficients gives

$$\boxed{d = 1-\alpha, \qquad c = \alpha\beta.}$$

The two parameters from Section 4 therefore receive a simple economic interpretation. The value $d = 1-\alpha$ determines the continuing effect of the firm's own earlier update. The value $c = \alpha\beta$ determines the strength with which an earlier change in the competitor's price affects the firm's next rule update.

The strength of the cross-firm adjustment channel therefore depends on two elements. The parameter β describes the strength of the response to a change in the competitor's price, while α describes how quickly the updating process adjusts toward that response. With a higher β and a higher α, a larger share of the competitor's earlier change carries over into the firm's next update.

For $0 < \alpha \le 1$ and $0 < \beta < 1$, we also obtain

$$d + c = 1 - \alpha(1 - \beta) < 1,$$

so successive updates die out over time, consistent with the conditions assumed in Section 4.

5.3. Economic Interpretation of the Results

The mapping $d = 1 - \alpha$ and $c = \alpha\beta$ allows us to interpret the results of Section 4 in terms of price adjustment. An initial change in firm 1's rule leads to a change in its price. This affects the next update of firm 2's rule and, through the subsequent change in firm 2's price, can return to firm 1. The path $1 \to 2 \to 1$ therefore describes successive responses through which part of the effect of the initial change returns to its source.

In this model, the ratio of the cumulative change in firm 2's rule to the cumulative change in firm 1's rule simplifies to

$$\frac{S_2}{S_1} = \beta.$$

The parameter β therefore also determines the relative scale of the cumulative adjustment of the two firms. The higher β, the larger the cumulative change in firm 2's rule relative to the cumulative change in firm 1's rule.

The multiplier of the cumulative effect from Section 4 takes a particularly simple form. Substituting the parameters of the price model gives

$$\boxed{M = \frac{1}{1 - \beta}}.$$

In this model, the relative increase in the cumulative effect generated by the mutual adjustment channel depends only on β. As the response to a change in the competitor's price becomes stronger, successive reactions sustain the adjustment process and increase the total change in rules relative to the case without the cross-firm adjustment channel. The parameter α affects the persistence of successive updates and, through it, the absolute size of the cumulative change. The rate at which the process dies out depends on the combination of the two parameters, as shown by the coefficient $d + c$.

The persistence of the adjustment process is governed by

$$\boxed{d + c = 1 - \alpha(1 - \beta)}.$$

This value increases with β, because a stronger response to a change in the competitor's price increases the share of the previous update carried into the next round. For a given $\beta < 1$, a higher α reduces $d + c$ and therefore makes successive updates die out faster. At the same time, a higher α reduces the own-update component $d = 1 - \alpha$. Because $\beta < 1$, the reduction in d is larger than the increase in c, so the combined coefficient $d + c$ falls. Faster adjustment in the updating process moves the next rule change more strongly toward the response implied by the latest change in the competitor's price and leaves a smaller share of the firm's own previous update to carry over into subsequent periods.

The parameter β determines the strength of the mutual amplification of the cumulative effect, while the persistence of the resulting adjustment process depends on the combination of α and β. The price model therefore shows that the abstract parameters d and c can arise from two economically interpretable elements of the behavioral rule: the speed of adjustment in the updating process and the strength of the response to a change in the competitor's price.

The model is deliberately simple. The two firms are symmetric, update their rules synchronously, and the relationship between their price changes is linear. This structure allows us to connect the economic adjustment rule directly to the formalization in Section 3 and the results in Section 4. With more firms, asymmetric responses, and more complex learning processes, the structure of interdependencies between updates can be substantially richer. Section 6 turns to these extensions and to the problems involved in modeling co-adaptation.

6. Modeling in a Complex Environment

The previous sections showed that the structure of interdependencies between the updates of different participants can itself be a source of system dynamics. It affects how the effects of changes propagate across agents, their cumulative magnitude, and the duration of the adjustment process. The price model additionally showed that the strength of these interdependencies can arise from simple, economically interpretable behavioral rules. In the two-agent case, we analyzed this structure directly. With more participants, however, the direction and strength of influence may differ across agents, and the interdependencies between their adaptation processes form a more complex system.

6.1. Multiple Agents, Asymmetry, and Networks of Interdependence

With more participants, the interdependencies between agents' rule updates can be much more complex. A change in one agent's rule may strongly affect the update of another, while the effect in the opposite direction may be weaker. Such asymmetry has a natural economic interpretation:

a large market participant may strongly shape the learning conditions of smaller participants, price changes by one firm may affect suppliers, customers, and competitors, and the decisions of one group may change the data later used by other groups. With many agents, the matrix C from Section 3 can be interpreted as a directed network of interdependencies between updating processes. Its structure determines possible transmission paths, cycles, the concentration of influence, and the position of individual agents in the network. Path length, the overlap of several channels, and differences in the speed of adjustment across participants may also matter.

This leads to a broader question of how the topology of interdependencies between adaptation processes affects economic dynamics. As the number of agents, the dimensionality of their rules, and the number of information channels increase, it also becomes more difficult to determine which interdependencies are responsible for the observed dynamics. Updates may be heterogeneous, stochastic, and asynchronous, while interactions may operate through networks or common markets. Analyzing such systems therefore requires tools that can represent many adapting participants at the same time and trace the interdependencies between their learning processes.

6.2. MARL as a Framework for Co-adaptation

Multi-agent reinforcement learning is a useful framework for representing systems in which many behavioral rules evolve simultaneously. MARL allows separate behavioral rules and updating processes to be represented for individual agents. This makes it possible to trace how a change in one participant's behavior modifies the experience of others and affects their subsequent updates (Barfuss, 2022; Charpentier et al., 2023; Shoham et al., 2007). With more agents, analytically tracing all paths of influence quickly becomes difficult. A multi-agent model makes the following sequence observable: a behavioral rule leads to an action, the action contributes to the experience of other participants, and that experience affects their subsequent rules. This representation makes it possible to trace how one agent's actions alter the experience of others and thereby affect their subsequent updates.

For the economic interpretation, the timing of learning within the model is crucial. Updates that occur during the modeled trajectory change agents' later behavior and become part of the dynamics of the economy. Joint training completed before that trajectory begins serves as a procedure for determining the rules subsequently used in the model. Simulation also allows us to examine whether the mechanism visible in the simple model remains important after its assumptions are relaxed. Computational models can therefore be used to analyze the mechanism and generate hypotheses even when a complete analytical solution is unavailable (Epstein, 2008; Gräbner et al., 2019; Lehtinen & Kuorikoski, 2007).

In complex models, it is important to be able to trace an economically meaningful adjustment path. In MARL, selected information channels can be removed, specific updates can be frozen, or a model with the relevant C_{ij} terms can be compared with a variant in which the corresponding channels are disabled. Such comparisons show how the dynamics change after a controlled modification of the mechanism under study (Gräbner et al., 2019).

6.3. Validation and Limitations

When analyzing co-adaptation, it is useful to distinguish three questions. The first is theoretical: can a particular structure of interdependence between adaptation processes alter system dynamics? The second is empirical: does such a mechanism operate in a particular economic environment? The third concerns its quantitative importance: how much of the observed process is shaped by co-adaptation? This paper focuses primarily on the first question. We show that mutual interdependence between adaptation processes can substantially change the cumulative effect of an initial disturbance and the duration of the adjustment process. Assessing the presence and importance of this mechanism in particular markets requires data and a separate empirical design. Similar outcomes can arise from different model structures. Validation should therefore also cover behavioral rules, learning processes, the information structure, and the mechanism of interaction between agents. As agent-based economics develops, calibration, validation, and empirical grounding of such models become increasingly important (Axtell & Farmer, 2025). In addition, the $C = 0$ variant used in the analysis isolates the channel of interdependence between updates. In empirical applications, interpreting such a comparison requires an additional specification of which elements of the environment, information, and behavior are held fixed.

7. Conclusions

Co-adaptation describes a situation in which the behavior of some participants changes the experience used by others to update their future behavioral rules. The evolution of these rules then becomes part of the dynamics of the economy. Interdependencies between adaptation processes shape how the effects of earlier changes propagate through the system and how long they continue to influence the subsequent decisions of market participants.

The aim of the paper was to isolate this mechanism and show how strongly it can affect adjustment dynamics. The two-agent analysis showed that mutual influence can increase the cumulative effect of an initial change several-fold and substantially prolong the resulting adjustment process. We also showed that these effects can arise even from simple, economically interpretable behavioral rules.

The structure of interdependencies between adaptation processes therefore becomes a key element of the dynamics: who affects the experience of other participants, how this influence passes through successive updates, and how long it continues to affect later decisions. With more participants, these interdependencies can form an extensive network through which the effects of changes in behavioral rules propagate. MARL makes it possible to represent and study such processes even when participants are heterogeneous and their adaptation rules take more complex forms.

References

Aghion, P., Espinosa, M. P., & Jullien, B. (1993). Dynamic duopoly with learning through market experimentation. *Economic Theory*, *3*(3), 517–539. https://doi.org/10.1007/BF01209700

Arthur, W. B. (2006). Out-of-Equilibrium Economics and Agent-Based Modeling. In *Handbook of Computational Economics* (pp. 1551–1564, Vol. 2). Elsevier. https://doi.org/10.1016/S1574-0021(05)02032-0

Arthur, W. B. (2021). Foundations of complexity economics. *Nature Reviews Physics*, *3*(2), 136–145. https://doi.org/10.1038/s42254-020-00273-3

Arthur, W. B., Holland, J. H., LeBaron, B., Palmer, R., & Tayler, P. (1997). Asset Pricing Under Endogenous Expectations in an Artificial Stock Market. In W. B. Arthur, S. N. Durlauf, & D. A. Lane (Eds.), *The Economy as an Evolving Complex System II* (pp. 15–44). Addison-Wesley.

Axtell, R. L., & Farmer, J. D. (2025). Agent-Based Modeling in Economics and Finance: Past, Present, and Future. *Journal of Economic Literature*, *63*(1), 197–287. https://doi.org/10.1257/jel.20221319

Bala, V., & Goyal, S. (1998). Learning from Neighbours. *Review of Economic Studies*, *65*(3), 595–621. https://doi.org/10.1111/1467-937X.00059

Barfuss, W. (2022). Dynamical systems as a level of cognitive analysis of multi-agent learning: Algorithmic foundations of temporal-difference learning dynamics. *Neural Computing and Applications*, *34*(3), 1653–1671. https://doi.org/10.1007/s00521-021-06117-0

Bergemann, D., & Välimäki, J. (1996). Learning and Strategic Pricing. *Econometrica*, *64*(5), 1125–1149. https://doi.org/10.2307/2171959

Bergemann, D., & Välimäki, J. (1997). Market Diffusion with Two-Sided Learning. *The RAND Journal of Economics*, *28*(4), 773–795. https://doi.org/10.2307/2555786

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *Journal of Economic Perspectives*, *12*(3), 151–170. https://doi.org/10.1257/jep.12.3.151

Bolton, P., & Harris, C. (1999). Strategic Experimentation. *Econometrica*, *67*(2), 349–374. https://doi.org/10.1111/1468-0262.00022

Brock, W. A., & Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, *22*(8-9), 1235–1274. https://doi.org/10.1016/S0165-1889(98)00011-6

Calvano, E., Calzolari, G., Denicolò, V., & Pastorello, S. (2020). Artificial Intelligence, Algorithmic Pricing, and Collusion. *American Economic Review*, *110*(10), 3267–3297. https://doi.org/10.1257/aer.20190623

Camerer, C., & Ho, T. H. (1999). Experience-weighted Attraction Learning in Normal Form Games. *Econometrica*, *67*(4), 827–874. https://doi.org/10.1111/1468-0262.00054

Charpentier, A., Élie, R., & Remlinger, C. (2023). Reinforcement Learning in Economics and Finance. *Computational Economics*, *62*(1), 425–462. https://doi.org/10.1007/s10614-021-10119-4

Conley, T. G., & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *American Economic Review*, *100*(1), 35–69. https://doi.org/10.1257/aer.100.1.35

Epstein, J. M. (2008). Why Model? *Journal of Artificial Societies and Social Simulation*, *11*(4), 12.

Evans, G. W., & Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.

Evans, G. W., & Honkapohja, S. (2003). Expectations and the Stability Problem for Optimal Monetary Policies. *Review of Economic Studies*, *70*(4), 807–824. https://doi.org/10.1111/1467-937X.00268

Farmer, J. D., & Geanakoplos, J. (2009). The virtues and vices of equilibrium and the future of financial economics. *Complexity*, *14*(3), 11–38. https://doi.org/10.1002/cplx.20261

Foerster, J., Nardelli, N., Farquhar, G., Afouras, T., Torr, P. H. S., Kohli, P., & Whiteson, S. (2017). Stabilising experience replay for deep multi-agent reinforcement learning. In D. Precup & Y. W. Teh (Eds.), *Proceedings of the 34th international conference on machine learning* (pp. 1146–1155, Vol. 70). PMLR.

Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games*. The MIT Press.

Gallegati, M., & Kirman, A. (2012). Reconstructing economics: Agent based models and complexity. *Complexity Economics*, *1*(1), 5–31. https://doi.org/10.7564/12-COEC2

Golub, B., & Jackson, M. O. (2010). Naïve Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics*, *2*(1), 112–149. https://doi.org/10.1257/mic.2.1.112

Gräbner, C., Bale, C. S. E., Furtado, B. A., Alvarez-Pereira, B., Gentile, J. E., Henderson, H., & Lipari, F. (2019). Getting the Best of Both Worlds? Developing Complementary Equation-Based and Agent-Based Models. *Computational Economics*, *53*(2), 763–782. https://doi.org/10.1007/s10614-017-9763-8

Hernandez-Leal, P., Kaisers, M., Baarslag, T., & Munoz de Cote, E. (2017). A survey of learning in multiagent environments: Dealing with non-stationarity. https://doi.org/10.48550/arXiv.1707.09183

Kirman, A. (1989). The Intrinsic Limits of Modern Economic Theory: The Emperor has No Clothes. *The Economic Journal*, *99*(395), 126–139. https://doi.org/10.2307/2234075

Kirman, A. (1992). Whom or What Does the Representative Individual Represent? *Journal of Economic Perspectives*, *6*(2), 117–136. https://doi.org/10.1257/jep.6.2.117

Kirman, A. P., & Vriend, N. J. (2001). Evolving market structure: An ACE model of price dispersion and loyalty. *Journal of Economic Dynamics and Control*, *25*(3-4), 459–502. https://doi.org/10.1016/S0165-1889(00)00033-6

Klos, T. B. (1999). Decentralized Interaction and Co-Adaptation in the Repeated Prisoner's Dilemma. *Computational and Mathematical Organization Theory*, *5*(2), 147–165.

Lehtinen, A., & Kuorikoski, J. (2007). Computing the Perfect Model: Why Do Economists Shun Simulation? *Philosophy of Science*, *74*(3), 304–329. https://doi.org/10.1086/522359

Marcet, A., & Sargent, T. J. (1989). Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory*, *48*(2), 337–368. https://doi.org/10.1016/0022-0531(89)90032-X

Morrison, S. A., & Winston, C. (1996). Causes and Consequences of Airline Fare Wars. *Brookings Papers on Economic Activity. Microeconomics*, *1996*, 85–131. https://doi.org/10.2307/2534747

Sato, Y., & Crutchfield, J. P. (2003). Coupled replicator equations for the dynamics of learning in multiagent systems. *Physical Review E*, *67*(1), 015206. https://doi.org/10.1103/PhysRevE.67.015206

Savit, R., Riolo, M., & Riolo, R. (2013). Co-Adaptation and the Emergence of Structure (M. Perc, Ed.). *PLoS ONE*, *8*(9), e71828. https://doi.org/10.1371/journal.pone.0071828

Shoham, Y., Powers, R., & Grenager, T. (2007). If multi-agent learning is the answer, what is the question? *Artificial Intelligence*, *171*(7), 365–377. https://doi.org/10.1016/j.artint.2006.02.006

Tesauro, G., & Kephart, J. O. (2002). Pricing in Agent Economies Using Multi-Agent Q-Learning. *Autonomous Agents and Multi-Agent Systems*, *5*(3), 289–304. https://doi.org/10.1023/A:1015504423309

Tesfatsion, L. (2002). Agent-Based Computational Economics: Growing Economies From the Bottom Up. *Artificial Life*, *8*(1), 55–82. https://doi.org/10.1162/106454602753694765

Tesfatsion, L. (2003). Agent-based computational economics: Modeling economies as complex adaptive systems. *Information Sciences*, *149*(4), 262–268. https://doi.org/10.1016/S0020-0255(02)00280-3

Zheng, S., Trott, A., Srinivasa, S., Parkes, D. C., & Socher, R. (2022). The AI Economist: Taxation policy design via two-level deep multiagent reinforcement learning. *Science Advances*, *8*(18), eabk2607. https://doi.org/10.1126/sciadv.abk2607

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl
ISSN 2957-0506