# Decentralized pricing in supply chain management: a blockchain-enabled multi-agent Reinforcement Learning approach

Panagiotis G. Giannopoulos
*School of Social Sciences*
*Hellenic Open University*
Patras, Greece
0000-0001-7530-2458

Vangelis Malamas
*Department of Informatics*
*University of Piraeus*
Piraeus, Greece
0000-0002-9154-8340

Thomas K. Dasaklis
*School of Social Sciences*
*Hellenic Open University*
Patras, Greece
0000-0003-1240-4822

***Abstract*—Pricing decisions in Supply Chain (SC) management play a crucial role in profit maximization. However, SC dynamics are increasingly complex and optimizing pricing-related decisions remains challenging due to centralized models that cause inefficiencies and slow adaptation. Trust issues between manufacturers and retailers, often driven by information asymmetry and data manipulation risks, also hinder proper SC coordination. In this paper we present a novel decentralized framework for dynamic pricing in SCs by integrating multi-agent reinforcement learning (MARL) with blockchain technology. We model the SC as a Stackelberg game where multiple manufacturers, acting as leaders, independently set wholesale prices and a single retailer, as the follower, adjusts retail prices to maximize profit. Each agent employs RL algorithms to iteratively learn optimal pricing strategies in response to evolving market conditions and competitor actions. The integration of a permissioned blockchain ensures that all pricing decisions, transactions and smart contract executions are immutably recorded, thus enhancing data integrity, security and transparency. Experimental evaluations under diverse market scenarios demonstrate that the proposed framework improves convergence rates, profit margins and system robustness. Our results highlight the potential of combining decentralized decision-making with tamper-proof ledger technology to foster trust and efficiency in complex SC ecosystems.**

***Index Terms*—Blockchain, Multi-Agent Systems, Reinforcement Learning, Dynamic Pricing, Supply Chain Management, Stackelberg Game, Smart Contracts, Decentralized Decision-Making**

## I. Introduction

In a supply chain (SC) scenario modeled as a Stackelberg game, manufacturers act as the leaders who set their wholesale prices to maximize their individual profits, anticipating the retailer's subsequent decisions. Each manufacturer's pricing strategy not only reflects its internal cost structure and competitive positioning but also incorporates expectations about the retailer's optimal response. The sequential decision-making process inherent in the Stackelberg framework means that manufacturers must balance the desire to secure higher margins against the risk of setting prices too high, which might discourage the retailer from promoting their products. At its core, this interaction aligns with the well-established two-echelon SC configuration, where manufacturers actively oversee inventory levels and pricing adjustments to ensure market efficiency and profitability.

Conversely, the retailer functions as the follower in the Stackelberg game by taking the wholesale prices as given and setting optimal retail prices to maximize its own profit while remaining competitive in the market. The retailer's decision process involves aggregating the impact of multiple manufacturers' pricing strategies and then determining a retail pricing strategy that covers the wholesale costs while appealing to end-customers. This equilibrium, often referred to as the Stackelberg equilibrium, highlights how the interplay between the upstream pricing decisions and the downstream retail response leads to a stable state where no single party can unilaterally improve its outcome. Analyzing this equilibrium provides valuable insights into the distribution of pricing power within the SC and the resultant market dynamics, thereby informing strategies for both profit maximization and competitive positioning.

SC dynamics in nowadays are becoming increasingly complex and decentralized, thus optimizing pricing strategies remains a formidable challenge. Traditional pricing models often rely on centralized decision-making, a critical aspect that leads to inefficiencies, lack of transparency and slow adaptability to market fluctuations. Moreover, trust issues between manufacturers and retailers further complicate coordination, as information asymmetry and potential data manipulation hinder seamless collaboration. To address these challenges, we introduce a novel decentralized framework that integrates MARL with blockchain technology to enhance pricing decision-making. In particular, integrating blockchain with multi-agent systems introduces a transformative framework for SC management by ensuring unparalleled transparency, security and efficiency in pricing strategies. By combining the decision-making capabilities of autonomous agents with blockchain's immutable ledger, the system guarantees that every pricing decision, transaction and contractual agreement is securely recorded and verifiable. This combination not only

979-8-3503-5348-8/24/$31.00 ©2024 IEEE

mitigates risks associated with data tampering and fraud but also facilitates real-time, decentralized coordination among manufacturers and retailers. Moreover, the use of smart contracts automates and enforces agreements, reducing administrative overhead and enhancing responsiveness to market dynamics. Arguably, this innovative approach leverages the strengths of both technologies to create a resilient, data-driven ecosystem that optimizes pricing strategies while fostering trust and collaboration across the SC.

## II. Literature review

The dynamics of SC pricing strategies have been extensively studied in the literature from both a theoretical and empirical perspective. For instance, [1] examined dynamic pricing decisions in a two-echelon SC with a dominant retailer. The study utilized a backward induction approach to determine equilibrium in both centralized and decentralized settings. Similarly, [2] analyzed a hierarchical pricing decision process in a dual-channel SC. Their model identified optimal wholesale and retail prices under various operational strategies, highlighting that dual-channel operations are optimal only under specific conditions. [3] explored a two-layer SC where a socially responsible manufacturer engages with duopolistic retailers. They found that social responsibility impacts pricing decisions and that a two-part tariff contract can resolve channel conflicts. Additionally, [4] investigated vertical cooperative advertising, showing that cooperative pricing and advertising strategies lead to increased profitability and lower retail prices. Similarly, [5] analyzed the effects of price discounts and advertising expenditures under Stackelberg and Nash equilibria, concluding that manufacturers typically prefer Stackelberg equilibrium due to greater. [6] developed a profit-maximization model considering service levels, pricing strategies and market structures in a manufacturer-giant retailer SC. They demonstrated that a cooperative market structure yields the highest profits and that bargaining models can effectively distribute gains between SC participants.

Several studies employed game-theoretic frameworks to model pricing interactions between manufacturers and retailers. [7] used a Stackelberg game to analyze the pricing strategies in a SC with one manufacturer and two retailers. They demonstrated that wholesale prices depend on order quantity and timing. [8] compared Nash, Stackelberg-Retailer and Cooperative games to assess their impact on pricing, advertising and service decisions. Their findings indicated that the Stackelberg-Retailer model provides higher profits than the Nash game, while cooperation results in the highest total profit. Several studies adopt game-theoretic frameworks to analyze equilibrium pricing strategies. [9] explore a SC comprising multiple manufacturers and a retailer, incorporating the role of online reviews. By employing a Stackelberg game model, the study finds that online reviews influence both wholesale and retail prices, enhancing the retailer's ability to optimize its pricing strategy. This suggests that external market factors such as consumer feedback can significantly alter equilibrium pricing. [10] extend equilibrium analysis by considering a SC with one common manufacturer and two competing retailers. Their study investigates how Bertrand competition and retailer collusion impact pricing outcomes.

The strategic interplay between manufacturers and retailers significantly influences pricing decisions. [11] develops an empirical model to infer vertical strategic interaction (VSI) between manufacturers and retailers alongside horizontal strategic interaction (HSI) among competing manufacturers. The study finds that the manufacturer-retailer relationship follows a Manufacturer-Stackelberg structure, where manufacturers set wholesale prices first and the retailer subsequently sets retail prices. In a similar vein, [12] examine the power dynamics between manufacturers and a common retailer in a local market. Their study, employing game-theoretic modeling, shows that retailer power often dominates, allowing the retailer to capture a significant portion of the channel profits.

The structure of pricing contracts between manufacturers and retailers plays a crucial role in determining overall profitability. [13] examine markup pricing strategies between a dominant retailer and competing manufacturers. Similarly, [14] analyze the effects of wholesale prices on a retailer's shelf-space allocation and retail pricing decisions. Their findings indicate that lower wholesale prices and greater price elasticity result in higher shelf-space allocation for a brand, which in turn affects the retailer's pricing strategy. Finally, understanding demand characteristics is essential for developing effective pricing strategies. [15] introduce a nested logit model to capture the effect of unobserved household heterogeneity on pricing decisions. This is particularly important in contexts where demand uncertainty exists, as observed in [13], who demonstrate that demand uncertainty intensifies the effect of markup pricing structures on order quantities and channel profits. [16] further expand the pricing literature by considering green promotion efforts in a retailer-owned dual-channel SC with multiple manufacturers.

## III. High-level architecture

This paper proposes an autonomous decision-making tool tailored to the need of two-echelon SCs, specifically by enabling convergence towards pricing strategies that optimize both profitability and inventory holding costs. Particularly, we elaborate upon CL-DE paradigms, to create a multi-agent RL module that addresses the dynamic needs of manufacturers and retailers within SC settings. The environment is structured as a multi-agent Stackelberg game, where manufacturers function as leaders who strategically determine wholesale prices considering their respective cost structures, projected market demand and anticipated retailer responses. On the other hand, the retailer agent act as follower in this game, with the purpose of dynamically processing wholesale prices and adjusting pricing plans to align with consumer purchasing behavior. To ensure efficient interaction and synchronization among these agents, an integration layer has been established. This layer facilitates seamless data-driven communication, allowing real-time adaptability in response to market fluctuations, consumer

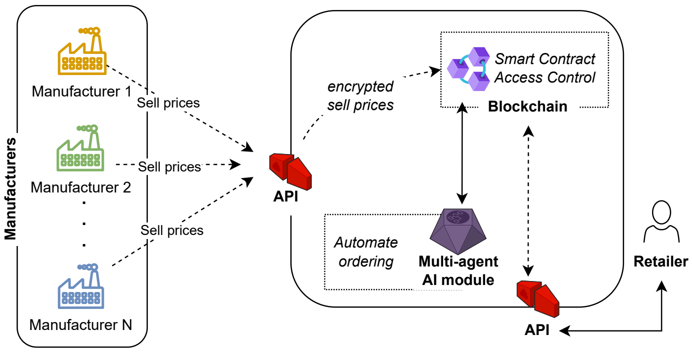


Fig. 1. A high level architecture leveraging intergration of blockchain and multi-agents

demand signals and cost variations, thus exemplifying a fully-functional operating scenario.

For complementing the AI module, we introduce a robust blockchain network that underpins the system with transparency and security. This permissioned blockchain records every transaction, pricing decision and contractual agreement in an immutable ledger, ensuring data integrity and fostering trust among participants. Smart contracts are employed to automate and enforce pre-defined agreements, such as payment settlements or price adjustments, thereby reducing administrative overhead and mitigating risks associated with human error or fraud. Together, the integration of blockchain with RL-driven agents establishes a resilient ecosystem that not only optimizes pricing strategies but also promotes decentralized coordination and real-time data sharing across the SC. The proposed architecture is seen in Figure 1.

### A. Blockchain module

The blockchain component is utilized to secure the invoicing and transactions in the proposed system, thereby ensuring the security of data and transactions. To regulate interactions throughout the SC, a role-based access control (RBAC) scheme is implemented that is embedded in the blockchain network. Through RBAC, manufacturers and the retailer are granted specific permissions and can take actions. For example, manufacturers can set wholesale prices and trigger transactions, while the retailer can view approved wholesale prices, close purchase contracts and trigger automatic payments. Automation of business processes and user action is achieved through the use of Smart Contracts. In this way, the blockchain module automatically ensures that the pricing and transaction policies -which are set in advance by the participants - cannot be overwritten.

All data of sale prices that are sent by the manufacturers are encrypted before being stored in a block in order to ensure confidentiality. When manufacturers submit their price quote in an encrypted blockchain format, the whole dataset is forwarded to multi-AI agents for processing. Then, once the AI pricing module decides further the purchase actions, these actions are recorded to the blockchain as a transaction. Once the players agree to a purchase order, a smart contract is automatically executed and irrevocably records the order. Finally, smart contracts implement the execution and settlement of the payment. To ensure ledger consistency and reliability, the blockchain network utilizes a Byzantine Fault-Tolerant (BFT) consensus protocol, such as Practical Byzantine Fault Tolerance (PBFT) or RAFT. These protocols validate transactions in real-time and preserve data integrity, even in adversarial conditions. Cryptographic techniques, including SHA-256 hashing and digital signatures, safeguard stored data, making the system tamper-proof. The blockchain seamlessly integrates with a MARL system through an API layer, enabling RL agents representing manufacturers and the retailer to securely interact with the blockchain. These agents can access verified pricing data, make pricing decisions, execute transactions and analyze historical pricing trends to optimize RL-based policies. Since all training and decision-making data originates from an immutable, tamper-resistant blockchain ledger, RL optimizations remain accurate and reliable.

### B. AI pricing module

As mentioned, the RL module consists of two agent types ($N$ manufacturer agents and a single retailer agent); is structured as a Stackelberg game and formulated as a Markov Decision Process (MDP). At each discrete time step $t$, each manufacturer agent $i \in \{1, \ldots, N\}$ observes a state $s_i(t)$, including its production costs, inventory levels and raw material prices, to choose an action $a_i(t)$, corresponding to a wholesale price from a discrete action set. Subsequently, the retailer agent, acting as a follower, observes its own state $s_r(t)$, comprising the manufacturers' wholesale prices, current inventory and market demand conditions and selects an action $a_r(t)$, namely the retail price. The AI module was intentionally selected to be built on a CL-DE paradigm towards mitigating the non-stationarity inherent in multi-agent settings.

Since this module is based on CL-DE, all of the states and sequential pricing decisions are stored in a centralized matrix. Equation 1 represents the vectorized form of storing the states of each agent in the central matrix. A similar vector is also designed for the actions undertaken for each of the agents. Both of these vectors are then stored in a central buffer, which corresponds to the main element used for CL purposes.

$$s(t) = \{s_1(t), s_2(t), \ldots, s_N(t), s_r(t)\} \tag{1}$$

The framework leverages the blockchain functionalities to facilitate a verifiable experience, by sharing the tuples $(s(t), a(t), r(t), s(t+1))$ and store them in the central experience buffer. In this context, $r(t)$ denotes the immediate reward obtained by each agent at time $t$, calculated based on its pricing decision, incurred costs, and resulting sales volume, as further detailed in Equations 4a and4b. The selection of transmitting vectorized data in the central buffer was made towards analyzing in a simplistic scenario the performance of our model, while in more detailed approaches continuous data could potentially transmit. Due to the format of the composed dataset, the training process is executed by incorporating the established Deep Q-Networks (DQN) approach.

Specifically, we employ multiple DQNs, each corresponding to a specific agent. The first DQN structure, applied within the centralized module, is used to train the retailer agent in optimizing its policy for selecting the best combination of products from manufacturers. This enables the retailer to make data-driven purchasing decisions that enhance SC efficiency. To ensure stability and convergence in the learning process, we also incorporate a soft update mechanism for synchronizing the online and target networks of each DQN. Specifically, the target network parameters $\theta_{\text{target}}$ are updated using the online network parameters $\theta$ through the exponential moving average rule, as given in Equation 2, in which $\tau \in (0, 1]$ is a smoothing factor controlling the degree of weight transfer per update. Specifically, lower $\tau$ promotes stable learning by reducing abrupt policy shifts, while a higher $\tau$ accelerates adaptation to new market conditions. This mechanism ensures that both the retailer and manufacturers maintain robust decision-making strategies, enabling dynamic adjustments to pricing and inventory levels in response to evolving SC dynamics.

$$\theta_{\text{target}} \leftarrow \tau\theta + (1-\tau)\theta_{\text{target}} \tag{2}$$

At the decentralized level, each manufacturer utilizes its own DQN to determine an optimal wholesaling policy, ensuring cost-efficient and stable inventory management. While decision-making remains decentralized, all agents benefit from centralized learning, as their state-action representations are processed in the shared buffer. The DQNs update their Q-values iteratively based on observed rewards and transitions, refining their policies over time to improve overall system performance in both simplified and complex scenarios. Once training is completed, each agent independently deploys its learned policy, making real-time pricing and inventory decisions based solely on local observations and its current production and inventory states.

## IV. Implementation details

### A. Blockchain-based implementation

The blockchain module serves as the backbone of our system, ensuring a secure and immutable record of all transactions and decision events. Implemented as a permissioned blockchain, it restricts participation to authorized entities—namely, the manufacturers and the retailer—thus establishing a controlled and trusted environment. Each participant operates a blockchain node that maintains a full copy of the ledger, ensuring that all pricing decisions, state transitions and smart contract executions are permanently recorded. The network employs a consensus algorithm, such as RAFT, to validate and order transactions in real time, which guarantees consistency and resilience even under adverse conditions. In addition, smart contracts are deployed to automate critical business processes, including the enforcement of pricing thresholds, automatic execution of payments and the imposition of penalties for deviations from contractual agreements. An API layer interfaces between the blockchain and the RL agents, enabling secure submission and retrieval of state-action-reward tuples for training, auditing and real-time de

### B. Agents' implementation: rewards and update

On the implementation side, DQNs are based on Multi-Layer Perceptron (MLP) architectures, leveraging fully connected layers to approximate the Q-function. These networks process the vectorized state representations stored in the central buffer, enabling the model to learn optimal action-value mappings. The MLP-based DQNs are trained using experience replay, where past experiences are sampled to stabilize learning and target networks, which mitigate instability in Q-value updates. Specifically, agents update their Q-values using the standard Bellman equation (see Eq. 3), where $\alpha$ denotes the learning rate, controlling the magnitude of Q-value updates, and $\gamma$ is the discount factor, determining the present value of future rewards. For the retailer agent, this update occurs within the centralized learning phase, incorporating data from all manufacturers to optimize its pricing strategy. In contrast, for the manufacturer agents, Q-value updates take place in a decentralized manner, relying on local state representations, where each manufacturer refines its wholesaling policy based on individual inventory levels, production rates, and previously observed responses of the retailer.

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right] \tag{3}$$

Since our approach incorporates multiple DQNs, all of them are accordingly updated based on the Equation 3. This rule could be considered as a temporal-difference approximation of the Bellman optimality, as commonly employed in model-free RL paradigms. Towards facilitating a fast-converged identification of the optimal pricing, we structure two different reward mechanisms. The reward function for each manufacturer $i$ is formulated by using Equation 4a, in which $a_i(t)$ is the wholesale price chosen by manufacturer $i$, $c_i(t)$ is its production cost and $Q_i(t)$ represents the quantity sold. For the retailer the corresponding reward function is presented in Equation 4b, in which $a_r(t)$ is the retailer's selling price, $\bar{w}(t)$ is the average wholesale price of the selected manufacturer and $Q_r(t)$ is the total quantity sold by the retailer. The term holding costs accounts for the expenses associated with maintaining inventory, while the term penalties is proportionally computed to the shortfall or inefficiency observed and is designed to discourage aggressive pricing towards system-wide coordination.

$$r_i(t) = \big(a_i(t) - c_i(t)\big) \times Q_i(t) - \text{penalties}, \tag{4a}$$

$$r_r(t) = \big(a_r(t) - \bar{w}(t)\big) \times Q_r(t) - \text{holding costs}. \tag{4b}$$

Beyond the reward mechanism, we have also incorporated an inventory replenishment mechanism with the aim of preventing stock depletion for manufacturers, which could potentially lead to mis-learned pricing plans. This mechanism is implemented by gradually adjusting stock levels toward a target inventory $I^*$ using a replenishment rate $\rho_i$, by following Equation 5.

$$I_i^{t+1} = I_i^t + \rho_i \cdot (I^* - I_i^t) \tag{5}$$

## V. Experimental Evaluation

For evaluation purposes, we orchestrate a Stackelberg game in which $N = 3$ manufacturers interact with a single retailer to optimize pricing decisions. At a proof-of-concept level several assumptions were made in order to simplify the complex interface in these up-stream SC settings. First, we assume that at the game's onset, all manufacturers hold the same inventory volume of fifty (50) units. Additionally, each manufacturer aims to gradually deplete its inventory to a target level of twenty (20) units over the course of the game. To regulate this process, we assume that the production rate of each manufacturer equals to 0.2, 0.4, and 0.3, respectively.

Within the game settings, manufacturers firstly choose their wholesale prices from a predefined discrete action space of $\{5, 10, 15, 20\}$ monetary units, while the retailer determines its selling price from a separate action space of $\{10, 20, 30, 40, 50\}$ monetary units. The base market demand at each time step is fixed at thirty (30) units, with demand sensitivity regulated by a demand elasticity coefficient ($\alpha$), which was set to 0.25 to simulate the variation of end-consumers needs. The cost of production for each manufacturer is randomly sampled from a uniform range of $[5, 10]$ monetary units.

The simulation runs over 1000 episodes each of them ran over twenty (20) maximum, with manufacturers and the retailer making pricing decisions sequentially. At each step, the retailer procures all demand from the cheapest available manufacturer, resolving ties randomly. The manufacturers then update their inventory levels, accounting for both sales volumes and replenishment rates. Throughout the game, the retailer seeks to maximize its profit margin, while manufacturers strategically adjust pricing to remain competitive while maintaining sustainable inventory levels.

After running the model, we manage to gather the results presented in Figures 2 and 3. Starting from Figure 2, it is evident that the model, after a reasonable number of episodes (around 500), has learned a policy that stabilizes both the selling price and profit margin. Initially, these two metrics exhibit high volatility as the retailer explores different pricing strategies. However, as training progresses, both converge within a stable range, indicating that the retailer has optimized its pricing policy in response to manufacturers' wholesale strategies. The buy price remains relatively stable, suggesting that manufacturers quickly settle on competitive pricing, reinforcing a predictable supply environment. This stabilization suggests the emergence of a Stackelberg equilibrium, where the retailer, acting as the leader, sets an optimal response strategy based on the manufacturers' pricing behaviors. The convergence pattern indicates that manufacturers, as followers, adjust their decisions accordingly, leading to a mutually stable pricing structure in the modeled SC interface.

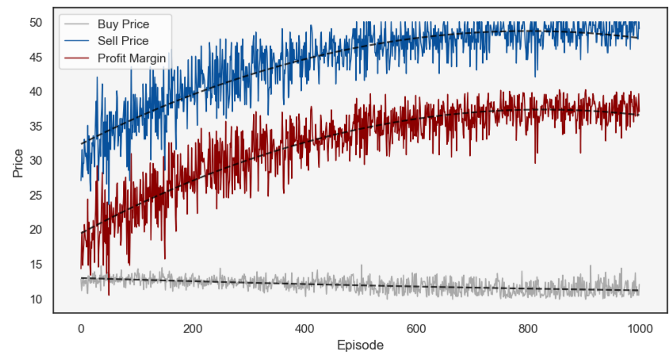


Fig. 2. Evolution of buy and sell prices throughout the game.

Beyond these indications, we also refer to Figure 3, in which we present the evolution of the per twenty episodes average wholesale price and the corresponding number of products in stock. Notwithstanding the volatility highlighted in the results, which holds true due to the dynamic nature of the perfect competition market orchestrated, some gradual stabilization trends can be observed. Particularly, for the wholesale pricing we can observe that both the first (MFG 0) and third manufacturer (MFG 2) tend to down-scaled pricing policies which leads to higher product sales. The effective self-adaptability of the model is also evidenced by the outliers, especially in the MFG 0 case, in which high values are clearly lead to shortened volumes of sales. Except for these two agents, the second one seems to have balanced on a middle pricing strategy, which suggests a moderate yet adaptive approach to wholesaling. Such behavior indicates that the agent has learned to balance profitability and demand responsiveness, allowing it to secure a consistent share of sales.

## VI. Discussion and Conclusions

In this paper we have proposed a blockchain-enabled multi-agent RL framework for decentralized pricing in SCs. The proposed framework has demonstrated significant improvements in both pricing efficiency and system robustness. Experimental results have indicated that the integration of blockchain technology with RL facilitate a faster convergence to stable pricing strategies where both manufacturers and the retailer are benefited. The proposed decentralized pricing approach has allowed manufacturers to autonomously set wholesale prices that are optimized to their respective cost structures and market conditions, thus leading to competitive yet sustainable pricing. The retailer, on the other hand, acting within a Stackelberg framework, has successfully adapted its pricing policies to these wholesale decisions and, therefore, has managed to maximize profit margins by dynamically responding to fluctuating market demands. Moreover, the blockchain-based approach has ensured transparency and data integrity, thus enhancing trust among SC participants.

The experimental evaluation has underscored several significant advantages. First, the integration of MARL has enabled efficient, real-time adaptability. This has significantly

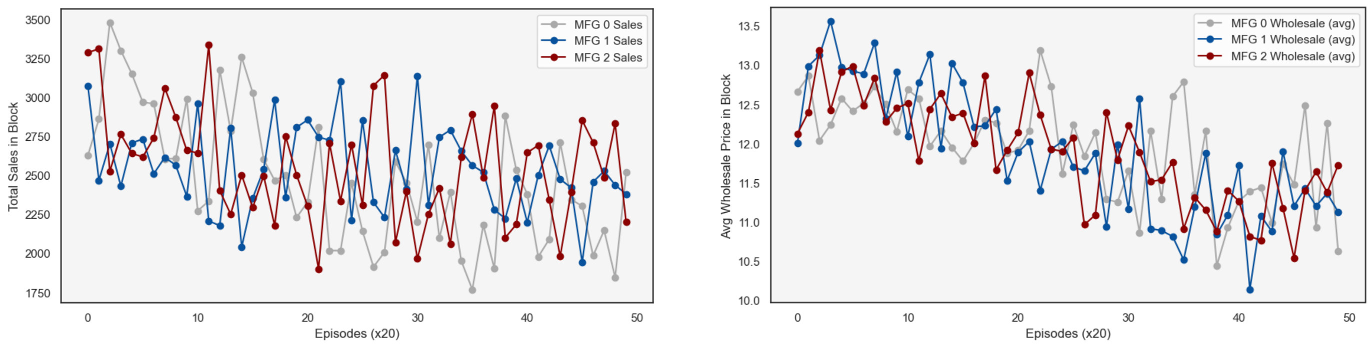


Fig. 3. Evolution of products sold (leftmost) and pricing plans (rightmost) for manufacturers.

improved the convergence rate of pricing strategies even under varying market scenarios, including demand surges and competitive conditions. In addition, the use of blockchain technology has provided an immutable ledger that securely recorded all transactions, pricing decisions and contractual agreements. This integration minimized the risk of data tampering and fraud, while smart contracts facilitated automated and error-free execution of business rules, considerably reducing administrative overhead and fostering seamless coordination among agents.

The presented framework also has several limitations that should be kept in mind. For example, the experimental evaluations were conducted under simplified assumptions, such as limited participants with fixed production rates, stable market conditions and a limited number of manufacturers. These simplifications may limit the generalizability of the findings to real-world SC conditions. Additionally, the CL-DE paradigm may introduce potential latency issues when scaling up the system or when deployed in geographically dispersed SC networks. Future research should therefore explore system scalability and performance in scenarios involving larger and more diverse agent populations and dynamic, unpredictable market fluctuations, to model “troubled” conditions. Also, expanding the framework to consider additional factors such as consumer heterogeneity and more complex inventory management scenarios by incorporating real data from SC settings could enhance its robustness and clarify its applicability. Finally, examining integration protocols for the proposed framework in legacy industrial systems presents a vital research area.

## Acknowledgment

This research has been funded by the Hellenic Open University under the project ”Multi-agent systems and Generative Artificial Intelligence in Management science: Innovative Applications in Human Resources and Operations management-PELOPAS”.

## References

[1] F. Ren, “Dynamic pricing decision in the two-echelon supply chain with manufacturer’s advertising and dominant retailer,” in *Proceedings of the 2011 Chinese Control and Decision Conference, CCDC 2011*, 2011, pp. 391 – 395.

[2] Q. Ding, C. Dong, and Z. Pan, “A hierarchical pricing decision process on a dual-channel problem with one manufacturer and one retailer,” *International Journal of Production Economics*, vol. 175, pp. 197 – 212, 2016.

[3] N. M. Modak, S. Panda, and S. S. Sana, “Pricing policy and coordination for a two-layer supply chain of duopolistic retailers and socially responsible manufacturer,” *International Journal of Logistics Research and Applications*, vol. 19, no. 6, pp. 487 – 508, 2016.

[4] G. Aust and U. Buscher, “Vertical cooperative advertising and pricing decisions in a manufacturer-retailer supply chain: A game-theoretic approach,” *European Journal of Operational Research*, vol. 223, no. 2, pp. 473 – 482, 2012.

[5] J. Yue, J. Austin, Z. Huang, and B. Chen, “Pricing and advertisement in a manufacturer-retailer supply chain,” *European Journal of Operational Research*, vol. 231, no. 2, pp. 492 – 502, 2013.

[6] R. Yan and J. Wang, “Service level, pricing strategy and firm performance in a manufacturer-giant retailer supply chain,” *Journal of Product and Brand Management*, vol. 19, no. 1, pp. 61 – 66, 2010.

[7] H.-C. Hung, J.-K. Kim, and C. C. Calugcug, “Dynamical pricing strategy for one-manufacturer and two-retailer supply chain model,” *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 4, no. 3, pp. 195 – 202, 2010.

[8] P. Esmaeili and M. Rasti-Barzoki, “Pricing, advertising, and service policies in a manufacturer-retailer supply chain with nash, stackelberg-retailer, and cooperative games,” *International Journal of Industrial Engineering and Production Research*, vol. 35, no. 4, pp. 143 – 166, 2024.

[9] X. Cai, J. Li, B. Dai, and T. Zhou, “Pricing strategies in a supply chain with multi-manufacturer and a common retailer under online reviews,” *Journal of Systems Science and Systems Engineering*, vol. 27, no. 4, pp. 435 – 457, 2018.

[10] H. Huang, H. Ke, and L. Wang, “Equilibrium analysis of pricing competition and cooperation in supply chain with one common manufacturer and duopoly retailers,” *International Journal of Production Economics*, vol. 178, pp. 12 – 21, 2016.

[11] K. Sudhir, “Structural analysis of manufacturer pricing in the presence of a strategic retailer,” *Marketing Science*, vol. 20, no. 3, pp. 244 – 264, 2001.

[12] V. Kadiyali, P. Chintagunta, and N. Vilcassim, “Manufacturer-retailer channel interactions and implications for channel power: An empirical investigation of pricing in a local market,” *Marketing Science*, vol. 19, no. 2, pp. 127 – 148, 2000.

[13] J.-C. Wang, A.-M. Wang, and Y.-Y. Wang, “Markup pricing strategies between a dominant retailer and competitive manufacturers,” *Computers and Industrial Engineering*, vol. 64, no. 1, pp. 235 – 246, 2013.

[14] G. Martín-Herrán, S. Taboubi, and G. Zaccour, “The impact of manufacturers’ wholesale prices on a retailer’s shelf-space and pricing decisions,” *Decision Sciences*, vol. 37, no. 1, pp. 71 – 90, 2006.

[15] P. K. Chintagunta and N. J. Vilcassim, “Empirical implications of unobserved household heterogeneity for manufacturer and retailer pricing,” *Journal of Retailing and Consumer Services*, vol. 5, no. 1, pp. 15 – 24, 1998.

[16] Y. Zhao, W. Huang, E. Xu, and X. Xu, “Pricing and green promotion decisions in a retailer-owned dual-channel supply chain with multiple manufacturers,” *Cleaner Logistics and Supply Chain*, vol. 6, 2023.