


Adaptive Agents in Spatial Double-Auction Markets: Modeling the Emergence of Industrial Symbiosis

Matthieu Mastio 

IRIT, UMR 5505 CNRS, Université Toulouse Capitole
Toulouse, France
matthieu.mastio@irit.fr

Benoit Gaudou 

IRIT, UMR 5505 CNRS, Université Toulouse Capitole
Toulouse, France
benoit.gaudou@ut-capitole.fr

Paul Saves 

IRIT, UMR 5505 CNRS, Université Toulouse Capitole
Toulouse, France
paul.saves@irit.fr

Nicolas Verstaevel 

IRIT, UMR 5505 CNRS, Université Toulouse Capitole
Toulouse, France
nicolas.verstaevel@ut-capitole.fr





ABSTRACT

Industrial symbiosis fosters circularity by enabling firms to re-purpose residual resources, yet its emergence is constrained by socio-spatial frictions that shape costs, matching opportunities, and market efficiency. Existing models often overlook the interaction between spatial structure, market design, and adaptive firm behavior, limiting our understanding of where and how symbiosis arises. We develop an agent-based model where heterogeneous firms trade byproducts through a spatially embedded double-auction market, with prices and quantities emerging endogenously from local interactions. Leveraging reinforcement learning, firms adapt their bidding strategies to maximize profit while accounting for transport costs, disposal penalties, and resource scarcity. Simulation experiments reveal the economic and spatial conditions under which decentralized exchanges converge toward stable and efficient outcomes. Counterfactual regret analysis shows that sellers' strategies approach a near Nash equilibrium, while sensitivity analysis highlights how spatial structures and market parameters jointly govern circularity. Our model provides a basis for exploring policy interventions that seek to align firm incentives with sustainability goals, and more broadly demonstrates how decentralized coordination can emerge from adaptive agents in spatially constrained markets.

KEYWORDS

Spatial Agent-Based Simulation, Industrial Symbiosis, Circular Economy, Distributed Market Analysis, Reinforcement Learning

ACM Reference Format:

Matthieu Mastio , Paul Saves , Benoit Gaudou , and Nicolas Verstaevel . 2026. Adaptive Agents in Spatial Double-Auction Markets: Modeling the Emergence of Industrial Symbiosis. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, Paphos, Cyprus, May 25 – 29, 2026, IFAAMAS, 10 pages. <https://doi.org/10.65109/EXII2056>



This work is licensed under a Creative Commons Attribution International 4.0 License.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). <https://doi.org/10.65109/EXII2056>

1 INTRODUCTION

Industrial Symbiosis (IS) refers to the exchange of byproducts between firms, turning one company's waste into valuable inputs for another one. By reusing materials such as heat, rubble, or chemical residues, IS reduces environmental impact while generating economic benefits, embodying the circular economy principle of closing resource loops. For instance, the RETDA eco-industrial park in China reported a 33% increase in input productivity, a 3650% rise in water use efficiency, and a 30.91% reduction in emissions [11].

Despite its potential, IS rarely emerges spontaneously. Exchanges depend on local supply-demand compatibilities, economic incentives, and logistical constraints [29]. Poor coordination often results in wasted resources and missed opportunities, highlighting the gap between individual strategies and collective circularity goals. Identifying the conditions under which decentralized firms can coordinate effectively is, therefore, a key research challenge.

Analytical solutions are often intractable due to the combinatorial complexity of interactions and the emergent nature of prices and traded quantities. Simulation-based approaches, by contrast, are particularly suited to study such nonlinear, decentralized, and heterogeneous systems [28]. By systematically varying system configurations, simulations can indeed reveal market behaviors and highlight factors that support efficient and resilient local circularity. Agent-based models, in particular, are able to capture heterogeneity among actors and test how different market designs, decision rules, and learning mechanisms affect the emergence of exchanges under realistic economic constraints [18]. Their ability to explicitly represent geosituated agents makes them especially suitable for studying spatially dependent interactions, such as proximity-based collaborations, transportation costs, and localized resource flows. Unlike theoretical game-theoretic models, which often rely on a small number of representative players to ensure tractability, agent-based models can accommodate many heterogeneous actors and interactions.

However, their effectiveness is often constrained by the simplifications made in representing agent behaviors. In many existing symbiosis models [7, 20, 22, 25, 44, 62], firms follow fixed behavioral rules or static pricing strategies, limiting their ability to capture the adaptive and uncertain nature of real-world exchanges. In practice, firms continuously adjust decisions in response to shifting demand, fluctuating costs, and evolving environmental regulations. This highlights the need to enhance decision-making mechanisms

within agent-based models, enabling agents to learn, adapt, and explore strategies in a more dynamic way.

In this paper, we present a spatially explicit, decentralized multi-agent model in which buyers and sellers of byproducts interact through multilateral spatial double auctions using co-evolving strategies. Sellers adaptively update their pricing strategies via reinforcement learning in a partially observable environment, while buyers select offers that maximize utility subject to spatial and economic constraints. A key feature of the model is that both transaction prices and traded quantities emerge endogenously from these interactions, allowing us to investigate how local decisions give rise to systemic patterns. While illustrated here through industrial symbiosis, this framework applies broadly to circular economy settings, such as heat recovery, wastewater reuse, or biomass valorization, where decentralized agents, ranging from firms to aggregated entities like neighborhoods, must coordinate to close resource loops under transport frictions. Our contributions are threefold. First, we develop a reinforcement learning-based market model for decentralized byproduct exchanges, based on a novel auction mechanism adapted to spatial and economic constraints of territorial circularity. Second, we validate the learned strategies via counterfactual regret analysis, showing convergence toward near-Nash equilibrium. Finally, we propose a simulation-based analysis of the conditions under which adaptive pricing and agent interactions lead to stable circular outcomes, providing insights for policymakers.

The paper first reviews related work (Section 2), then presents the decentralized market model (Section 3) and simulation setup (Section 4). Results on price formation and local circularity are analyzed in Section 5, followed by discussion (Section 6) and conclusions (Section 7).

2 RELATED WORK

2.1 Industrial Symbiosis and Circular Economy Modeling

Industrial symbiosis in Eco-Industrial Parks (EIPs) has traditionally been studied using system dynamics to capture network level behavior and feedback loops [6, 35, 43, 54, 66]. Complementary approaches have treated EIPs as complex adaptive systems, highlighting dynamic business interactions and emergent symbiotic relationships [13, 19, 32, 48, 64]. Other studies focused on material and energy flows, byproduct exchanges, and social network analysis [12, 17, 19, 49, 63]. Recent reviews [16, 45] emphasize the potential of agent-based modeling to capture local, adaptive, and self-organizing behaviors. Emerging agent-based studies have integrated input-output models [62], explored spatially networked exchanges [25, 38, 44], and examined sector specific synergies [7, 20, 22]. Yet, despite these advances, the explicit representation of economic incentives and market coordination mechanisms remains underexplored. Additionally, these works do not explicitly model markets where prices, trade volumes, and resource allocations emerge endogenously from decentralized interactions.

2.2 Auctions and Market Mechanisms in Spatial Settings

Auction and negotiation mechanisms in Multi-Agent Systems (MAS) have become canonical tools for decentralized coordination. From early negotiation platforms [9] to double auctions—market mechanisms in which multiple buyers and sellers simultaneously submit bids and offers—used to study bidding strategies and price discovery [51, 57], research has demonstrated how decentralized negotiations can generate robust outcomes. Extensions have examined repeated competition [4, 34], peer-to-peer trading [42], social influence [60], and multilateral bargaining across multiple resources [1, 10]. Spatial aspects are frequently acknowledged, for instance through costly trade links [3], heterogeneous bilateral transport costs [14], or land markets where localized interactions induce differentiated prices [21]. Yet these contributions remain largely theoretical, typically involving few participants and limited attention to adaptive mechanisms that capture the evolving dynamics of real markets.

2.3 Learning Agents for Adaptive Market Participation

Traditional learning agents often rely on fixed rules [15], or supervised training [33], limiting their ability to adapt to dynamic market conditions and strategic interactions. Reinforcement Learning (RL) overcomes these limitations by allowing agents to learn optimal policies through trial-and-error, balancing exploration and exploitation in complex, evolving environments. In the context of double auctions, RL agents are particularly well-suited due to the complex interactions between buyers and sellers and the continuous feedback provided by order book dynamics. Agents can be designed to adopt specific roles, such as deciding when to buy, sell, or set prices, enabling cooperative and strategic trading behaviors [5, 37]. Multi-Agent Reinforcement Learning (MARL) extends this approach to settings with multiple interacting agents, allowing the simulation of bottom-up market dynamics [23, 31, 36, 53]. MARL models have been shown to reproduce emergent behaviors in order books and centralized market microstructures by capturing interactions among heterogeneous agents [31]. This enables detailed analysis of price formation, liquidity, and strategic adaptation. However, existing approaches largely ignore spatial constraints, as these are typically absent in the stock market. Yet, such constraints can be critical in markets where transportation or delivery costs influence trading decisions [47]. RL with spatial considerations has been primarily explored in logistics, where carriers and shippers act as bidders in dynamic auctions mediated by centralized brokers [2, 41, 61, 65]. Even when shippers themselves are modeled as learning agents [59], exchanges remain intermediary mediated. In contrast, many real-world IS transactions are inherently local and broker-free. To the best of our knowledge, RL has not yet been applied to decentralized, spatially constrained markets where transport costs and policy incentives influence agents' strategies.

2.4 Positioning of This Work

Taken together, three main research gaps emerge:

- (1) **Industrial symbiosis models** rarely provide an in-depth mechanism for endogenous market-based price formation.

- (2) **Auction and negotiation models** seldom incorporate adaptive learning of agents.
- (3) **Reinforcement learning in markets** has not been applied to decentralized settings where spatial topology influences agents' behavior.

Our contribution is to bridge these strands by introducing a decentralized auction model for industrial symbiosis. While RL and double auctions have been studied in isolation, their integration into a spatial multilateral double-auction, where transport costs and location directly drive strategic behavior, constitutes a novel contribution. Specifically, our model ensures that (i) sellers adapt prices through reinforcement learning, (ii) spatial constraints are explicitly internalized, and (iii) both prices and traded quantities emerge from multilateral interactions. Our objective is to reproduce the systemic complexity of inter-firm interactions within industrial symbiosis as a subset of the broader circular economy, providing a foundational basis for modeling circular economic dynamics. Beyond the specific application, the contribution lies in the rigorous methodology and implementation of this multi-agent RL framework, offering a reusable and generalizable approach for studying spatial market interactions.

3 MODEL

3.1 Environment and Agents

To concentrate on the essential dynamics of the model, we restrict our analysis to the exchange of a single byproduct. This simplification enables us to isolate and assess the effects of multilateral interactions and sellers' strategies, without additional complexity introduced by multiple interdependent byproducts. We model a decentralized market for one byproduct, composed of buyers $\mathcal{B} = \{B_1, \dots, B_{N_B}\}$ and sellers $\mathcal{S} = \{S_1, \dots, S_{N_S}\}$. Each agent has a spatial location, and the distance between a buyer B_i and a seller S_j is denoted d_{ij} . Buyers have heterogeneous initial demands q_i drawn from a scaled uniform distribution and evaluate incoming offers based on their price sensitivity β_i . Specifically, buyer B_i accepts an offer p_{ij} from seller S_j only if $p_{ij} \leq \beta_i p_m$, where p_m is the reference market price. This mechanism ensures that buyers only engage in transactions that meet their cost expectations relative to the external market. We adopt here a "Small Open Economy" [26] assumption: we hypothesize that local industrial park internal exchanges are too small to impact the global commodity market price. Consequently, p_m is treated as a fixed exogenous parameter.

Sellers are endowed with quantities q_j and propose personalized prices to buyers, defined as

$$p_{ij} = \phi_j p_m + c_{ij}, \quad (1)$$

The term c_{ij} is originally defined as $c_{ij} = d_{ij} c_t$, where c_t represents the transportation cost per kilometer (expressed as a percentage of the market price). This term can be further extended to capture the total logistics cost required for the byproduct provided by the seller to become usable by the buyer, including handling, treatment, or adaptation costs. Finally, ϕ_j is an adaptive pricing parameter that evolves during the learning phase.

In addition to market and transportation costs, sellers face a penalty c_d for unsold quantities (also expressed as a percentage of the market price), representing the cost of sending residual flows to

landfill and reflecting both economic and environmental impacts. If a seller j ends a timestep with q_j^{unsold} units remaining, its reward is reduced by $q_j^{\text{unsold}} c_d$, imposing a direct financial penalty that encourages sellers to adjust pricing strategies to minimize waste while still aiming to maximize profit.

To characterize market conditions, we define the scarcity level of the byproduct as

$$s = \frac{\sum_{i \in \mathcal{B}} q_i}{\sum_{j \in \mathcal{S}} q_j}. \quad (2)$$

This measure allows us to systematically investigate the effects of supply-demand imbalances on agents' behavior in the decentralized market. Each seller S_j aims to maximize its cumulative expected profit at each simulation step. Let $\mathbf{q}_j^t = (q_{ij}^t)_{i \in \mathcal{B}}$ denote the vector of quantities bought by B_i from S_j at step t , and $\mathbf{p}_j^t = (p_{ij}^t)_{i \in \mathcal{B}}$ the corresponding price vector. Defining $\mathbf{c}_j = (c_{ij})_{i \in \mathcal{B}}$ as the transport cost vector, the instantaneous profit can be written compactly as

$$r_j^t = \langle \mathbf{q}_j^t, \mathbf{p}_j^t - \mathbf{c}_j \rangle - q_j^{\text{unsold},t} c_d, \quad (3)$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product. The first term corresponds to the net revenue from executed contracts, while the second term penalizes unsold inventory through the landfill cost c_d . This spatially defined market closely resembles a *spatial Bertrand oligopoly* [30], where multiple sellers compete by setting prices, and buyers prefer lower-cost options. Unlike the classical Bertrand model, where identical products drive prices toward marginal cost, the inclusion of transport costs differentiates sellers spatially, allowing local price variation and creating room for adaptive strategies to emerge.

3.2 Decentralized Multilateral Auction

We model the interactions between buyers and sellers using a decentralized multilateral auction described in Algorithm 1. Each seller submits a set of bids to each buyer, where each bid specifies a price and quantity for a given product p , computed on the basis of the logistic costs required to deliver the product to the buyer's location and the external market price. Buyers evaluate the incoming bids and select the one that maximizes their utility, subject to their price sensitivity β_i and demand q_i . Selected bids are then sent back as proposals to sellers. Each seller chooses a subset of the received bids to accept in order to maximize expected profit, taking into account transport costs c_{ij} and remaining inventory q_j . Contracts are executed, updating both buyers' needs and sellers' available quantities. This process iterates until no further mutually acceptable bids can be matched. Sellers adapt their pricing strategy ϕ_s over time using a strategy based on historical rewards that will be described in 3.3.

3.3 Seller Reinforcement Learning and Action Selection

It has been shown by [4] that Bertrand competition problems can be effectively learned using bandit algorithms. Accordingly, each seller's sequential pricing decision is modeled as a contextual multi-armed bandit. The action space is discretized into K intervals, $\forall 0 < k \leq K$:

$$\Phi = \{\phi^{(1)}, \dots, \phi^{(K)}\}, \quad \phi^{(k)} = \phi_{\min} + \frac{k-1}{K-1} (1 - \phi_{\min}), \quad (4)$$

Algorithm 1 Decentralized Multilateral Auction

Require: Buyers \mathcal{B} , Sellers \mathcal{S} , product p

- 1: **while** contracts have been made at the previous timestep **do**
- 2: **Seller bids:** each $S_j \in \mathcal{S}$ submits a bid (q_j, p_{ij}) to each $B_i \in \mathcal{B}$, where q_j is the maximum quantity available for S_j , and $p_{ij} = \phi_j p_m + d_{ij} c_t$
- 3: **Buyer choice:** each B_i evaluates the received bids, selects the most attractive one within price tolerance $\beta_i p_m$, and sends an acceptance proposal to the corresponding seller
- 4: **Seller allocation:** each S_j sorts accepted bids by expected profit per sold unit, confirms contracts until inventory is exhausted, updates sold quantity and profit
- 5: Remove satisfied buyers and empty sellers from \mathcal{B} and \mathcal{S}
- 6: **end while**
- 7: **Return** executed contracts

where $\phi_{\min} = -\frac{c_d}{p_m}$. At every timestep t , seller S_j selects a discrete pricing action $\phi_j^t \in \Phi$ based on the current market context and receives a reward r_j^t . Each interval is associated with a weight w_k representing the expected reward. When an action is chosen and the associated reward is obtained, the weights are updated via an exponential moving average (EMA):

$$w_k^{(t+1)} = \begin{cases} \alpha w_k^t + (1 - \alpha)r_t, & \text{if } \phi_j = \phi^{(k)}, \\ w_k^t, & \text{otherwise.} \end{cases} \quad (5)$$

Here, $\alpha \in [0, 1]$ is the smoothing factor, which balances the influence of past and recent observations. By prioritizing recent data, EMA allows for rapid adaptation to changing conditions while remaining computationally efficient, requiring only constant-time updates per action. Although it does not offer formal convergence guarantees, EMA provides a practical compromise between adaptivity and scalability and allows for maintaining high learning efficiency [39]. Sellers select their next action using Boltzmann (softmax) sampling [56] with a declining temperature τ_t :

$$P_k = \frac{\exp(w_k/\tau_t)}{\sum_{l=1}^K \exp(w_l/\tau_t)}, \quad \tau_t = \max(\tau_{\min}, \tau_0 \cdot \text{decay}^t), \quad (6)$$

where τ_0 sets the initial exploration level, decay controls the annealing rate, and τ_{\min} prevents premature convergence. This ensures exploration dominates early in learning, while the policy gradually becomes more conservative over time. The subsequent action $\phi_j^{(t+1)}$ is drawn at random from the set of actions according to probabilities P_k . This sampling strategy balances exploration and exploitation in a multi-agent, stochastic environment. In our simulations, Boltzmann exploration consistently converged faster and with less oscillation than ϵ -greedy or UCB, which often struggled with noise from multi-agent interactions [40]. By combining a discretized action space, adaptive reward estimates, and Boltzmann sampling, the model captures the decentralized emergence of price formation and allocation dynamics in a spatially distributed market for byproducts. While Boltzmann does not always guarantee optimal regret bounds in multi-agent stochastic settings, its smooth adaptation makes it practical and effective for this scenario.

3.4 Symbiosis Index

We finally introduce a *symbiosis index* (SI) to quantify the efficiency of local byproduct exchanges. It is defined as:

$$\text{SI} = \frac{q_{\text{bought}}}{\min(q_{\text{toSell}}, q_{\text{needed}})}, \quad (7)$$

where q_{bought} denotes the total quantity actually purchased, q_{toSell} the total quantity offered by the sellers, and q_{needed} the total quantity required by the buyers. The denominator normalizes the indicator so that it cannot exceed 1, even when the sold quantity is greater than the buyer’s need. This ensures that the symbiosis index reflects the proportion of demand satisfied rather than rewarding oversupply. In contrast to existing technical indicators [24, 27, 58, 66], which often incorporate numerous parameters (e.g., material compatibility, life cycle impacts, or multi-criteria environmental scores), our simplified formulation deliberately focuses only on exchanged quantities. This avoids unnecessary complexity while remaining sufficient for capturing efficiency in the present modeling framework.

4 EXPERIMENTS

4.1 Simulation Framework

The model is implemented in a simulation framework that supports decentralized markets with heterogeneous agents. Buyers and sellers are spatially distributed across a territory, which can be represented either with an abstract 2D environment or with a real geographic area with road network constraints. Distances, travel times, and transport costs are computed based on the underlying spatial representation, allowing realistic modeling of logistical constraints. The businesses interact through the auction mechanism described earlier, with offers, acceptances, and executed contracts dynamically updating inventories and demands.

The framework tracks a variety of systemic metrics over time, including the local prices and their evolution, the proportion of demand satisfied locally (symbiosis index) and the trading patterns. These indicators form the backbone of our analysis, as they connect local agent decisions to measurable systemic dynamics. This flexibility enables us to investigate a wide range of experimental settings, from dense and highly interconnected territories to sparse environments where exchanges are scarce. Visualization tools provide spatial and temporal representations of market dynamics, price formation, and circularity indicators. The code implementation is open source¹.

4.2 Experimental Conditions

For similar reasons behind our decision to focus on a single product, we decided to focalize on virtual environments. This approach provides precise control over key parameters such as firm density, territorial extent, and spatial clustering, thereby enabling a systematic exploration of their influence on circularity indicators and overall market performance. One of the key factors shaping the potential for cooperation is the spatial dispersion of firms. We therefore define the cluster spread parameter (cs) as the standard deviation of companies’ positions around cluster centers, normalized by the environment width: small values lead to compact, well-separated

¹<https://github.com/matthieu-mastio/MASTIO>

clusters, while setting it to 1 produces a configuration comparable to a fully uniform distribution across the environment, effectively removing any clustering effect. Figure 1 shows the difference between a low (1a) and high (1b) spread value. In our experiments, businesses were allocated into four spatial clusters. By default, the cluster spread parameter was set to 1, indicating a broad dispersion and minimal clustering.

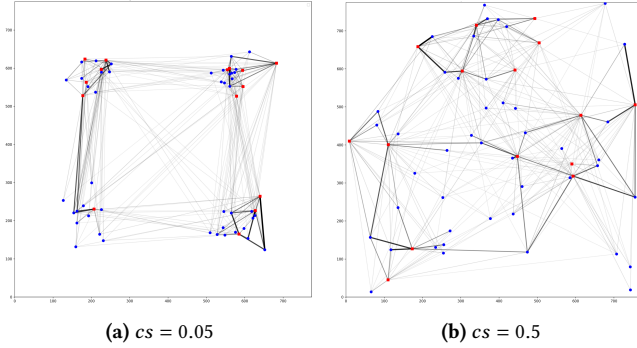


Figure 1: Comparison of spatial agent distributions for different cs values.

The results we obtained came from simulations conducted with a population of 40 firms, representing the typical scale of a standard EIP. The market price was fixed at 100, and the transportation cost was set to 0.1 per kilometer. This transport cost is a normalized parameter relative to the market price, calibrated to ensure that spatial distance acts as a meaningful economic friction without completely blocking trade, a choice further supported by the sensitivity analysis in Section 5.3. Sellers adjusted their strategies through Boltzmann exploration, with a temperature decay parameter set to 0.996 and the action space discretized into 30 bins. The relatively small scale of the simulated market was chosen to balance computational tractability with the ability to capture complex multi-agent interactions. This size of population allows us to conduct thousands of simulation runs for systematic sensitivity analyses and counterfactual experiments, which would be computationally prohibitive at larger scales. Furthermore, each simulation runs for a large number of timesteps (1000 steps) to ensure that seller strategies have sufficient time to adapt and converge, and that emergent patterns of price formation and allocation stabilize. By prioritizing repeated experimentation and long-horizon dynamics over absolute market size, we focus on understanding the fundamental mechanisms driving spatial effects, adaptive bidding behavior, and the emergence of circularity. Our implementation is optimized for performance, allowing on a standard laptop processor (Intel(R) Core(TM) Ultra 9 185H processor with 21 cores) to achieve approximately 100 simulations of 1000 timesteps per minute.

5 RESULTS

In this section, we examine the core dynamics of the decentralized, spatially explicit market, focusing on conditions for price convergence and the efficiency of adaptive strategies using counterfactual regret analysis. We then explore how spatial organization, resource scarcity, disposal costs, and density interact to shape market outcomes and local circularity. Finally, a global sensitivity analysis

quantifies the influence of individual parameters and their interactions on emergent symbiosis patterns.

5.1 Convergence Toward Equilibrium

We first explore empirically the evolution of transaction prices as sellers adapt their strategies under different scarcity and disposal cost conditions, with a density fixed to 0.001. Figures 2a and 2b show the mean price trajectories over time compared to the external market price and the disposal cost. As anticipated given the decaying temperature, the curves gradually stabilize, approaching an equilibrium after sufficient simulation steps. In the high scarcity and low disposal cost setting (Figure 2a), prices converge toward an equilibrium close to the external market price. By contrast, in the low scarcity and medium disposal cost scenario (Figure 2b), prices decline steadily and stabilize near the disposal cost, reflecting buyers’ stronger bargaining power and the reduced competitive pressure on resource demand.

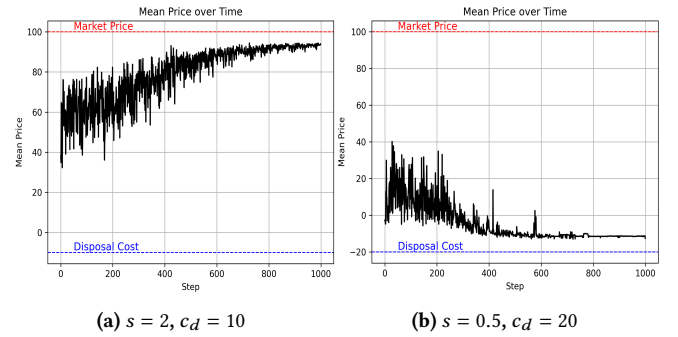


Figure 2: Evolution of mean transaction prices under different scarcity and disposal cost conditions, where s denotes the scarcity ratio and c_d the disposal cost.

To better understand the efficiency and stability of the sellers’ adaptive strategies, we next examine the dynamics at the level of individual firms using counterfactual regret analysis [67], which quantifies how much a seller could have improved its payoff by choosing alternative actions.

Let $\phi_j^{(t)} \in \Phi_j$ denote the action used by seller j at timestep t , and $r_j^{(t)}(\phi_j^{(t)})$ denote the reward received at that step. The *optimal reward* at time t is

$$r_j^{*(t)} = \max_{\phi \in \Phi_j} r_j^{(t)}(\phi), \tag{8}$$

where $r_j^{(t)}(\phi)$ is the reward that the seller j would have obtained at step t if it had chosen action ϕ while all other sellers’ actions remained unchanged. The *regret* of seller j at time t is then

$$\text{Regret}_j^{(t)} = r_j^{*(t)} - r_j^{(t)}(\phi_j^{(t)}). \tag{9}$$

A positive regret indicates that the seller could have achieved a higher payoff at that step by selecting a different action, whereas a regret of zero indicates that the chosen action was optimal given the choices of the other sellers at that moment.

Empirically, we observe that the per-step total regret, the aggregate deviation from a joint optimum, declines over time and converges toward zero in both scenarios (Figure 3). Since per-step

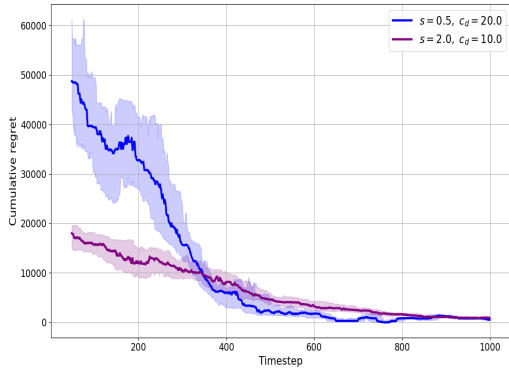


Figure 3: Rolling median ($window = 50$) of per-step regret under different scarcity and disposal cost conditions.

regret measures the immediate payoff loss relative to the best fixed action in hindsight at that step, a vanishing regret implies that sellers’ strategies are nearly optimal given the current actions of others. Consequently, the joint strategy profile of all sellers appears to converge toward an approximate *Nash equilibrium*, providing justification for the stability of the emergent market dynamics under our MARL learning scheme.

5.2 Price and Symbiosis Dynamics

In this section, we examine how firm density, resource scarcity, and disposal costs shape the dynamics of transaction prices and circularity. We focus on these output metrics because they directly quantify the level of local circularity achieved in the market and the average exchange price, making them relevant proxies for assessing the effectiveness of different policy levers. By systematically varying firm density, resource scarcity, and disposal costs, we analyze how structural and economic factors influence sellers’ adaptive strategies, shaping both the emergence of local symbiosis and overall market outcomes. Figures 4 and 5 illustrate the evolution of equilibrium prices and local symbiosis levels as a function of disposal cost, scarcity, and density. Each point is the mean of 10 simulation runs, with shaded areas indicating one standard deviation to visualize variability across runs.

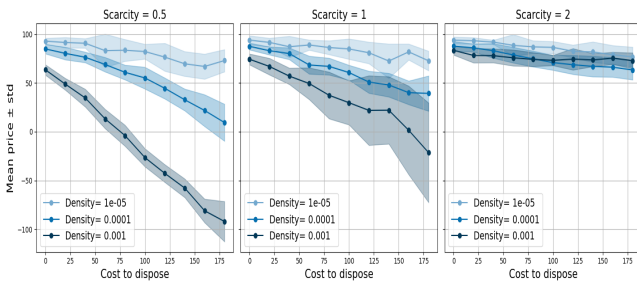


Figure 4: Evolution of the exchange price

On the price side (Figure 4), we observe that, as expected, higher disposal costs induce sellers to lower their prices in order to avoid the landfill penalty. This effect is strongest in low scarcity environments ($s = 0.5$), where intense competition can even drive equilibrium prices below zero in dense territories. At intermediate

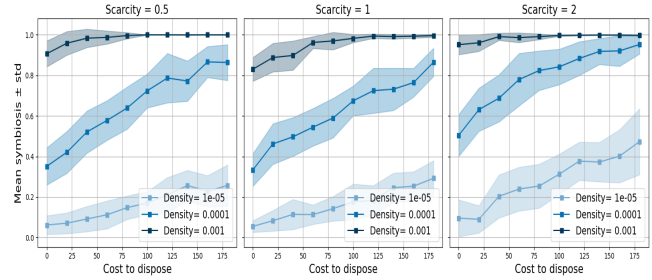


Figure 5: Evolution of symbiosis index

scarcity ($s = 1$), the downward trend remains but is attenuated, as sellers retain partial market power. Under high scarcity ($s = 2$), prices remain largely insensitive to c_d : the excess demand ensures that buyers continue to pay close to the outside option p_m , leaving sellers with little incentive to adjust their strategies. Interestingly, at the demand equilibrium ($s = 1$), we observe higher variance, reflecting the instability and potential oscillations in companies’ strategies as they adapt to competitive pressures.

Symbiosis levels (Figure 5) exhibit a complementary pattern: the share of demand met through local exchanges increases with c_d , as the disposal penalty strengthens the incentive to form local contracts. Here, density plays a critical role: in dense environments, local exchanges dominate rapidly, and symbiosis saturates close to one; in sparse environments, by contrast, symbiosis remains low even under high c_d , reflecting structural limitations in firms’ connectivity. We observe that this time scarcity has only a marginal effect on the evolution of symbiosis: higher scarcity slightly enhances local exchanges, but the dominant factors remain disposal cost and firm density.

5.3 Sensitivity Analysis

The previous analysis showed that market outcomes are shaped by interacting mechanisms, but it did not allow us to fully disentangle their relative influence. To address this limitation and account for the influence of additional variables, we conducted a variance-based global sensitivity analysis [50].

We evaluate the influence of the simulation parameters on the symbiosis index and exchanged prices measured at the end of each simulation run after 1000 timesteps. By analyzing which parameters exert the greatest influence on the price and symbiosis index, we can identify the factors that policymakers should prioritize when designing interventions to foster circularity while keeping prices at a low level. We simulate varying values for the following five variables:

- **Disposal cost** ($c_d \in [0, 200]$), representing the penalty for unsold byproducts;
- **Scarcity level** ($s \in [0.25, 2]$), controlling the relative abundance of resources in the market;
- **Firm density** ($\rho \in [10^{-5}, 10^{-1}]$), defining the number of firms per unit area;
- **Cluster spread** ($cs \in [0, 0.5]$), determining how tightly companies are spatially grouped within clusters;
- **Transportation cost per kilometer** ($c_t \in [0, 10]$), capturing the economic friction associated with distance.

To efficiently explore the parameter space, we employed surrogate models, which act as fast emulators of the full simulation while capturing the key input-output relationships [52]. Specifically, we implemented a sparse Polynomial Chaos Expansion (PCE), enabling the analytical derivation of Sobol’ indices from its coefficients [55]. The surrogate models demonstrated strong predictive performance, achieving an R^2 of 95% for the symbiosis and 92% for the price on a 20% test sample. Parameter sampling was performed using Latin hypercube sampling to train surrogate models, with s sampled on a \log_2 scale and $density$ on a \log_{10} scale. For each parameter set, two independent simulations were executed to account for stochastic variability in agent decisions, resulting in a total of 20,000 independent simulations. Sobol’ sensitivity analysis decomposes the variance of the model output into orthogonal contributions attributable to individual input parameters and their interactions. We computed the full set of Sobol’ indices: first-order indices quantify the direct effect of each parameter on output variance, while higher-order indices capture interaction effects among parameters [8]. Taken together, Sobol’ indices provide a rigorous and comprehensive assessment of the relative importance of independent inputs and their interactions, with their sum accounting for 100% of the variance.

In Table 1, out of 31, we show only the 12 indices greater than 1% of the total variance for either symbiosis or price indices. The Sobol’ analysis confirms the strong influence of firm density: higher density reduces the average distance between agents and intensifies competition, explaining its impact on both model outcomes. As expected, price is also strongly affected by scarcity, and both outputs are influenced by transport costs. Interestingly, higher-order interactions contribute substantially to the variance, emphasizing the importance of interplay effects between parameters.

Table 1: Comparison of Sobol’ sensitivity indices for the symbiosis and price outputs, including unaccounted variance from Sobol’ indices < 1%.

Variables	Price Sobol’	Symbiosis Sobol’
ρ	$S_1 = 20.5\%$	$S_1 = 63.2\%$
s	$S_1 = 24.6\%$	$S_1 = 1.3\%$
c_t	$S_1 = 8.5\%$	$S_1 = 16.7\%$
s, ρ	$S_2 = 14.0\%$	$S_2 \approx 0$
c_d	$S_1 = 11.0\%$	$S_1 = 2.4\%$
ρ, c_t	$S_2 \approx 0$	$S_2 = 7.5\%$
s, c_t	$S_2 = 7.0\%$	$S_2 \approx 0$
cs	$S_1 = 1.5\%$	$S_1 = 3.8\%$
c_d, s	$S_2 = 4.6\%$	$S_2 \approx 0$
c_d, ρ	$S_2 = 1.9\%$	$S_2 \approx 0$
ρ, cs	$S_2 \approx 0$	$S_2 = 2.2\%$
c_d, s, ρ	$S_3 = 1.7\%$	$S_3 \approx 0$
Unaccounted	$\approx 4.7\%$	$\approx 3.0\%$

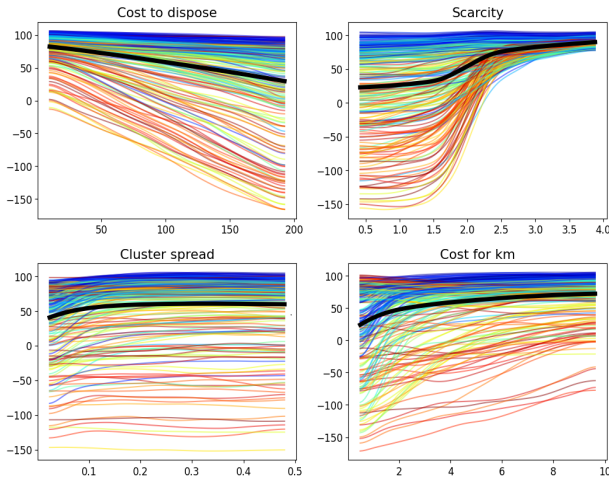
Although firm density emerges as the most influential factor for both symbiosis and price, its overwhelming effect tends to dominate the variance and obscure the individual contributions of the other parameters. To better isolate and interpret the impact of the potentially actionable economic and spatial levers, we therefore fix

density at different levels and focus the subsequent analysis on the remaining four input variables. Therefore, in the subsequent analysis, we focus on the 4 other input variables under different fixed density scenarios, since their interaction effects with density are all significant even at a third order. To better capture the non-linear effects, we replace the sparse PCE by a Multi-Layer Perceptron neural network surrogate model (R^2 of 97% for the symbiosis and 95% for the price), and use it to predict both the Partial Dependence Plots (PDP) and Individual Conditional Expectations (ICE) [46] as illustrated in Figure 6. In these plots, each colored line corresponds to an ICE, showing how the outcome of a single simulation responds when varying one parameter while holding others fixed; the black line is the PDP, representing the average effect across all simulations, and the color of each ICE line encodes the underlying firm density.

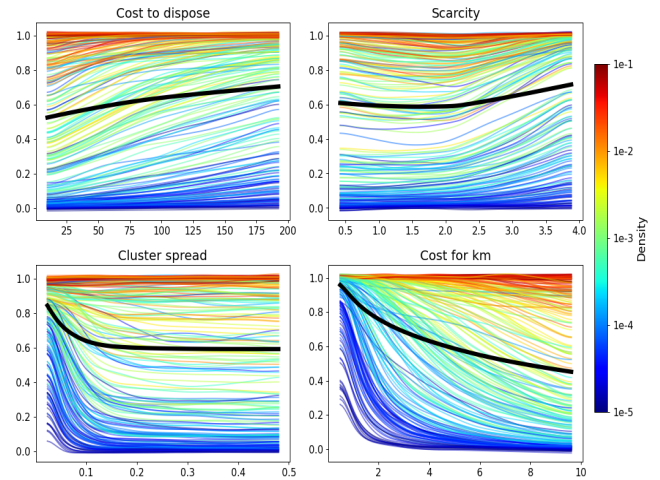
First, for the price in Figure 6a, the PDP and ICE plots are consistent with the Sobol’ analysis in Table 1. In the blue low-density cases, the exchange price remains close to the market price and is only weakly influenced by other variables. As density increases, however, the price tends to decrease. The most influential variable is Scarcity, which exhibits a threshold effect: at low levels, scarcity has little impact, but between 1.5 and 2.5 it strongly drives the price upward. Beyond 2.5, the effect persists but grows at a slower rate. We also observe that increasing the disposal cost provides a strong and linear incentive for price reduction; the higher the density, the stronger the effect. Finally, cluster spread and kilometer cost have more limited influence, suggesting that they are less effective levers for intervention on the price outcome.

Second, for circularity in Figure 6b, we again observe a pronounced density effect: higher density leads to greater symbiosis, largely independent of other variables. At low density levels, cluster spread emerges as an important factor, but only when it takes on small values. As spread increases, its influence fades rapidly, suggesting that spatial compactness can partially compensate for low density but only within narrow limits. Reducing travel costs appears more consistently actionable: at low density, for example, a per-kilometer cost below 2 is required to reach any desirable symbiosis value. Under the same low density conditions, increasing disposal costs also provide a clear incentive to exchange byproducts, while at high density its effect is only marginal. Finally, scarcity has a notable effect on low-to-mid density values only above the critical threshold of 2.5.

In summary, firm density is a primary determinant of both symbiosis and price, though it is mostly exogenous and difficult to influence directly through policy. Among the controllable parameters, transport cost (c_t) and disposal penalty (c_d) emerge as effective levers. The sensitivity results indicate that reducing per-kilometer transport costs, through measures such as shared logistics or targeted subsidies, can substantially enhance exchange opportunities in low-density areas. Similarly, higher disposal penalties consistently incentivize firms to seek symbiotic exchanges instead of wasteful disposal.



(a) PDP and ICE illustrating the factors influencing the price index



(b) PDP and ICE illustrating the factors influencing the symbiosis index

Figure 6: PDP and ICE of c_d , s , cd and c_t colored by fixed densities for price and symbiosis.

6 DISCUSSION

Our results demonstrate that the proposed mechanism generates stable, economically grounded behaviors. Emergent pricing patterns reveal that sellers adapt efficiently to local conditions, while counterfactual regret analysis confirms convergence toward a near-Nash equilibrium rather than stochastic oscillation. Furthermore, global sensitivity analysis establishes clear causal links between model parameters and market outcomes. Significant high-order interactions suggest that circularity arises from the synergy of spatial configuration and economic incentives, highlighting the necessity for integrated, multivariate policy frameworks. Ultimately, these findings indicate that decentralized adaptive learning produces rational outcomes without centralized clearing, validating the mechanism’s robustness.

Beyond methodology, the model offers a framework for evaluating policy instruments by mapping simulation parameters to regulatory levers. For instance, increasing disposal costs mimics landfill taxes, while the cluster spread parameter can be interpreted as a proxy for industrial zoning or policies that promote firm density. This mapping enables the simulator to function as a virtual laboratory, allowing stakeholders to perform “what-if” analyzes and compare interventions before real-world deployment.

While these insights highlight the applicability of our model, we must recognize several limitations that temper the generalizability of the results. We focused on a single byproduct to isolate the core mechanisms of decentralized exchange, abstracting away product heterogeneity and multi-market interdependence. Likewise, to ensure control and replicability, our experiments relied on stylized virtual geographies rather than empirically calibrated territories. We also abstract from temporal dynamics (production cycles, storage) and agent memory (trust, long-term contracts). Nonetheless, this work provides an indispensable first step toward systematically modeling spatially constrained markets for industrial symbiosis, upon which richer and more realistic frameworks can be built.

Future work will address these limitations by introducing multi-product environments, temporal dynamics, and agent memory to capture substitution effects and relational contracts. Grounding simulations in empirical case studies will further validate predictions and refine the policy decision tool. Finally, we also plan to include rigorous runtime benchmarking to formally showcase the simulator’s scalability.

7 CONCLUSION

We introduced a decentralized multi-agent model for local byproduct exchanges, capturing how sellers adapt prices through RL, based on spatial and economic constraints. Simulations exhibit that adaptive decentralized strategies can produce stable circular outcomes. Our results highlight two complementary contributions. First, from a *policy design perspective*, the simulator is the initial step toward a decision support tool. By linking agent-level parameters to macro-level outcomes such as the symbiosis index, the model allows policymakers to identify which levers most strongly affect circularity. This provides a controlled environment where different regulatory scenarios can be tested before real-world implementation. The emergent insights from our simulations help policymakers prioritize interventions that effectively promote circular practices across heterogeneous firms and contexts. Second, from a *MAS perspective*, the work provides a new methodological understanding of decentralized exchange mechanisms under spatial and economic constraints. We show how RL agents, interacting in a double auction market with heterogeneous resources and transport costs, can converge to stable trading patterns. Beyond the application domain, our contribution lies in revealing how market based coordination and learning dynamics shape emergent equilibrium.

ACKNOWLEDGMENTS

This work was supported by the MUTTEC project (France 2030 – TIRIS, contract 23-AAP-TIRIS-01-062) and by the MIMICO project (ANR-24-CE23-0380).

REFERENCES

- [1] Biao An, Victor Lesser, and Kwang Mong Sim. 2011. Strategic Agents for Multi-Resource Negotiation. *Autonomous Agents and Multi-Agent Systems* 23, 1 (2011), 114–153.
- [2] Bilge Atasoy, Frederik Schulte, and Alex Steenkamp. 2020. Platform-based collaborative routing using dynamic prices as incentives. *Transportation Research Record* 2674, 10 (2020), 670–679.
- [3] Moshe Babaioff, Noam Nisan, and Elan Pavlov. 2005. Mechanisms for a spatially distributed market. In *Proceedings of the 4th ACM Conference on Electronic Commerce*. Association for Computing Machinery, San Diego CA, USA, 9–20.
- [4] Martin Bichler, Julius Durmann, and Matthias Oberlechner. 2025. Online Optimization Algorithms in Repeated Price Competition: Equilibrium Learning and Algorithmic Collusion. <https://arxiv.org/abs/2412.15707>
- [5] Aleksander Bjerkoey and Mikael Kvalvaer. 2019. *Replicating Financial Markets using Reinforcement Learning: An Agent-Based Approach*. Master's thesis. NTNU.
- [6] Kai Cao, Xiao Feng, and Hui Wan. 2009. Applying agent-based modeling to the evolution of eco-industrial systems. *Ecological Economics* 68, 11 (2009), 2868–2876.
- [7] Ghayda Chahla and Amine Zoughaib. 2019. Agent-Based Conceptual Framework for Energy and Material Synergy Patterns in a Territory with Non-Cooperative Governance. *Computers & Chemical Engineering* 131 (2019), 106596.
- [8] Gaelle Chastaing, Fabrice Gamboa, and Clémentine Prieur. 2015. Generalized Sobol' sensitivity indices for dependent variables: numerical methods. *Journal of statistical computation and simulation* 85, 7 (2015), 1306–1333.
- [9] Anthony Chavez and Pattie Maes. 1997. Kasbah: An Agent Marketplace for Buying and Selling Goods. In *Proceedings of the 1st International Conference on Autonomous Agents*. Association for Computing Machinery, New York, NY, USA, 8–12.
- [10] Siqi Chen and Gerhard Weiss. 2015. Toward Efficient Agreements in Real-Time Multilateral Agent-Based Negotiations. In *Proceedings of the 27th IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE Computer Society, Vietri sul Mare, Italie, 896–903.
- [11] Xinyi Chen, Miaoxin Dong, Long Zhang, Xiaoyu Luan, Xiaowei Cui, and Zhaojie Cui. 2022. Comprehensive evaluation of environmental and economic benefits of industrial symbiosis in industrial parks. *Journal of Cleaner Production* 354 (2022), 131635.
- [12] Marian Chertow. 2000. Industrial symbiosis: literature and taxonomy. *Annual Review of Energy and the Environment* 25, 1 (2000), 313–337.
- [13] Marian Chertow and John Ehrenfeld. 2012. Organizing self-organizing systems: Toward a theory of industrial symbiosis. *Journal of Industrial Ecology* 16, 1 (2012), 13–27.
- [14] Leon Chu and Zuo-Jun Shen. 2008. Truthful double auction mechanisms. *Operations research* 56, 1 (2008), 102–120.
- [15] David Cliff. 2024. Parameterised Response Zero Intelligence Traders. *Journal of Economic Interaction and Coordination* 19 (2024), 439–492.
- [16] Melissa Demartini, Flavio Tonelli, and Kannan Govindan. 2022. An Investigation into Modelling Approaches for Industrial Symbiosis: A Literature Review and Research Agenda. *Cleaner Logistics and Supply Chain* 3 (2022), 100020.
- [17] Teresa Doménech and Michael Davies. 2009. The social aspects of industrial symbiosis: The application of social network analysis to industrial symbiosis networks. *Progress in Industrial Ecology, An International Journal* 6, 1 (2009), 68–99.
- [18] Joshua Epstein and Robert Axtell. 1996. *Growing artificial societies: social science from the bottom up*. Brookings Institution Press, Washington, DC, USA.
- [19] Marilia Felicio, Debora Amaral, Karine Esposto, and Xavier G. Durany. 2016. Industrial Symbiosis Indicators to Manage Eco-Industrial Parks as Dynamic Systems. *Journal of Cleaner Production* 118 (2016), 54–64.
- [20] Hugo Fernandez-Mena, Benoit Gaudou, Simon Pellerin, Gavin K. MacDonald, and Thomas Nesme. 2020. Flows in Agro-Food Networks (FAN): An Agent-Based Model to Simulate Local Agricultural Material Flows. *Agricultural Systems* 180 (2020), 102718.
- [21] Tatiana Filatova, Dawn C Parker, and Anne van der Veen. 2007. Agent-based land markets: heterogeneous agents, land prices and urban land use change. In *4th Annual Conference of the European Social Simulation Association, ESSA 2007*. The European Social Simulation Association (ESSA), Toulouse, France, 263–275.
- [22] Vincent Gan and Jackie Cheng. 2015. Formulation and Analysis of Dynamic Supply Chain of Backfill in Construction Waste Management Using Agent-Based Modeling. *Advanced Engineering Informatics* 29, 4 (2015), 878–888.
- [23] Sumitra Ganesh, Nelson Vadori, Mengda Xu, Hua Zheng, Prashant Reddy, and Manuela Veloso. 2019. Reinforcement Learning for Market Making in a Multi-agent Dealer Market. <https://arxiv.org/abs/1911.05892>
- [24] Xiao Gao, Rui Li, and Ran Li. 2013. Study on byproducts recycling in eco-industrial parks. *Advanced Materials Research* 788 (2013), 288–292.
- [25] Mohamed Raouf Ghali, Jean-Marc Frayret, and Chahid Ahabchane. 2017. Agent-based model of self-organized industrial symbiosis. *Journal of cleaner production* 161 (2017), 452–465.
- [26] S. Gulati and D. Kellenberg. 2013. Political Economy and the Interaction Between International Trade and Environmental Policies. In *Encyclopedia of Energy, Natural Resource, and Environmental Economics*, Jason F. Shogren (Ed.). Elsevier, Waltham, 288–295. <https://doi.org/10.1016/B978-0-12-375067-9.00020-6>
- [27] Catherine Hardy and Thomas E Graedel. 2002. Industrial ecosystems as food webs. *Journal of Industrial Ecology* 6, 1 (2002), 29–38.
- [28] Xue-Zhong He, Youwei Li, and Min Zheng. 2019. Heterogeneous agent models in financial markets: A nonlinear dynamics approach. *International Review of Financial Analysis* 62 (2019), 135–149.
- [29] Juan Henriques, Paulo Ferrão, Rui Castro, and João Azevedo. 2021. Industrial symbiosis: A sectoral analysis on enablers and barriers. *Sustainability* 13, 4 (2021), 1723.
- [30] Harold Hotelling. 1929. Stability in Competition. *The Economic Journal* 39, 153 (1929), 41.
- [31] Michaël Karpe, Jin Fang, Zhongyao Ma, and Chen Wang. 2020. Multi-agent reinforcement learning in a realistic limit order book market simulation. In *Proceedings of the First ACM International Conference on AI in Finance*. Association for Computing Machinery, New York, NY, USA, 1–7.
- [32] Alfred Lambert and Frank Boons. 2002. Eco-industrial parks: stimulating sustainable development in mixed industrial parks. *Technovation* 22, 8 (2002), 471–484.
- [33] Marcelo S. Lauretto, Barbara B. C. Silva, and Pablo M. Andrade. 2013. Evaluation of a Supervised Learning Approach for Stock Market Operations. <https://arxiv.org/abs/1301.4944>
- [34] Andrew Ledvina and Ronnie Sircar. 2011. Dynamic Bertrand oligopoly. *Applied Mathematics & Optimization* 63, 1 (2011), 11–44.
- [35] Tobias Lorenz and Andreas Jost. 2006. Towards an orientation framework in multi-paradigm modeling. In *Proceedings of the 24th International Conference of the System Dynamics society*. System Dynamics Society Albany, Nijmegen, The Netherlands, 2134–2151.
- [36] Johann Lussange, Ivan Lazarevich, Sacha Bourgeois-Gironde, Stefano Palminteri, and Boris Gutkin. 2021. Modelling stock markets by multi-agent reinforcement learning. *Computational Economics* 57, 1 (2021), 113–147.
- [37] Iwao Maeda, David DeGraw, Michiharu Kitano, Hiroyasu Matsushima, Hiroki Sakaji, Kiyoshi Izumi, and Atsuo Kato. 2020. Deep reinforcement learning in agent based financial market simulation. *Journal of Risk and Financial Management* 13, 4 (2020), 71.
- [38] Matthieu Mastio and Benoit Gaudou. 2025. Modélisation multi-agents d'un marché local de flux résiduels. In *Congrès Interdisciplinaire sur l'Economie Circulaire 2025 (CIEC 2025)*. AIFREC. Amiens, France, online.
- [39] Daniel Morales-Brotons, Thijs Vogels, and Hadrien Hendrikx. 2024. Exponential moving average of weights in deep learning: Dynamics and benefits. *Transactions on Machine Learning Research Journal* online (2024), 1–27.
- [40] Brendan O'Donoghue. 2021. Variational bayesian reinforcement learning with regret bounds. *Advances in Neural Information Processing Systems* 34 (2021), 28208–28221.
- [41] Bin Qiao, Shenle Pan, and Eric Ballot. 2019. Dynamic pricing model for less-than-truckload carriers in the Physical Internet. *Journal of Intelligent Manufacturing* 30, 7 (2019), 2631–2643.
- [42] Dawei Qiu, Jianhong Wang, Junkai Wang, and Goran Strbac. 2021. Multi-Agent Reinforcement Learning for Automated Peer-to-Peer Energy Trading in Double-Side Auction Market. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*. International Joint Conferences on Artificial Intelligence, Montréal, Canada, 2913–2920.
- [43] Qingling Qu, Xin Qian, and Jin Wang. 2010. A system dynamics model for the environment management of eco-industrial park. In *Proceedings of the 4th International Conference on Bioinformatics and Biomedical Engineering*. Association for Computing Machinery, Chengdu, China, 1–4.
- [44] Juste Raimbault, Julian Broere, Matthieu Somveille, Jose Serna, Erik Strömböm, Christopher Moore, Barry Zhu, and Lydia Sugar. 2020. A Spatial Agent Based Model for Simulating and Optimizing Networked Eco-Industrial Systems. *Resources, Conservation and Recycling* 155 (2020), 104538.
- [45] Massimiliano Rizzati and Matteo Landoni. 2024. A systematic review of agent-based modelling in the circular economy: Insights towards a general model. *Structural Change and Economic Dynamics* 69 (2024), 617–631.
- [46] Mohammad Robani, Paul Saves, Pramudita Palar, Lavi Zuhail, and Joseph Morlier. 2025. SMT-EX: An Explainable Surrogate Modeling Toolbox for Mixed-Variables Design Exploration. In *AIAA SCITECH 2025 Forum*. American Institute of Aeronautics and Astronautics, Orlando, Florida, USA, 0777.
- [47] Jean-Paul Rodrigue. 2020. *The geography of transport systems*. Routledge, New York, USA.
- [48] Elena Romero and Carmen Ruiz. 2013. Framework for applying a complex adaptive system approach to model the operation of eco-industrial parks. *Journal of Industrial Ecology* 17, 5 (2013), 731–741.
- [49] Elena Romero and Carmen Ruiz. 2014. Proposal of an Agent-Based Analytical Model to Convert Industrial Areas in Industrial Eco-Systems. *Science of The Total Environment* 468-469 (2014), 394–405.
- [50] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. 2010. Variance based sensitivity analysis of model

- output. Design and estimator for the total sensitivity index. *Computer physics communications* 181, 2 (2010), 259–270.
- [51] Mark Satterthwaite, Steven Williams, and Konstantinos Zachariadis. 2022. Price Discovery Using a Double Auction. *Games and Economic Behavior* 131 (2022), 57–83.
- [52] Paul Saves, Rémi Lafage, Nathalie Bartoli, Youssef Diouane, Jasper Bussemaker, Thierry Lefebvre, John Hwang, Joseph Morlier, and Joaquim Martins. 2024. SMT 2.0: A Surrogate Modeling Toolbox with a focus on hierarchical and mixed variables Gaussian processes. *Advances in Engineering Software* 188 (2024), 103571.
- [53] Ali Shavandi and Majid Khedmati. 2022. A multi-agent deep reinforcement learning framework for algorithmic trading in financial markets. *Expert Systems with Applications* 208 (2022), 118124.
- [54] John Sterman. 2000. *Business dynamics: systems thinking and modeling for a complex world*. Irwin McGraw-Hill, Columbus, USA.
- [55] Bruno Sudret. 2008. Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety* 93, 7 (2008), 964–979.
- [56] Richard Sutton and Andrew Barto. 1998. *Reinforcement learning: An introduction*. Vol. 1. MIT press, Cambridge, USA.
- [57] Gerald Tesauro and Rajarshi Das. 2001. High-Performance Bidding Agents for the Continuous Double Auction. In *Proceedings of the 3rd ACM Conference on Electronic Commerce*. Association for Computing Machinery, Tampa Florida, USA, 206–209.
- [58] Dai Tiejun. 2010. Two quantitative indices for the planning and evaluation of eco-industrial parks. *Resources, Conservation and Recycling* 54, 7 (2010), 442–448.
- [59] Wouter Van Heeswijk. 2019. Smart containers with bidding capacity: A policy gradient algorithm for semi-cooperative learning. In *Lecture Notes in Computer Science*, E. Lalla-Ruiz, M. R. K. Mes, and S. Voß (Eds.), Vol. 12433. Springer, Cham, 469–491.
- [60] Andrew Wainwright and Paolo Zeppini. 2024. Social Tipping Points: from Polarisation to Sustainability Transitions. <https://ssrn.com/abstract=5101001>
- [61] Yingfei Wang, Juliana Martins Do Nascimento, and Warren Powell. 2019. Reinforcement Learning for Dynamic Bidding in Truckload Markets: an Application to Large-Scale Fleet Management with Advance Commitments. <https://arxiv.org/abs/1802.08976>
- [62] Devrim Yazan and Luca Fraccascia. 2020. Sustainable Operations of Industrial Symbiosis: An Enterprise Input-Output Model Integrated by Agent-Based Simulation. *International Journal of Production Research* 58, 2 (2020), 392–414.
- [63] Devrim Yazan, Vincenzo Alessio Romano, and Vito Albino. 2016. The design of industrial symbiosis: an input–output approach. *Journal of cleaner production* 129 (2016), 537–547.
- [64] Chang Yu, Martin de Jong, and Gerard Dijkema. 2013. Process analysis of eco-industrial park development: the case of Tianjin, China. *Journal of Cleaner Production* 64 (2013), 464–477.
- [65] Liteng Zha, Yafeng Yin, and Yuchuan Du. 2017. Surge pricing and labor supply in the ride-sourcing market. *Transportation Research Procedia* 23 (2017), 2–21.
- [66] Yan Zhao, Jin-cheng Shang, Chong Chen, and He-nan Wu. 2008. Simulation and evaluation on the eco-industrial system of Changchun economic and technological development zone, China. *Environmental Monitoring and Assessment* 139 (2008), 339–349.
- [67] Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. 2007. Regret minimization in games with incomplete information. *Advances in neural information processing systems* 20 (2007), 1–8.