Leveraging Transfer Learning to Improve Convergence in All-Pay Auctions

Luis Eduardo Craizer[1][a], Edward Hermann[1][b] and Moacyr Alvim Silva[2][c]

[1] *Pontifícia Universidade Católica, 22451-900, Rio de Janeiro, RJ, Brazil*

[2] *Fundação Getulio Vargas, 22250-145, Rio de Janeiro, RJ, Brazil*

Keywords: Transfer Learning, Auction Theory, Nash Equilibrium, Deep Reinforcement Learning, Multi-Agent, All-Pay.

Abstract: In previous research on Multi-Agent Deep Deterministic Policy Gradient (MADDPG) in All-Pay Auctions, we identified a key limitation: as the number of agents increases, the tendency for some agents to bid 0.0 — resulting in local equilibrium — grows, leading to suboptimal bidding strategies. This issue diminishes the effectiveness of traditional reinforcement learning in large, complex auction environments. In this work, we propose a novel transfer learning approach to address this challenge. By training agents in smaller N auctions and transferring their learned policies to larger N settings, we significantly reduce the occurrence of local equilibrium. This method not only accelerates training but also enhances convergence toward optimal Nash equilibrium strategies in multi-agent settings. Our experimental results show that transfer learning successfully overcomes the limitations observed in previous research, yielding more robust and efficient bidding strategies in all-pay auctions.

1 INTRODUCTION

In multi-agent all-pay auctions, agents are tasked with bidding strategies that maximize their expected payoffs in a highly competitive environment.[1] Previous research utilizing Multi-Agent Deep Deterministic Policy Gradient (MADDPG) has demonstrated that agents can converge to Nash equilibrium in smaller auctions with fewer participants (Craizer et al., 2025). However, as the number of agents increases, a phenomenon of local equilibrium often emerges, where certain agents bid optimally while others fall into simplistic strategies, such as bidding 0.0, effectively opting out of competition. This behavior undermines the strategic complexity of the auction and leads to suboptimal overall outcomes.

The difficulty of achieving global equilibrium increases exponentially with the number of agents, as the strategy space becomes more complex. This presents a significant challenge in ensuring that agents learn optimal bidding behaviors in larger auctions, particularly when initializing from a random state. To address this issue, we propose leveraging transfer learning as a means to improve training in these high-dimensional environments. Specifically, we train agents in lower-N auctions, where convergence to Nash equilibrium is more feasible, and then transfer the learned models and parameters to auctions with a greater number of agents. While we initially attempted a direct approach to larger auctions, applying the transfer process stepwise, gradually increasing the number of agents, proved to yield better results.

The primary contributions of this work are twofold. First, we propose a transfer learning framework designed for multi-agent all-pay auctions, which effectively mitigates the emergence of local equilibrium in higher-N settings. By leveraging pretrained models from simpler auction scenarios, we significantly improve the scalability of the training process, enabling robust convergence to near-Nash equilibrium strategies. Second, we adapt the critic network architecture to accommodate the increased complexity of larger agent populations, ensuring the model's effectiveness in higher-dimensional environments. These contributions lay the groundwork for extending transfer learning techniques to other complex auction formats and multi-agent systems, offering new insights into strategic decision-making in competitive scenarios.

[a] https://orcid.org/0009-0001-5112-2679
[b] https://orcid.org/0000-0002-4999-7476
[c] https://orcid.org/0000-0001-6519-1264

[1] This section was revised for grammar and wording with assistance from ChatGPT-3.

Craizer, L. E., Hermann, E. and Silva, M. A.
Leveraging Transfer Learning to Improve Convergence in All-Pay Auctions.
DOI: 10.5220/0013294000003929
In *Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025) - Volume 1*, pages 534-543
ISBN: 978-989-758-749-8; ISSN: 2184-4992

2 RELATED WORK

Deep Reinforcement Learning (DRL), an approach that integrates deep learning with reinforcement learning principles, enables agents to learn decision-making strategies through cumulative reward maximization in an environment, largely without explicit supervision (Sutton, 2018). DRL's effectiveness has been propelled by major contributions from OpenAI and DeepMind, whose development of environments like Gymnasium and pioneering models such as DQN (Mnih et al., 2015), AlphaZero (Schrittwieser et al., 2020), A3C (Mnih, 2016), and PPO (Schulman et al., 2017) have significantly advanced the field. With the rise of multi-agent reinforcement learning (MARL), algorithms such as MADDPG and MAPPO have been developed to manage the challenges of non-stationary and partially observable environments, making these approaches highly applicable to competitive and cooperative multi-agent scenarios.

In auction theory, DRL has become a prominent tool for simulating and understanding strategic behaviors in complex auction types. Recent studies by Kannan (Kannan et al., 2019) and Luong et al. (Luong et al., 2018) use agent-based simulations powered by DRL to analyze human decision-making patterns within auction frameworks. Gemp's research explores DRL application in all-pay auctions, focusing on scenarios where traditional equilibrium analysis is computationally infeasible (Gemp et al., 2022). Moreover, Dütting (Dütting et al., 2021) and Feng advance auction models by employing neural networks to bridge theoretical gaps in expected and observed outcomes, notably in multi-item auction settings.

Relevant to our study are the contributions by Bichler, whose Neural Pseudo-Gradient Ascent (NPGA) algorithm offers innovative ways to estimate equilibrium in symmetric auctions, especially within all-pay environments (Bichler et al., 2021). Bichler's work highlights the potential for DRL algorithms to identify and approximate equilibrium strategies in auctions lacking explicit equilibrium formulas, underscoring the robustness of DRL for analyzing complex auction formats (Ewert et al., 2022). Furthermore, his work provides insights into human behavioral deviations from neutral to risk-averse equilibrium, a phenomenon our study also examines in all-pay auctions. This cumulative body of research validates DRL's versatility and relevance, positioning it as an essential tool for addressing increasingly intricate auction dynamics.

Transfer learning (TL) in deep reinforcement learning (DRL) has gained attention as an approach to address some of RL's core challenges, such as sample inefficiency and the exploration-exploitation trade-off. By enabling agents to apply previously acquired knowledge to new, related tasks, TL accelerates learning and improves performance in complex environments where direct training is costly or impractical (Zhu et al., 2023). Traditional RL methods often rely on agents learning from scratch, a process that can be inefficient, particularly in high-dimensional tasks where tabula rasa learning can be prohibitive (Taylor and Stone, 2009). By leveraging knowledge from earlier tasks, TL enables generalization across tasks rather than just within a single task, a concept rooted in psychology and cognitive science (Lazaric, 2012). In DRL, various transfer methods, such as policy distillation and representation disentanglement, have demonstrated success in applying generalizable strategies to complex domains like robotics and autonomous systems, highlighting TL's potential to enhance RL performance across diverse applications.

3 BACKGROUND

Auctions are highly popular mechanisms for allocating goods and services to economic agents.[2] There is a wide variety of auction designs concerning participation rules, types of items being auctioned, bidding rules for participants, allocation of goods based on bids, and payment regulations.

Despite the wide variety of auctions, it is sufficient for the purposes of this work to focus on the simple case of auctions with private and independent values. In these auctions, each participant assigns a unique value to the item. An art auction serves as a good example. One participant might find the painting being auctioned beautiful and therefore highly valuable, while another might consider it unattractive and assign it a low value. Additionally, we will consider first-price sealed-bid auctions. In these auctions, participants submit their bids without knowing the bids of others. The participant who submits the highest bid wins the item being auctioned.

Auction theory is a branch of game theory, as each participant's payoff depends on their own action (bid) as well as the actions of others. Participants aim to maximize their payoff. Auction theory boasts a vast literature, where optimal strategies or Nash equilibrium for various types of auctions are studied. The strategies are described in terms of the ”bid function” $b(v)$, where the bid is a function of the value assigned to the auctioned item. In the following section, we

[2] This section was revised for grammar and wording with assistance from ChatGPT-3.

present some theoretical results for the auction designs chosen for the experiments in this work. The results can be found in (Klemperer, 1999), (Krishna, 2009) or (Menezes and Monteiro, 2008).

3.1 Algorithm Design

This research examines sealed-bid auctions involving a single item. Here, the auctioneer determines the winning bid from N participating agents. We conduct n auction rounds to observe the agents' behaviour and learning patterns, seeking convergence in their bids for each given value or signal over time. Each player i has a value v_i for the item. In private value auctions these values may differ among participants. The profit function for each agent is defined based on their bids: $\pi_i : \mathbb{R} \to \mathbb{R}$, where $\mathcal{B}$ is the vector space of possible bids of all agents. For example, in a sealed first-price auction of private values, a (risk-neutral) participant i's profit function is:

$$\pi_i(b_i, b_{\_i}) = \begin{cases} v_i - b_i & \text{if } b_i > \max(b_{\_i}) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $b_{\_i}$ represents the bids of other participants, excluding b_i.

3.2 The Rational Bid

Each participant i receives a value v_i, representing the value that she privatelly attributes for the item. Based on this value v_i, participant i formulates a bid $b_i(v_i)$. The expected payoff for participant i is given by:

$$E[u_i|v_i] = \int_B u(\pi(b_i(v_i), y)) f_{b_{\_i}}(y) dy$$

Here, $f_{b_{\_i}}(y)$ is the probability density function of the vector y, which contains the bids of other participants. Participants aim to maximize their expected reward, which requires knowledge of the function $f_{b_{\_i}}(y)$, dependent on other players' policies.

3.3 Types of Auctions

3.3.1 First Price Auction

The first-price auction is the most well-known auction design. The allocation rule is straightforward: the item is awarded to the participant with the highest bid (ignoring any tie-breaking rules for simplicity). The payment rule is also simple: the winning participant pays the amount they bid, while the non-winning bidders do not pay anything. Therefore, the payoff of a participant i is

$$\Pi_i = \begin{cases} v_i - b_i & \text{if } b_i > \max_{j \neq i}(b_j) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where v_i is player i's valuation, b_i is their bid, and b_j are the bids of other players. The Nash equilibrium of this auction when the private values come from uniform [0,1] independent distributions and the participants are risk-neutral is the same bid function for all (Krishna, 2009)

$$b_i^* = \frac{(N-1)v_i}{N}.$$

3.3.2 Second Price Auction

The allocation rule of the second-price auction (also known as a Vickrey auction, after the seminal work of Vickrey [1961]) is the same as that of the first-price auction: the winner is the participant with the highest bid. However, the payment rule differs: the winner pays the amount of the second-highest bid, not their own bid. The interest in this type of auction stems from the fact that the Nash equilibrium strategy for each participant is to bid their true valuation of the item, i.e., the bid function is the identity (Krishna, 2009):

3.3.3 All-Pay Auction

In an all-pay auction, the allocation rule is the same as previously described: the item is awarded to the participant with the highest bid. The unique aspect of this auction lies in its payment rule: all participants must pay their bids, regardless of whether they win. In this scenario, the Nash equilibrium for risk-neutral participants is determined by the bid function

$$b_i^* = \frac{(N-1)}{N} v_i^N.$$

4 METHODOLOGY

4.1 Agents Training and Evaluation

In this study, we employ the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) algorithm to train agents in auction environments, specifically focusing on optimizing bidding strategies.[3] Each agent is equipped with its own actor and critic networks, where the critic is trained using the observations and actions of all agents, capturing the interdependent nature of multi-agent environments like auctions. The

[3]This section was revised for grammar and wording with assistance from ChatGPT-3.

training process involves iterative learning, where agents receive private values at the beginning of each auction round and choose actions (bids) to maximize their expected utility based on rewards determined by the auction's payment rules. This setup allows agents to refine their strategies over time. The MADDPG architecture and workflow used in our approach is illustrated in Figure 1.

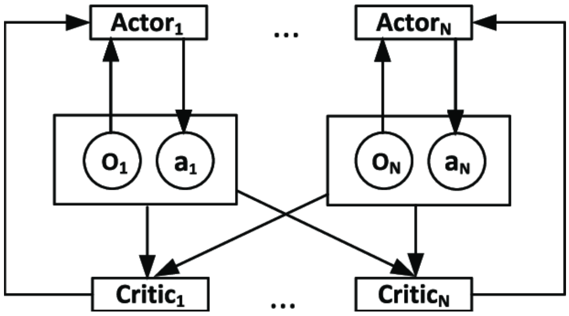


Figure 1: MADDPG Architecture - Figure taken from (Zheng and Liu, 2019).

To stabilize the learning process, we implement a Replay Buffer, which stores past experiences to break the correlation between consecutive interactions, allowing the agents to learn more effectively from a diverse set of experiences. We experiment with various buffer configurations, including a Combined Experience Replay Buffer (Zhang and Sutton, 2017), which merges historical experiences with the most recent interactions. This approach is particularly useful in dynamic environments, where the agent's policy is continually evolving. Additionally, we introduce dynamic noise in the agents' actions to balance exploration and exploitation. Early in training, higher noise encourages exploration of different bidding strategies, while later stages reduce the noise to focus on exploiting optimal strategies.

The neural network architecture consists of both actor and critic networks, each containing two layers with 100 neurons. The output layer uses a sigmoid activation function. During training, a batch size of 64 is used, with an actor learning rate set to 0.000025 and a critic learning rate of 0.00025. To aid the learning process, a decay factor is also applied. All hyperparameters, including the number of neurons per layer, were chosen based on preliminary experiments to ensure a balance between model performance and computational efficiency.

4.2 Transfer Learning in Auctions

In auction scenarios, particularly in multi-agent settings like all-pay auctions, finding a good starting point for training can significantly improve the efficiency and success of the learning process. In our previous research, we observed that as the number of agents N increased, agents often converged to a local equilibrium, where some of them would bid 0.0 for any private value, thus underperforming. This issue became especially prominent when $N \geq 3$, as random initialization did not provide sufficient guidance for the agents to explore more effective strategies. In this work, we propose the use of transfer learning to overcome this challenge by using trained models from lower-N auctions as a starting point for training agents in higher-N auctions. This allows us to provide agents with better initial conditions, reducing the risk of falling into suboptimal equilibrium.

Transfer learning in this context involves training agents in a smaller game space, with fewer players, where they can more easily learn stable strategies. Once these agents have been trained in a lower-N auction scenario, such as $N = 2$, we replicate their models for use in higher-N auctions. For instance, if we transition from $N = 2$ to $N = 5$, we can choose one or both of the initial agents, duplicate their parameters, and use them to populate the additional agents in the new setting. By starting with agents who have already converged to near-optimal policies, we can reduce the need for extensive retraining, and more importantly, avoid the instability that arises when starting from random initialization. This technique significantly accelerates convergence and leads to more efficient training.

The beauty of this approach lies in the flexibility of how these agents are duplicated and retrained. Since agents in lower-N scenarios often converge to very similar policies, there is little difference in which agents are duplicated for higher-N games. We can use any combination of the original agents' parameters to initialize the new ones. After duplication, the agents are then retrained to account for the new competitive environment with more participants. This method not only speeds up the learning process but also helps to address issues seen in previous work, such as local equilibrium that result from poorly initialized parameters.

A key technical challenge arises when increasing the number of agents: the input structure of the critic network must be adjusted. In multi-agent reinforcement learning, the critic takes into account the states and actions of all agents to evaluate each agent's decision-making process. When we scale from a lower-N to a higher-N auction, the input size of the critic increases accordingly, leading to dimension mismatches. Initially, this posed a programming error, as the critic's input was built for a smaller input space. This mismatch had to be addressed to ensure seamless scaling of the model.

To resolve this, we exploited the deep neural networks' capacity for handling redundancy in inputs. For the critic network, we duplicated the states and actions of the original agents as needed to match the higher N. By duplicating these inputs, we ensured the critic could still evaluate the joint actions of all agents, even if the inputs were redundant. Furthermore, to prevent future programming errors when scaling the model, we modified the critic's architecture during the training of the initial agents. We added extra input slots to accommodate additional agents, ensuring that when scaling up, the critic would already be prepared for the increased input size. This adaptation allowed us to maintain the critic's function without compromising the model's performance, providing a stable and scalable approach for transfer learning in auction settings.

However, this one-step transfer learning approach does not always yield optimal results or achieve near Nash equilibrium, especially as the number of agents increases. To address this, we introduced a more robust step-by-step heuristic for scaling up the number of agents while maintaining stability. Starting with trained agents from an $N = 2$ auction, we incrementally introduced one new agent at a time, progressively moving to $N = 3$, $N = 4$, and so on. At each step, the new agent's initial parameters were taken from one of the existing agents, ensuring a consistent starting point, while the entire ensemble was retrained to adapt to the new environment. This gradual increase allowed the agents to adjust more seamlessly to the added complexity, facilitating smoother convergence.

This iterative method proved particularly effective in preventing agents from defaulting to suboptimal behaviors, such as bidding 0.0 for all private values. The step-by-step integration helped the network manage increased strategic interactions without destabilizing the learning process. By methodically expanding the training environment, agents had the opportunity to adapt incrementally, resulting in more robust policy learning and a higher likelihood of achieving a global Nash equilibrium, as described in Section 5.

5 RESULTS

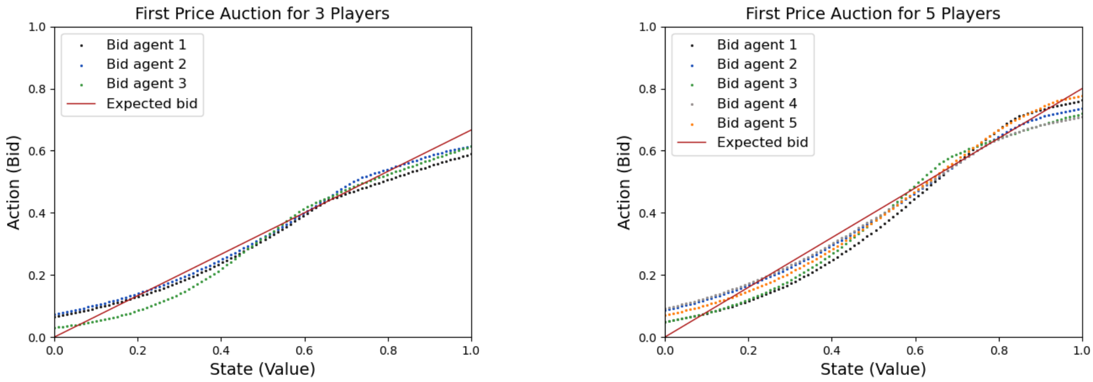
This section presents the experimental results, starting with basic auction types—first-price and second-price auctions—to validate the efficacy of the multi-agent deep reinforcement learning (DRL) approach in relatively straightforward settings.[4] In these standard auctions, transfer learning was not required for the DRL agents to reach equilibrium strategies. The agents naturally converged toward near-Nash equilibrium without getting stuck in local solutions, which often occurs in more complex auction types. This lack of dependence on transfer learning in basic auctions highlights the algorithm's ability to learn optimal bidding behavior when the strategic landscape presents fewer challenges.

Figures 2 (first-price auction with $N = 3$ and $N = 5$) and 3 (second-price auction with $N = 3$ and $N = 6$) illustrate the agents' steady convergence to expected equilibrium behaviors. These results align closely with theoretical predictions, confirming the robustness of the DRL model in simpler auction environments and demonstrating that agents can efficiently learn and adapt without needing additional techniques such as transfer learning. This strong foundational performance in standard auctions sets the stage for examining the more complex dynamics of all-pay auctions, where the benefits of transfer learning become essential for overcoming challenges such as local equilibrium.

In more complex settings, such as the all-pay auction, transfer learning proved valuable for enabling DRL agents to reach equilibrium despite the added strategic complexity. Instead of directly transitioning from a smaller-scale $N = 2$ auction to a larger $N = 5$ auction, we employed a step-by-step transfer learning approach. The model was first trained with two agents, which successfully converged to a near-Nash equilibrium. This learned strategy was then incrementally scaled by introducing one new agent at a time and retraining the ensemble at each step (e.g., $N = 2 \rightarrow N = 3 \rightarrow N = 4 \rightarrow N = 5$). This gradual approach allowed agents to adapt incrementally to the increasing strategic complexity, mitigating the risk of falling into suboptimal local equilibrium. The progression of this approach is illustrated in Figure 4, where the left subfigure (4a) shows the two-agent equilibrium and the other players bidding zero for any private value, while the right subfigure (4b) displays the successful equilibrium achieved in the four-agent setup. This example demonstrates how transfer learning allows learned strategies to scale effectively, enabling agents to adapt efficiently even in increasingly complex environments.

As the auction size increased further, the strategic complexity grew, and new challenges became evident. In the all-pay structure, each participant must pay their bid regardless of winning, which creates a strate-

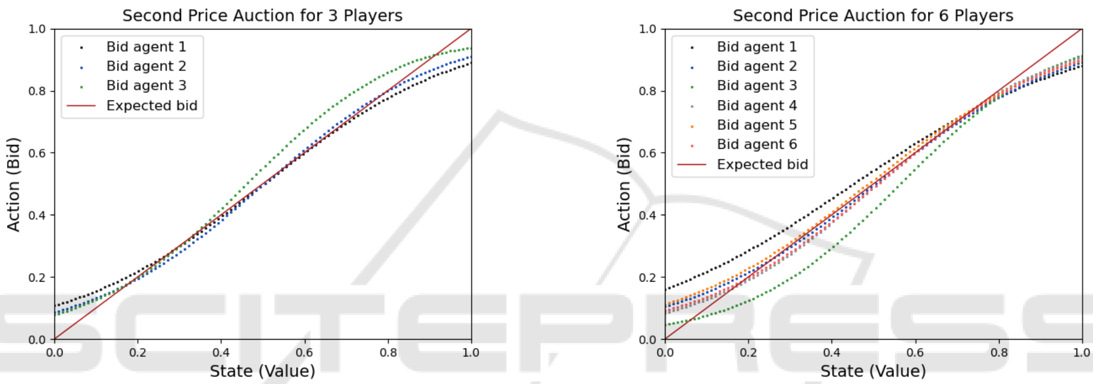
[4] This section was revised for grammar and wording with assistance from ChatGPT-3.



(a) First Price with 3 agents.

(b) First Price with 5 agents.

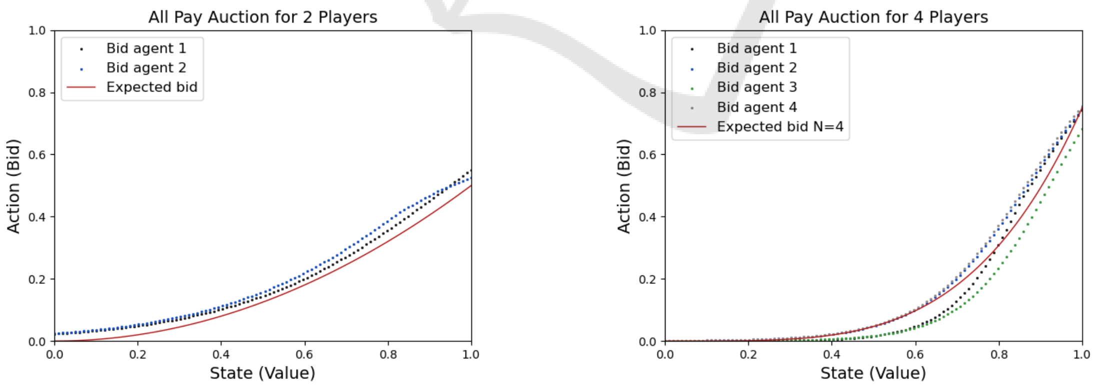
Figure 2: First Price Auction Results.



(a) Second Price with 3 agents.

(b) Second Price with 6 agents.

Figure 3: Second Price Auction Results.



(a) All Pay with 2 agents.

(b) All Pay with 4 agents starting from 2 agents.

Figure 4: Transfer Learning in All-Pay Auctions: convergence from $N = 2$ to $N = 4$ Agents.

gic landscape where suboptimal behaviors, such as zero-bidding, are more likely as the number of agents increases. This growth in complexity often led agents to fall into local minimum, deviating from Nash equilibrium strategies. Without additional support, such as transfer learning, agents in larger-scale settings struggled to maintain optimal policies, highlighting the limitations of reinforcement learning alone in handling the increasing strategic demands.

Transfer learning addressed these challenges by initializing agents with pre-trained policies from simpler, smaller-scale auctions. This approach allowed agents to begin from a more informed position rather than random initialization, helping them avoid common pitfalls. As shown in Figures 5, 6, and 7, agents that benefited from transfer learning not only avoided local equilibrium but also exhibited competitive strategies, with all agents converging to near-equilibrium behavior. The improvement in convergence was particularly pronounced in the $N = 5$ auction.

While due to the inherent stochasticity of these algorithms, convergence is not always guaranteed, transfer learning significantly improves convergence rates and helps agents avoid suboptimal bidding strategies. As the number of participants (N) increases, the likelihood of some agents getting stuck in local minimum also grows, reflecting the heightened strategic complexity in larger settings. Despite this, the transfer learning approach consistently outperforms random initialization, enabling more agents to converge to near-equilibrium strategies even in challenging scenarios.

In summary, transfer learning has proven to be a highly effective technique for addressing key challenges in multi-agent DRL for all-pay auctions. By initializing agents with strategies learned in smaller-scale auctions, we facilitated more efficient learning and achieved stable convergence across increasingly complex environments. The approach demonstrated strong results, particularly in settings with participant numbers ranging from $N = 2$ to $N = 6$, showing its robustness in navigating the strategic complexity of all-pay auctions. While some limitations remain, particularly in scenarios with larger participant numbers where convergence can still be challenging, the overall performance underscores the potential of transfer learning as a powerful strategy for enhancing learning efficiency and equilibrium convergence in competitive multi-agent systems.

6 DISCUSSION

The aim of this study was to introduce and evaluate a novel approach to training agents in multi-agent auction environments using transfer learning techniques.[5] Specifically, we focused on improving the convergence of deep reinforcement learning (DRL) agents in all-pay auctions, where previous research encountered challenges in finding equilibrium strategies as the number of participants (N) increased. Our transfer learning method proved highly effective in addressing these issues by using pre-trained agents from smaller auctions and using them as a starting point for training agents in more complex, higher-N auction scenarios.

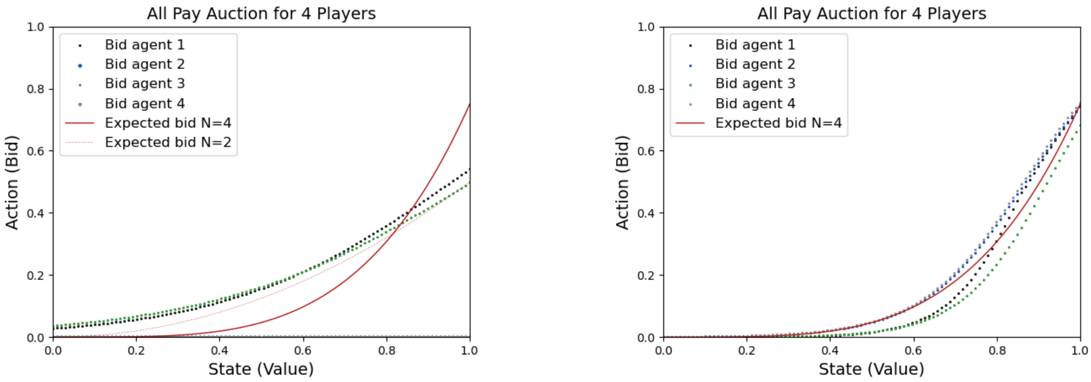
Our results demonstrate that this approach offers significant advantages over traditional random initialization methods, particularly in complex auction types like all-pay auctions, where convergence to a Nash equilibrium is notoriously difficult. By employing a stepwise transfer learning approach, where models were incrementally trained starting from lower-N auctions and moving to higher-N scenarios, agents began the training process with a strong initial condition. This method enabled them to avoid suboptimal bidding strategies, such as consistently bidding zero, which were prevalent in previous studies. The stepwise strategy effectively mitigated the occurrence of local equilibrium, where agents placed zero bids regardless of their private values, a problem that plagued our earlier research.

The novelty of this approach lies in its ability to solve the critical issue of local equilibrium by initializing agents with parameters that are more aligned with optimal strategies. By reusing agents trained in auctions with fewer participants, we provided them with a near-optimal bidding strategy that could be adapted to higher-N environments. This not only improved the efficiency of the training process, but also increased the agents' ability to learn robust bidding strategies in larger and more complex auctions. Additionally, we adapted the critic network to account for the changing number of participants in higher-N auctions. This modification involved designing the critic network to accommodate additional agent inputs by duplicating existing inputs, allowing the network to effectively process the higher-dimensional input space without requiring a complete reinitialization of network parameters.

However, as N becomes larger, the algorithm faces increasing challenges in maintaining performance and achieving convergence. The added complexity of interactions among a higher number of agents creates a larger strategy space, making it more difficult for agents to reach an optimal equilibrium. In particular, the probability of local minimum increases as agents struggle to adapt to the expanding competitive environment. Although transfer learning significantly improves scalability, the diminishing returns observed for very large N highlight the need for further refinements, such as adaptive learning mechanisms or more sophisticated initialization strategies.

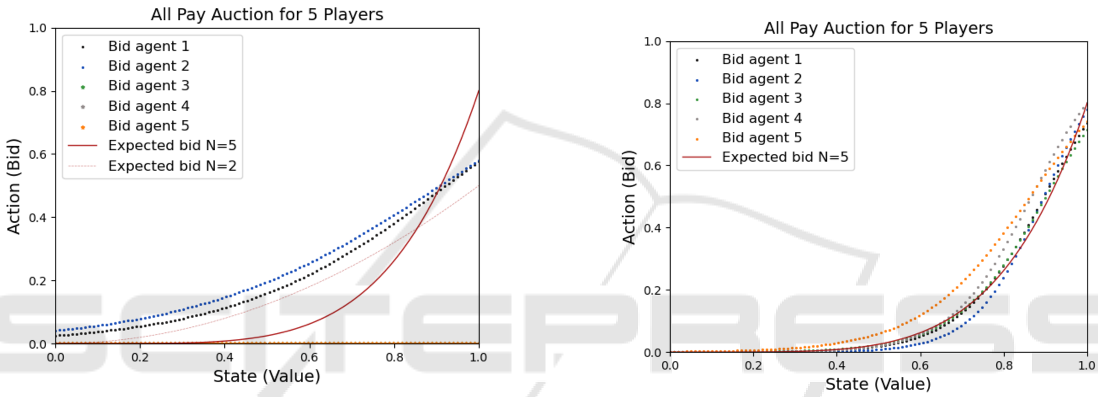
In summary, the transfer learning approach we implemented in this study represents a substantial im-

[5]This section was revised for grammar and wording with assistance from ChatGPT-3.



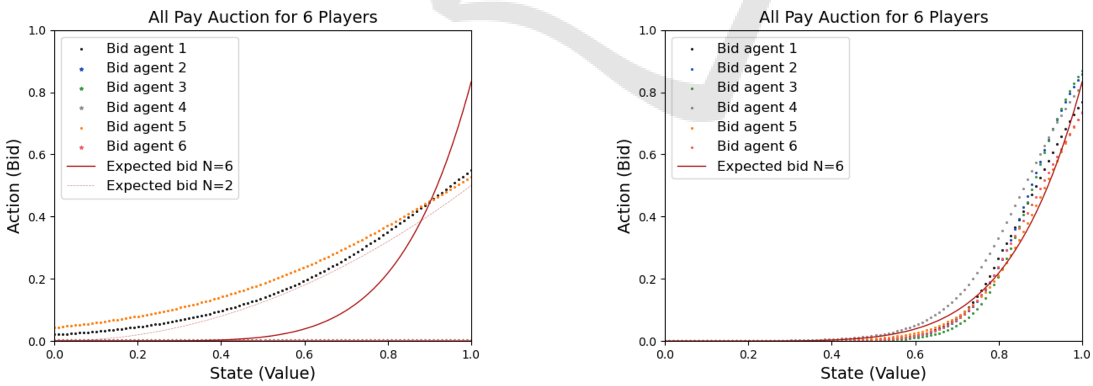
(a) All Pay with 4 agents - Regular Learning. (b) All Pay Auction with 4 agents - Transfer Learning.

Figure 5: All Pay with 4 agents.



(a) All Pay with 5 agents - Regular Learning. (b) All Pay with 5 agents - Transfer Learning.

Figure 6: All Pay Auction with 5 agents.



(a) All Pay with 6 agents - Regular Learning. (b) All Pay with 6 agents - Transfer Learning.

Figure 7: All Pay Auction with 6 agents.

provement in training DRL agents in all-pay auctions and similar multi-agent settings. By leveraging pre-trained models, we not only solved critical issues from previous research but also laid the groundwork for further applications of transfer learning in other complex auction formats. This method holds promise for improving learning efficiency in a variety of auction types and multi-agent systems, ultimately broadening the applicability of DRL in strategic decision-making environments.

7 CONCLUSIONS

This study introduced a novel transfer learning approach for training agents in multi-agent auction environments, specifically focusing on all-pay auctions.[6] The results demonstrated strong performance in enabling agents to converge toward Nash equilibrium strategies by leveraging pre-trained models from smaller auctions. This method effectively mitigated challenges associated with local equilibrium and significantly enhanced the efficiency of the learning process.

Our findings indicate that transfer learning is particularly effective even when there is a substantial difference in the number of agents between the pre-trained and new models, especially when using a step-by-step transfer approach. By incrementally introducing one agent at a time, we observed enhanced performance and scalability, allowing for better adaptation to larger agent populations. Again, as N increases, the growing strategy space and heightened risk of convergence to local minimum pose challenges, emphasizing the need for enhanced techniques to ensure efficiency in high-N environments.

Future work will explore scaling the algorithm to handle auctions with significantly larger N, as well as extending its application to auctions with interdependent values. In interdependent value settings, the valuation of the item depends not only on private signals but also on shared external factors, creating additional complexity in learning optimal strategies. Investigating how transfer learning performs in these environments will provide valuable insights into its adaptability and robustness. Additionally, comparative experiments with other transfer learning methods and alternative DRL architectures are planned to evaluate the effectiveness of the proposed approach against state-of-the-art techniques. Furthermore, we aim to refine the proposed method by incorporating adaptive learning rates, exploring curriculum learning, and testing it in broader multi-agent environments. These enhancements will help generalize the approach to a wider range of auction formats, ultimately contributing to more effective strategic decision-making in competitive and cooperative systems.

The incremental approach used in this study aimed to mitigate the emergence of local equilibrium by starting from a simpler problem and gradually transforming it into the target problem. This technique is inspired by methods like numerical continuation (Allgower and Georg, 2012), where a problem is solved incrementally by starting with a simpler, well-understood version and progressively increasing its complexity. In our case, agents trained in lower-N auctions adapted their strategies step by step as new agents were introduced, avoiding the abrupt strategy shifts often associated with random initialization in higher-N settings. While this approach proved effective for the scenarios tested, we recognize that the efficiency and success of this method may depend on the specific auction format and the way the incremental transition is implemented.

Moreover, we envision applying this technique to broader DRL applications, particularly in scenarios where agents often achieve suboptimal strategies and lack incentives to leave such states, exemplified by local equilibrium. In general, the promising results of our experiments suggest that transfer learning can play a crucial role in enhancing the training of agents in complex auction scenarios. By building on the foundation established in this study, we aim to further investigate the application of this approach across a broader range of auction types and multi-agent environments, ultimately contributing to more effective strategic decision-making in competitive settings.

[6]This section was revised for grammar and wording with assistance from ChatGPT-3.

REFERENCES

Allgower, E. L. and Georg, K. (2012). *Numerical continuation methods: an introduction*, volume 13. Springer Science & Business Media.

Bichler, M., Fichtl, M., Heidekrüger, S., Kohring, N., and Sutterer, P. (2021). Learning equilibria in symmetric auction games using artificial neural networks. *Nature machine intelligence*, 3(8):687–695.

Craizer, L. E., Hermann, E., and Alvim, M. (2025). Learning optimal bidding strategies in all-pay auctions. In *Proceedings of the International Conference on Artificial Intelligence and Soft Computing (ICAISC)*, volume 15165 of *Lecture Notes in Computer Science*. Springer. To appear.

Dütting, P., Feng, Z., Narasimhan, H., Parkes, D. C., and Ravindranath, S. S. (2021). Optimal auctions through deep learning. *Communications of the ACM*, 64(8):109–116.

Ewert, M., Heidekrüger, S., and Bichler, M. (2022). Approaching the overbidding puzzle in all-pay auctions: Explaining human behavior through bayesian optimization and equilibrium learning. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, pages 1586–1588.

Gemp, I., Anthony, T., Kramar, J., Eccles, T., Tacchetti, A., and Bachrach, Y. (2022). Designing all-pay auctions using deep learning and multi-agent simulation. *Scientific Reports*, 12(1):16937.

Kannan, K. N., Pamuru, V., and Rosokha, Y. (2019). Using machine learning for modeling human behavior and analyzing friction in generalized second price auctions. *Available at SSRN 3315772*.

Klemperer, P. (1999). Auction theory: A guide to the literature. *Journal of economic surveys*, 13(3):227–286.

Krishna, V. (2009). *Auction theory*. Academic press.

Lazaric, A. (2012). Transfer in reinforcement learning: a framework and a survey. In *Reinforcement Learning: State-of-the-Art*, pages 143–173. Springer.

Luong, N. C., Xiong, Z., Wang, P., and Niyato, D. (2018). Optimal auction for edge computing resource management in mobile blockchain networks: A deep learning approach. In *2018 IEEE international conference on communications (ICC)*, pages 1–6. IEEE.

Menezes, F. and Monteiro, P. (2008). An introduction to auction theory: Oxford university press.

Mnih, V. (2016). Asynchronous methods for deep reinforcement learning. *arXiv preprint arXiv:1602.01783*.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. (2015). Human-level control through deep reinforcement learning. *nature*, 518(7540):529–533.

Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., Guez, A., Lockhart, E., Hassabis, D., Graepel, T., et al. (2020). Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 588(7839):604–609.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Sutton, R. S. (2018). Reinforcement learning: An introduction. *A Bradford Book*.

Taylor, M. E. and Stone, P. (2009). Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, 10(7).

Zhang, S. and Sutton, R. S. (2017). A deeper look at experience replay. *arXiv preprint arXiv:1712.01275*.

Zheng, S. and Liu, H. (2019). Improved multi-agent deep deterministic policy gradient for path planning-based crowd simulation. *Ieee Access*, 7:147755–147770.

Zhu, Z., Lin, K., Jain, A. K., and Zhou, J. (2023). Transfer learning in deep reinforcement learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11):13344–13362.