CIVICA RESEARCH

Co-funded by the Horizon 2020 Programme of the European Union

Call for Research Proposals 2022

APPLICATION FORM

CIVICA Partners are invited to submit proposals under the 3rd Call for Proposals for research projects jointly executed by CIVICA partner universities.

The overarching aim of the initiative is to encourage joint, cross-disciplinary research in the priority areas of CIVICA.

Basic information

Which priority area(s) does your project pertain to?

1. ☐ Societies in Transition, Crises of Earth
2. ☐ Democracy in the 21st Century
3. ☐ Europe Revisited
4. x **Data Driven Technologies for Social Sciences**
5. ☐ None of the above (specify which area your project is in)

Project title (mandatory):	Multi-agent learning and equilibrium
Acronym (mandatory):	EquiLearn

Project team

Please list all the participants in the project and their affiliation.

Project team leader – PI (must be a permanent faculty with one of the CIVICA Partners – responsible for submitting the application and reporting)

Name	CIVICA institution	Unit	Position	Email	Gender[1]
Bernhard von Stengel	LSE	Mathematics (Game Theory)	Professor	b.von-stengel@lse.ac.uk	Man

Team members

Name	CIVICA institution	Unit	Position	Email	Gender[1]
Galit Ashkenazi-Golan	LSE	Mathematics (Game Theory)	Assistant Professor	G.Ashkenazi-Golan@lse.ac.uk	Woman
Katerina Papadaki	LSE	Mathematics (Operations Research)	Associate Professor	K.P.Papadaki@lse.ac.uk	Woman
Andrea Celli	Bocconi	Computing Sciences	Assistant Professor	andrea.celli2@unibocconi.it	Man
Mark Voorneveld	Stockholm School of Economics	Economics	Associate Professor	Mark.Voorneveld@hhs.se	Man

[1] This data is collected purely for statistical purposes. Please choose of the following options: “Man”, “Woman”, “Non-binary”, “Prefer not to say”.

Designated contact person for administration and budget (Each participating institution needs to name one person responsible for administering the institution's part of the budget)

Name	CIVICA institution	Unit	Position	Email
Aygen Kurt-Dickson	LSE	Research and Innovation (R&I)	R&I Strategy Manager	A.S.Kurt-Dickson@lse.ac.uk
Chiara Monetti	Bocconi	Artificial Intelligence Lab	Grants Officer	c.monetti@unibocconi.it
Göran Lindqvist	SSE	Research Office	Director of Research	Goran.Lindqvist@hhs.se

Detailed information

Expected duration

Please specify your proposed start and end dates. Please note that the spending deadline is 1 November 2023.

Start date:	17 October 2022	End date:	1 November 2023

Brief description

Please briefly describe your project objective(s), planned activities, and expected results (max 300 words)

This project studies machine learning by multiple agents in an interactive environment. It starts with a specific model of competition between firms which price a good over a number of time periods. The learned pricing strategies then evolve as a population over time. The population is represented with fractions of these strategies in a "mixed equilibrium". The research may explain how learning algorithms lead to *price collusion* rather than full competition.

The novel approach is that the system is modeled in two stages. The first stage models reinforcement learning in the dynamic pricing game, in order to find a good pricing strategy. The second stage models a population of strategies which are in *equilibrium* and define the learning environment for the first stage. A newly learned strategy, if successful, is added to the population, for which a new equilibrium is computed.

This novel two-stage approach has several advantages. First, the underlying pricing game is already more complex and therefore realistic because strategies are *learned* rather than explicitly listed in advance. Second, it is *modular* rather than a large monolithic simulation. The general software framework (a deliverable) can and will be adapted to other models. Explicitly, we will test the methodology with a different patrolling game instead of the pricing game for the first learning stage, and varying models of equilibrium for the second population stage.

The applicants are complemented by their expertise in equilibrium computation (von Stengel), dynamic games (Ashkenazi-Golan), patrolling games and applied operations research (Papadaki), machine learning (Celli), and general solution concepts in economics and games (Voorneveld).

The planned activities are travels for in-person joint research and a final workshop. Reproducible data and open-source software for the computational experiments will be made public. Assuming success, this should be seed-funding for a (currently explored) larger grant.

CIVICA RESEARCH

Co-funded by the Horizon 2020 Programme of the European Union

Extended description

Please describe the project in more detail, including: the state of the art in the respective area(s) of work or research; contributions and roles of the faculty members in the team; the methodology; the added value of your project for developing CIVICA as the European social science university, and the anticipated outputs, deliverables, impact, and beneficiaries (target audience/end users) (max. 3000-4000 words).

Purpose of research and state of the art

Learning algorithms make millions of economic decisions automatically, such as *pricing* airline seats or hotel stays, or bidding for advertisement slots next to online search results. These algorithms are designed to optimize revenue given expected demand and own supply capacities, or to use an advertisement budget most effectively. Without being explicitly programmed to do so, they may also lead to *collusion* among competitors, with higher prices for consumers or lower revenue for the advertisement auctioneer, which raises challenging questions for welfare considerations and even competition law [1]. In order to understand why this happens, one needs to study learning algorithms in interactive settings, and how their actions influence each other.

This project is a fundamental study of automated learning in interactive environments with the tools of algorithmic game theory. It addresses two current shortcomings of game theory and multi-agent learning in this context: First, complex games in realistic settings have too many strategies to write down explicitly, which makes the analysis of such games very difficult; here, these strategies will be generated by learning algorithms. Second, learning in interactive scenarios should be trained against opponent strategies that already interact in a reasonable way, not arbitrary ones that lead to slow simulations, too much noise, and lack of convergence; here, these opponent strategies will be computed from the current set of strategies with an automated *equilibrium analysis*, which is the main novelty of this research. **Equilibrium computation** is the particular strength of the applicants across the CIVICA institutions.

Games as studied in game theory are mathematical models of conflict and cooperation. Algorithms have been most successful for playing complex games with well-defined rules such as Chess, and recently in the board game Go with deep learning algorithms trained on billions of computer-generated game plays. Even in Poker, with its lack of information because players do not know their opponents' hands, and necessary deception with randomized bluffing, algorithms now compete with the best human players (aided by an algorithm by the PI von Stengel [2] that allows to compute perfect play).

As realistic models of economic interaction, however, games are necessarily simplified and still mostly analyzed "by hand". The main solution concept for games is *equilibrium*, which recommends an action to each agent that is optimal under the assumption that the other agents follow their recommendations. Unless the game is zero-sum (as in Chess or Poker), there is in general no "optimal play" because equilibrium is an interdependent concept that is not unique – agents may be stuck in a "bad" equilibrium.

In this project, a twofold approach is taken to analyze more realistic games as economic models, and to study the role of machine learning in this context. As a starting point we consider a classic model of competition between firms, an oligopoly with "demand inertia", which is played over a number of rounds where firms set prices for a product and lose (or gain) customers in the next round in proportion to how much more expensive (or cheaper) they were compared to their competitors. The classic analysis of this model by Selten in 1965 prescribes very competitive behaviour. However, in later experiments where game theorists submitted strategies that played against each other in a round-robin tournament [3], much more cooperative behaviour was observed, with collusion and higher prices, and also strategies that exploited naive pricing strategies. At the time, von Stengel participated very successfully in this strategic tournament.

The first new research question is **if a successful pricing strategy in this oligopoly game can be found by machine learning**. Unlike the large number of states (due to discretization) in the Q-learning approach in [1], we hope that a much smaller neural net can exploit the inherent continuity of setting prices. This should lead to much faster learning.

CIVICA RESEARCH
Co-funded by the Horizon 2020 Programme of the European Union

The second, important conceptual new research question is **how to *evaluate* the performance of the pricing strategy that is to be learned**. This should not be done against some arbitrary set of possible competitor strategies (e.g., as submitted in the experiments in [3]), but against strategies that are successful and eventually prevail. There are good reasons to assume that this will be a *mixture* of strategies as in a mixed equilibrium (whose existence was proved by Nash). The envisaged approach is to start with some set of simple initial strategies, against which the learning algorithm finds a successful strategy, which is then entered to the existing pool of strategies. For the resulting game, an *equilibrium* is found, against which the next iteration of the learning algorithm is then trained. One then needs to see if this process converges, and how diverse the resulting set of strategies will be.

In case the game has multiple equilibria, learning can be applied to each of them, which may lead to different "evolutions" of successful strategies. This is an important conceptual problem that will also be investigated with varying, random starting points for equilibrium computation as used in [2], or evolutionary dynamics. This addresses the known problem of *equilibrium selection* in games.

Contributions and roles of the faculty members in the team

The PI Bernhard von Stengel (at LSE since 1998) is a mathematical game theorist and an expert on equilibrium computation. He found an algorithm for solving very large extensive games (dynamic games with imperfect information) that has been used to compute perfect play in two-person Poker, and allows for equilibrium selection [2]. In [4], he extended Aumann's concept of *correlated equilibrium* to extensive games as "extensive-form correlated equilibrium". Unlike Nash equilibrium, correlated equilibrium allows for coordinated actions among agents, and is the natural outcome of regret-based learning algorithms.

Galit Ashkenazi-Golan (at LSE since 2021) is an expert on dynamic games and probability theory [5]. The present project is meant to be a seed project **for a larger UKRI Future Leadership grant** that she will apply for, which has as its subject the study of machine learning in general dynamic games.

Katerina Papadaki (at LSE since 2002) works in applied operations research and is a specialist in search and patrolling games [6]. Her current interests are in applying artificial intelligence to model and analyze more realistic games of this kind. All three LSE researchers and their PhD students are currently actively collaborating on games and machine learning.

Andrea Celli (at Bocconi since 2021) is a computer scientist working on game theory. He extended von Stengel's concept of extensive-form correlated equilibrium to learning algorithms in [7], which won the Best Paper Award at NeurIPS 2020, the most prestigious machine learning conference.

Mark Voorneveld (at SSE since 2001) is a mathematical economist. He is an expert on equilibrium concepts, and in [8] found a novel simple proof for the existence of fixed points that promises a much simpler computation of an equilibrium in the challenging scenario of games with many players.

The team does not yet have any joint publications. This project has the promise of many synergies due to the joint expertise and interests of the team, and fortuitous timing. Ashkenazi-Golan's arrival at LSE coincided with an unprecedented start of four PhD students in game theory in the Department of Mathematics at LSE, who will contribute with enthusiasm and interest to programming and research. Papadaki has recently expanded her research into artificial intelligence. Von Stengel will be on sabbatical leave for the entire academic year 2022/23 with full time for research on this project. He has supervised several MSc theses on the work [8] of Voorneveld, and corresponded with him on possible novel directions. Voorneveld also co-organized, on behalf of the Swedish Nobel Foundation, the symposium "100 years of game theory". It took place in December 2021 in Stockholm and was attended by a large number of Nobel laureates; PI von Stengel chaired the session on mathematical game theory. Celli is an expert on machine learning (and game theory), and will help the rest of the team with getting quickly up-to-date on developments in this extremely active area.

The research will be conducted by in-person meetings, with most of the applied funds intended for travel, as we believe spending time together will help us develop the theory and apply the model more effectively. We will adhere to LSE's net carbon zero emission and sustainability policies by planning our travels carefully.

CIVICA RESEARCH
Co-funded by the
Horizon 2020 Programme
of the European Union

This will help us adhere to keep LSE's status as the first carbon-neutral certified UK university [10].

The team will explore larger-scale funding, such as an **ERC Synergy Grant**, to continue this cooperation. Its initial seed funding via CIVICA should provide an excellent opportunity to produce early results from prototypes of the model, for which the scientific approach is very suitable.

Methodology and more detailed research questions

As outlined above, the approach has two main parts: To start with, a "pricing game" defines an interaction between two players in a duopoly, or between several players in an oligopoly. Due to the "demand inertia" in this model [3], this is a dynamic game over many rounds that influence each other (which is not the case in the statically defined interaction in [1]). Apart from some simple "manually" designed strategies to start with (possibly taken from [3]), machine learning models are applied that learn how to play well in this game. They can be Q-learning, neural nets, or regret minimization, and these various approaches will be compared with each other. We conjecture that a neural net that uses only the last few pricing periods as inputs, with its potential to compute the next price as a continuous function, leads to much faster learning than the hundreds of thousands of iterations needed for the discretized Q-learning model in [1].

The second main part is to define the environment in which this learning takes place. If the agents learn simultaneously and treat the other agents as part of the environment, the resulting overall behaviour often does not converge, and individual strategies tend to be over-fitted against each other [9]. Instead, the novel approach taken here is to learn a new strategy against an existing *equilibrium* of other strategies, and only enter the new strategy to the pool of existing strategies once (and if) it performs successfully. This is akin to an evolutionary entrant that starts to invade a population once it is successful. In fact, a "realistic" model of competition would be to model this as a dynamical system and study its evolution. We plan a comparison in this regard (see below) with our main approach, namely an equilibrium analysis.

To repeat, our model has two stages: The first stage is to apply a learning algorithm in a dynamic pricing game in order to find a strategy that sets prices over time periods. The *performance* of this strategy, which is needed to guide the reinforcement learning, is measured against a population of existing other strategies that are currently in an (approximate) equilibrium, typically as a population mix (a mixed Nash equilibrium). This *population game* defines the second stage of our model. This process is iterated by adding a successfully learned strategy in the first stage to the population game. In that extended game, a new *equilibrium is computed*, which defines the next learning environment. We will study the iteration for its convergence; for example, it may terminate if no better entrant is found at the first stage. For the equilibrium computation at the second stage we will compare appropriate notions of equilibrium, such as Nash equilibrium as in [2], correlated equilibrium as in [4] or [7], or equilibrium as defined by fixed points as in [8], where the latter should be of particular interest when investigating an oligopoly rather than a duopoly at the first stage.

Equilibrium computation is challenging from a theoretical perspective. Already two-player games that are not zero-sum are known to be "PPAD-hard", which means that an equilibrium of a large game may take prohibitively long to be found. In addition, for a game with more than two players the equilibrium property will only hold approximately, not exactly. However, these objections are mitigated by our approach because the population of competing firms is not represented by a large number of agents, but by their numerical fractions of *strategies* (the strategies that are learned and added to the population game). Early investigations show that only a small number of strategies have a significant fraction (technically, the mixed equilibria of the game have "small support"). A complete equilibrium analysis of the population game at the second stage seems therefore very feasible. Similarly, the learning environment for the first stage defined by any such equilibrium is easily evaluated, so that this learning stage is not too costly (in contrast to large multi-agent simulations).

Another interesting aspect for study is the non-uniqueness of an equilibrium. The pricing game has elements of competition and cooperation, and therefore has multiple equilibria. This can be observed already with rather simple strategies (not yet found by machine learning): One equilibrium may be given by a single very aggressive strategy for each player, another by a mixture of several colluding strategies that support each other together with an aggressor that is not given much weight in equilibrium. Equilibrium-finding algorithms can be designed to follow a *path* from a *starting point* that represents a known equilibrium of a simpler game,

CIVICA RESEARCH
Co-funded by the Horizon 2020 Programme of the European Union

which is then gradually adjusted to incorporate new strategies until it represents an equilibrium of the full game. In our two-stage approach, a natural starting point is the old equilibrium against which the new entrant has been trained. However, it is also possible to test many randomly chosen starting points in order to see how the resulting equilibrium varies, or is robust with relation to this starting point. Varying starting points have been used in [2] for two-player games. For more general games, a concrete research question is if the path-following algorithm of Voorneveld in [8] for finding an approximate fixed point can be adapted to an arbitrary starting point. We also plan to compare the starting points of evolutionary dynamics, such as the replicator dynamics, with these path-following equilibrium computations.

The particular attraction of the two-stage methodology is that it is **modular** rather than a big monolithic simulation. Various learning methods can be tested and compared at the first stage. The effect of new entrants can be compared with various equilibrium notions at the second stage. A great number of technical options are available here, on which the team has strong expertise. The modular approach also promises ***transferable* insights**: The underlying pricing game at the first stage, for which learning algorithms are devised, can be replaced by other games. As a specific example, we test the methodology by applying them to zero-sum (that is, strictly competitive) games such as patrolling games, the expertise of Papadaki. Another example is a competition for advertisement space as in online ad auctions; Celli worked on such models as a postdoctoral researcher at Facebook Core Data Science in 2020/21. Similarly, the computation of an equilibrium in the second stage (which then defines next the learning environment at the first stage) is of general purpose.

Anticipated outputs, deliverables, impact, and beneficiaries (target audience/end users)

The general academic interest in multi-agent reinforcement learning is extremely large and timely; the papers [1] and [9] have been cited hundreds of times per year. Deliverables will be academic papers, first aimed at conferences such as NeurIPS and Economics and Computation (EC). In addition, the designed computer programs will be made public as open-source software on an appropriate platform such as Github. The aim is to document the computational experiments thoroughly and *reproducibly*.

Intended beneficiaries are also people who are not software experts, where the programs can be run with a friendly user interface on the web platform http://www.gametheoryexplorer.org, long maintained by von Stengel and collaborators. These users could be economists or, say, business students who want to experiment with a pricing model to see if collusive behaviour emerges.

Added value of the project for developing CIVICA as the European social science university

The members of the project team of this project are diverse in their career stages and their disciplines: mathematics, computer science, and economics. This fits the cross-disciplinary character of the study: Understanding the use of machine learning in multi-agent environments, with an application - the pricing game - that highlights how collusion may arise from machine learning in an environment that should be competitive. **Machine learning** has been extremely successful in single-agent scenarios (such as image recognition), but a fundamental study **in the context of multiple agents** is highly important for social science. Significantly, this study is *quantitative* and draws on the strong expertise and authority of the applicants in mathematical models of interaction, in particular game theory. This quantitative emphasis should be an excellent complement to the standard social-science activities of CIVICA.

Selected references

[1] Calvano E. et al (2020). Artificial intelligence, algorithmic pricing, and collusion. American Economic Review 110(10), 3267–3297.

[2] von Stengel B. et al (2002). Computing normal form perfect equilibria for extensive two-person games. Econometrica 70(2), 693–715.

[3] Keser C. (1993). Some results of experimental duopoly markets with demand inertia. The Journal of

CIVICA RESEARCH
Co-funded by the Horizon 2020 Programme of the European Union

Industrial Economics 41(2), 133–151.

[4] von Stengel B., Forges F. (2008). Extensive-form correlated equilibrium: Definition and computational complexity. Mathematics of Operations Research 33(4), 1002–1022.

[5] Ashkenazi-Golan G., Lehrer E. (2019). Blackwell's comparison of experiments and discounted repeated games. Games and Economic Behavior 117, 163–194.

[6] Papadaki K. P. et al (2016). Patrolling a border. Operations Research 64(6), 1256–1269.

[7] Celli A. et al (2020). No-regret learning dynamics for extensive-form correlated equilibrium. Advances in Neural Information Processing Systems 33, 7722–7732. (Best paper award)

[8] Petri H., Voorneveld M. (2018). No bullying! A playful proof of Brouwer's fixed-point theorem. Journal of Mathematical Economics 78, 1–5.

[9] Lanctot M. et al (2017). A unified game-theoretic approach to multiagent reinforcement learning. Advances in Neural Information Processing Systems 30.

[10] LSE (2021) LSE becomes the first Carbon Neutral verified university in the UK. Online. Available at: https://www.lse.ac.uk/News/Latest-news-from-LSE/2021/k-November-21/LSE-becomes-the-first-Carbon-Neutral-verified-university-in-the-UK [last accessed on 14 September 2022].

CIVICA RESEARCH

Co-funded by the Horizon 2020 Programme of the European Union