AI-driven Prices for Externalities and Sustainability in Production Markets

Panayiotis Danassis
Harvard University
Cambridge, MA, USA
pdanassis@g.harvard.edu

Aris Filos-Ratsikas
University of Edinburgh
Edinburgh, United Kingdom
aris.filos-ratsikas@ed.ac.uk

Haipeng Chen
William & Mary
Williamsburg, VA, USA
hchen23@wm.edu

Milind Tambe
Harvard University
Cambridge, MA, USA
milind_tambe@harvard.edu

Boi Faltings
EPFL
Lausanne, Switzerland
boi.faltings@epfl.ch

ABSTRACT

Markets do not account for negative externalities; indirect costs that some participants impose on others, such as the cost of over-appropriating a common-pool resource (which diminishes future stock, and thus harvest, for everyone). Quantifying appropriate interventions to market prices has proven to be quite challenging. We propose a practical approach to computing market prices and allocations via a deep reinforcement learning policymaker agent, operating in an environment of other learning agents. Our policy-maker allows us to *tune* the prices with regard to diverse objectives such as sustainability and resource wastefulness, fairness, buyers' and sellers' welfare, etc. As a highlight of our findings, our policymaker is significantly more successful in maintaining resource *sustainability*, compared to the market equilibrium outcome, in scarce resource environments.

KEYWORDS

Sustainability, markets, market failure, multi-agent learning

ACM Reference Format:
Panayiotis Danassis, Aris Filos-Ratsikas, Haipeng Chen, Milind Tambe, and Boi Faltings. 2023. AI-driven Prices for Externalities and Sustainability in Production Markets. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Competitive markets, founded in the works of Walras (1874) and Fisher (1892), constitute the fundamental mechanism of allocation; the means that products are sold and bought. Market theory [1] suggests that free markets will reach an efficient stable outcome, the *market equilibrium*, in which supply equals demand, and all participants are maximally satisfied by the bundles of goods that they buy or sell at the chosen prices.

Nevertheless, free markets fail to account for negative *externalities* [19], which lead to market failure [2]. These externalities refer to indirect costs that are not reflected in the market equilibrium prices. A representative example of such inefficiencies is the environmental harm caused by pollution and overexploitation of natural resources, e.g., air pollution from burning fossil fuels, water pollution from industrial effluents, antibiotic resistance due to overuse of antibiotics in industrial farms, etc. Another prominent example of a negative externality – which we will use as a real-world, indicative test-case throughout the paper – is the depletion of the stock of fish due to overfishing (e.g., according to OECD, about 25% of fish stocks globally are at risk [21]). With these "exogenous" objectives being of paramount importance, it is only natural to assume some form of intervention to the reign of free markets.

There are many approaches to reconcile the economics of the free market with societal and environmental externalities. For example, policy-makers can correct for the inefficiencies by employing command-and-control legislation (e.g., [27]), permit markets [8, 9] (e.g., [13]), or taxation (e.g., [14]). A classic example of the latter is Pigouvian taxes [22], i.e., taxes that are equal to the external damage caused by the production decisions. While such interventions are clearly necessary, *selecting and quantifying* the appropriate ones has proven to be quite challenging. For instance, in the case of common-fisheries, approaches aiming to determine the "optimal" level of annual harvest and subsequently control fishing to achieve that quota have often failed to prevent overfishing [6]. Similarly, determining the marginal social cost of a negative externality and converting it into a monetary value can be quite impractical [3].

An added complication when it comes to devising effective policies for sustainability and combating externalities, or any societal objective for that matter, is the fact that the interactions between the different entities in the market ecosystem are rather complex and of a repeated nature. Indeed, the appropriate mathematical modeling of these systems is that of a *Markov* (or *Stochastic*) Game, in which the actors (i.e., the policy-maker and the harvesters of natural resources) are both aiming to optimize their individual utilities over a fixed horizon. To do that, they need to optimize over their future rewards, taking into the account the effect of the actions of the other actors on their own. Solving these games analytically is both conceptually and computationally hard, even for relative simple variants of those games and well-behaved equilibrium notions (e.g., see [5, 11, 12]). For this reason, most classic works in economics and mathematics (e.g., [16, 24, 25]) have only gone as far as identifying conditions that merely establish the existence of some equilibrium, without providing any guarantees about its properties.

For the full version of the paper see [10].

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

Besides the computational burden, another significant hurdle to the analytical approach is that it typically requires full observability of the environment and the actions of the other participants, which is most often not the case in practice.

Reinforcement learning (see [18]) has been proposed and extensively used as an alternative approach for computing optimal strategies in Markov Games [20]. The idea is that the actors, as learners, interact with their environment exclusively via signals of limited information: they typically observe their rewards based on their past actions, and update their current actions accordingly, via the employment of some carefully devised learning algorithm. Note that this approach does not require observable information about the parameters of the environment or the other actors. It also does not require the derivation of analytical solutions to complex optimization problems, as "off the shelf" reinforcement learning algorithms are readily available. For these reasons, an established line of work has considered reinforcement learning as a form of bounded rationality [23] which is much more conceivable for complex environments in practice, compared to the standard "perfect" rationality of traditional economic agents. Finally, reinforcement learning has been shown to be generally robust to changes in a range of input parameters, making it very suitable for complex and volatile environments.

Motivated by the arguments above, we propose a practical and effective technique for calculating concrete market prices and allocations via a *deep reinforcement learning policymaker*, operating in an environment of other learning agents. These prices can serve as a clear-cut guideline for intervention, and can then be implemented by a variety of mechanisms; e.g., policy-makers can tax (or subsidize) the difference between the current market price and the computed price, or buy/sell from reserves.[1] This new approach grants us the ability to abstract real-world situations into a form that makes them amenable to research, and allows for advances in the state-of-the-art. In particular, it enables us to design and test novel policies (via tuning the parameters and simulating the multi-agent environment) to tackle a plethora of real-world problems in various disciplines under a host of objectives, such as the problem of sustainable production (renewable energy, CO_2 markets, natural resource preservation, etc.). Our work falls into the very recent research agenda of building agent-based models to inform policy in socio-economic environments (see [17, 29]).

2 OUR CONTRIBUTIONS

We use deep reinforcement learning for policy making, and study the emergent behaviors as a group of deep learners interact in a complex and realistic market, where both the pricing policy and the harvesting behaviors are learned *simultaneously*. Neither the policy maker nor the harvesters have prior knowledge / assumptions of domain dynamics or economic theory, and every agent only makes use of information that it can individually observe. In particular:

(1) We propose a practical approach to computing market prices and allocations via a deep reinforcement learning policymaker agent, that allows us to *tune* the prices with regard to diverse objectives such as *sustainability* and resource *wastefulness*, *fairness* and buyers' and sellers' welfare. Our goal is to investigate the feasibility of using traditional deep reinforcement learning agents as (i) a practical alternative to classical notions of rationality and market equilibria, and (ii) a means to reach stable outcomes that are comparable with the idealized market equilibrium outcome from economics, while optimizing exogenous objectives.

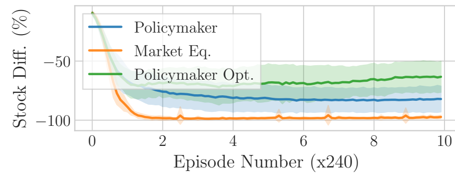

Figure 1: Sustainability (maximum negative deviation from the equilibrium stock) over the training episodes. The orange line is the market equilibrium prices (MEP). The blue line is the vanilla policymaker that optimizes each objective in the reward equally, while the green line optimizes sustainability. This is a *scarce resource environment.* The policymaker achieves a *dramatic improvement in sustainability.* The MEP maintain a population stock that is 97.3% below the equilibrium population (on average), while the policymaker that optimizes sustainability improves to 63.3%; almost 35% improvement. The MEP permanently *deplete* the resources in 9.79% of the episodes, with episodes lasting as low as 48 time-steps (out of 500). In contrast, the policymaker fails in only 2.24% of the episodes (min episode length of 258 time-steps).

(2) We introduce a novel multi-agent socio-economic environment and *prove* a necessary condition for market failure. Our environment combines established principles of competitive markets with the challenges of *resource scarcity* and the tragedy of the commons. This is paramount to understand the impact of self-interested appropriation and develop sustainable strategies. While we use a common-fishery as an indicative, real-world test-bench, our approach is general and can be employed in *any* production market. To demonstrate the need for intervention via our policymaker, we provide an analytical example where leaving the market entirely "free" to act according to the market equilibrium will result in the depletion of the resource in a short period of time. Our analysis also highlights the inherent challenges of theoretically computing optimal strategies for either the policymaker or the harvesters, and justifies modeling both as learning agents instead.

(3) We provide a thorough (quantitative & qualitative) analysis on the learned policies and demonstrate that they can achieve significant improvements over the market equilibrium benchmark for several objectives, while maintaining comparable performance for the rest. As a highlight of our results, we show that our policymaker fares notably better in terms of sustainability of resources, essentially without compromising any of the remaining objectives.

Given that it is often quite hard to experiment with real-world pricing policies, traditional work in economics often results to simplifying assumptions which are hard to validate. Our approach provides an alternative route, enabling experimentation (via tuning of the parameters and simulating the multi-agent environment) to find the best possible policies.

[1] There are many such examples of influencing the supply/demand [4, 7, 28].

REFERENCES

[1] Kenneth J Arrow and Gerard Debreu. 1954. Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society* (1954), 265–290.

[2] Francis M Bator. 1958. The anatomy of market failure. *The quarterly journal of economics* 72, 3 (1958), 351–379.

[3] William J Baumol. 1972. On taxation and the control of externalities. *The American Economic Review* 62, 3 (1972), 307–322.

[4] BBC. 2021. US to release oil reserves in attempt to lower prices. https://www.bbc.com/news/business-59353194 [Online; accessed 7-February-2022].

[5] Krishnendu Chatterjee, Rupak Majumdar, and Marcin Jurdziński. 2004. On Nash equilibria in stochastic games. In *International workshop on computer science logic*. Springer, 26–40.

[6] Colin W Clark. 2006. *The worldwide crisis in fisheries: economic models and human behavior*. Cambridge University Press.

[7] CNN. 2021. Facing shortages, Canada taps its strategic reserves of maple syrup. https://edition.cnn.com/2021/11/28/business/canada-maple-syrup-reserves-shortage/index.html [Online; accessed 7-February-2022].

[8] Thomas D Crocker. 1966. The structuring of atmospheric pollution control systems. *The economics of air pollution* 61 (1966), 81–84.

[9] JH Dales. 1968. Pollution, property & prices; an essay in policy-making and economics. (1968).

[10] Panayiotis Danassis, Aris Filos-Ratsikas, Haipeng Chen, Milind Tambe, and Boi Faltings. 2022. AI-driven Prices for Externalities and Sustainability in Production Markets. *CoRR* abs/2106.06060 (2022). arXiv:2106.06060 https://arxiv.org/abs/2106.06060

[11] Constantinos Daskalakis, Noah Golowich, and Kaiqing Zhang. 2022. The complexity of markov equilibrium in stochastic games. *arXiv preprint arXiv:2204.03991* (2022).

[12] Xiaotie Deng, Yuhao Li, David Henry Mguni, Jun Wang, and Yaodong Yang. 2021. On the complexity of computing markov perfect equilibrium in general-sum stochastic games. *arXiv preprint arXiv:2109.01795* (2021).

[13] EU. 2022. EU Emissions Trading System. https://ec.europa.eu/clima/eu-action/eu-emissions-trading-system-eu-ets_en [Online; accessed 7-February-2022].

[14] EU. 2022. EU Green Taxation. https://ec.europa.eu/taxation_customs/green-taxation-0_en [Online; accessed 7-February-2022].

[15] Irving Fisher. 1892. *Mathematical Investigations in the Theory of Value and Prices*. Ph.D. Dissertation. Yale University.

[16] CJ Himmelberg, Thiruvenkatachari Parthasarathy, TES Raghavan, and FS Van Vleck. 1976. Existence of p-equilibrium and optimal stationary strategies in stochastic games. *Proc. Amer. Math. Soc.* 60, 1 (1976), 245–251.

[17] Michael Bradley Johanson, Edward Hughes, Finbarr Timbers, and Joel Z. Leibo. 2022. Emergent Bartering Behaviour in Multi-Agent Reinforcement Learning. https://doi.org/10.48550/ARXIV.2205.06760

[18] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. 1996. Reinforcement learning: A survey. *Journal of artificial intelligence research* 4 (1996), 237–285.

[19] Jean-Jacques Laffont. 2008. Externalities. *New Palgrave Dictionary of Economics* (2008).

[20] Michael L. Littman and Peter Stone. 2002. *Implicit Negotiation in Repeated Games*. Springer Berlin Heidelberg, Berlin, Heidelberg, 393–404. https://doi.org/10.1007/3-540-45448-9_29

[21] OECD. 2020. OECD Review of Fisheries 2020. https://www.oecd-ilibrary.org/agriculture-and-food/oecd-review-of-fisheries-2020_7946bc8a-en [Online; accessed 7-February-2022].

[22] Arthur Cecil Pigou. 1924. *The Economics of Welfare*. Macmillan.

[23] Ariel Rubinstein and Carl-johann Dalgaard. 1998. *Modeling bounded rationality*. MIT press.

[24] Matthew J Sobel. 1971. Noncooperative stochastic games. *The Annals of Mathematical Statistics* 42, 6 (1971), 1930–1935.

[25] Eilon Solan and Nicolas Vieille. 2002. Correlated equilibrium in stochastic games. *Games and Economic Behavior* 38, 2 (2002), 362–399.

[26] Leon Walras. 1874. *Éléments d'Économie Politique Pure*.

[27] Wikipedia. 2022. Clean Air Act (United States) — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Clean_Air_Act_(United_States)&oldid=1069225593 [Online; accessed 7-February-2022].

[28] WPR. 2018. USDA Plans To Buy $50M In Milk To Reduce Surplus. https://www.wpr.org/usda-plans-buy-50m-milk-reduce-surplus [Online; accessed 7-February-2022].

[29] Stephan Zheng, Alexander Trott, Sunil Srinivasa, David C. Parkes, and Richard Socher. 2022. The AI Economist: Taxation policy design via two-level deep multiagent reinforcement learning. *Science Advances* 8, 18 (2022), eabk2607. https://doi.org/10.1126/sciadv.abk2607 arXiv:https://www.science.org/doi/pdf/10.1126/sciadv.abk2607