Price of Anarchy in a Double-Sided Critical Distribution System

David Sychrovský
Charles University
Prague, Czechia
sychrovsky@kam.mff.cuni.cz

Jakub Černý
Nanyang Technological University
Singapore
cerny@disroot.org

Sylvain Lichau
University of Bordeaux
Bordeaux, France
sylvain.lichau@etu.u-bordeaux.fr

Martin Loebl
Charles University
Prague, Czechia
loebl@kam.mff.cuni.cz

ABSTRACT

Measures of allocation optimality differ significantly when distributing standard tradable goods in peaceful times and scarce resources in crises. While realistic markets offer asymptotic efficiency, they may not necessarily guarantee fair allocation desirable when distributing the critical resources. To achieve fairness, mechanisms often rely on a central authority, which may act inefficiently in times of need when swiftness and good organization are crucial. In this work, we study a hybrid trading system called *Crisdis*, introduced by Jedličková et al., which combines fair allocation of buying rights with a market – leveraging the best of both worlds. A frustration of a buyer in Crisdis is defined as a difference between the amount of goods they are entitled to according to the assigned buying rights and the amount of goods they are able to acquire by trading. We define a Price of Anarchy (PoA) in this system as a conceptual analogue of the original definition in the context of frustration. Our main contribution is a study of PoA in realistic complex double-sided market mechanisms for Crisdis. The performed empirical analysis suggests that in contrast to market free of governmental interventions, the PoA in our system decreases.

KEYWORDS

Price of Anarchy; auction; fairness

1 INTRODUCTION

Most of the goods available to the general public are meant to increase the quality of life of individuals or count as luxuries, and are traded using standard market mechanisms. Other resources serve a more social purpose – when allocated well, they increase the well-being of the entire society like public housing, school seats, or healthcare products. Among those, some are desirable to be readily available to everyone, e.g., essential medicines, various equipment, or even vaccines that enable to reach herd immunity in the population only when enough people have developed protective antibodies against future infections. In times of need like disasters, local epidemics, or even conflicts and wars, these resources need to be distributed swiftly and in a highly organized manner to reach as many eligible people as possible in a limited timeframe.

Allocating such public resources is commonly reserved for governmental services and done at prices below market-clearing or even free of charge. However, leaving the competitive markets out of the allocation process often results in inefficiencies, both economic and temporal, caused by problems inherent to centralized planning [11]. On the other hand, real-world trading markets, frequently modeled as large double auctions with many sellers and buyers on each side, are capable of distributing the goods flexibly and reliably. The problem remains that even though, with increasing size, the participants are incentivized to be truthful (which leads to asymptotic efficiency [4]), the resulting goods reallocation is not necessarily socially optimal in terms of being available to everyone. Any discrepancy in wealth is then only exacerbated by crises similar to the coronavirus pandemic or the war in Ukraine we experienced in recent years. In such settings, scarce resources necessary for keeping the society up and running could be easily swayed by its more fortunate members, which has to be countered by carefully designed measures.

As an attempt to combine the best of both worlds, the following hybrid distribution system called Crisdis is suggested in [14]: a trustworthy central authority provides a marketplace where buyers and sellers engage in two-sided repeated trading over a period of many days. At the beginning of each market day, the authority allocates buying rights to the participating members (e.g., individual hospitals), which are traded together with the goods. Everything else (pricing, storage, delivery, etc.) is left up to the sellers and buyers themselves, with one requirement only: at the end of each trading day, each buyer needs to possess the number of rights greater or equal to the number of goods. The straightforward motivation for this arrangement is that the traders selling some of their assigned rights obtain extra funds, which they can use in future markets to satisfy their demand for the critical goods better. Another motivation is that the needs of individual participants used to allocate the rights can be evaluated independently by the central authority using real-time crisis data, thus sidestepping the bottleneck of many auction mechanisms – the proneness to strategic manipulation.

The utmost priority of Crisdis is to improve the accessibility of critical goods to all eligible buyers during crises in a trading system which is as realistic as possible. For this purpose, [14] introduced a measure of the social efficiency of the allocations realized by

the semi-distributed system called *frustration*. Frustration can be seen as a scaled negative difference between fairness and reality: for a participating buyer, it is the scaled difference between the (potential) allocation of rights to the buyer and the number of goods purchased by them if the value is at least zero, and zero otherwise. Assuming the market attains its equilibrium, the sum of frustrations of the traders describes the system's *Price of Anarchy*, i.e., the price the society pays for allocating the goods through the market and not directly as suggested by the fairness mechanism.

In [14], the authors study how frustration evolves during repeated interactions in the system under a single-sided auction mechanism based on activities of buyers. This work contributes by a thorough study of a more realistic double-sided mechanism.

1.1 Contributions

We study trading in a system consisting of a sequence of complex double-sided markets combined with a fairness mechanism for allocating the rights designed to improve social good. Following [14], we focus on the well-known and thoroughly studied *contested garment distribution* [2] for fairly allocating the rights[1], but expand on their work by analyzing multiple double-auction market mechanisms instead of a simple English auction. We introduce these double-auction mechanisms in Section 3, ranging from random acceptable allocations to maximum clearing under average-price bids.

Our priority in this work is to study the behavior of a large complex system, which makes it difficult to analyze the traders' behavior theoretically. For this reason, in Section 4, we introduce a reinforced-learning algorithm in an attempt to approximate the system's equilibrium. In Section 5 we present the empirical results. First, we perform a thorough numerical analysis demonstrating how close to the equilibrium we are able to converge to. Then we carry out a series of ablation experiments, showing that the Price of Anarchy in the system without the fairness mechanism may be high. We confirm that together with intuitive governmental regulations akin to increased storing prices for the goods, the system with the fairness mechanism is able to decrease the Price of Anarchy. In the last part of the paper, we summarize the desired features of the trading system enlightened by the experiments.

1.2 Related work

Our work belongs to the literature on redistributive mechanisms, especially those mitigating inequalities. Perhaps the most related paper studies a two-sided market trading goods of homogeneous quality, optimizing the traders' total utilities [5]. The difference lies in our explicit incorporation of the buyers' varying needs into the consideration and the fact that in our model, the utilities are a common knowledge. This work was recently generalized into a setting with heterogeneous quality of tradable objects, more diverse measures of allocation optimality, and imperfect observations about the traders [1]. Another related work presents multiple markets and non-market mechanisms for allocating a limited number of identical goods to several buyers [3]. The author shows that when the buyers' willingness to pay coincides with the designer's allocation preferences, market mechanisms are optimal, and vice versa. In

[1]Our experimental results show that this fairness mechanism performs well in practice.

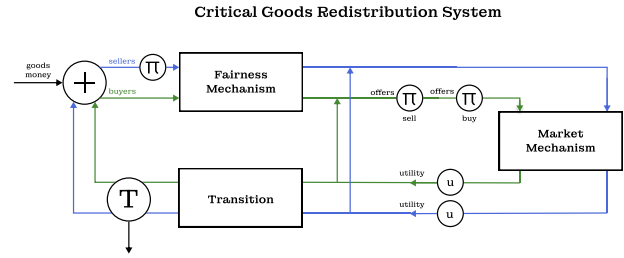Critical Goods Redistribution System

Figure 1: Fairness and market mechanisms positioned in a feedback loop of our redistribution system. One iteration of the outer loop corresponds to one Market. π refers to the strategies and T checks the termination condition.

crises environments studied in our work, it is reasonable to assume that the critical resources are highly valuable to all participants, yet, some may lack the money to obtain them. Together with the fact that it is in society's best interests to allocate the goods fairly, these results suggest that leaving the distribution solely to unregulated markets is rather inadvisable.

2 PROBLEM DEFINITION

We assume the existence of a centralized marketplace where critical goods are traded periodically during the entire distribution crisis among the buyers and the sellers using an internal currency. We consider only one type of good and call it Good. To simplify the presentation we assume here that the Good is *divisible*[2]. We refer to each trading period of the distribution crisis as a *Market*. A finite sequence of Markets then forms a *Crisis*. The structure of the entire system is depicted in Figure 1.[3]

In order to reduce volatility and promote fairness during trading, similarly as in [14], we introduce a new type of tradable resource called Right. In each Market, in order to buy the Good, the buyer also needs to possess an equivalent amount of the Right. The Rights are allocated to the buyers before the trading in a Market begins by a centralized *Fairness mechanism* using the sellers' declared offers and buyer's declared demands. The traders then engage in a series of interactions resulting in their announcement of bids. A dedicated *Market mechanism* then allocates the Goods and Rights based on the bids.

The residual resources of Good after consuming the demands, and the Money are then transferred to the next Market, as we model the shortage of the critical Good over an extended period of time. However, the residual of Right disappears after each Market.

2.1 One trading period: the Market

Formally, we model the trading of the goods during the Market as a double-auction represented as a parametric[4] imperfect-information sequential game $\mathbb{G}(\mathcal{M}, \mathcal{G}) = (\mathcal{T}, \mathcal{M}, \mathcal{G}, \mathcal{D}, \phi, \mu, u, k)$. We use calligraphic letter to denote ordered sets and non-calligraphic letters to denote individual elements of these sets. The set of traders $\mathcal{T}$

[2]The results easily generalize to periodic trading of larger quantities of indivisible Good; see [14] for a more detailed explanation.

[3]See full paper at https://arxiv.org/abs/2302.09959 for simple example.

[4]Why this market game and its strategies are parametric in $(\mathcal{M}, \mathcal{G})$ will become apparent later during the definition of the sequence of markets constituting the crisis.

consists of buyers $\mathcal{B}$ and sellers $\mathcal{S}$; the sets of buyers and sellers are assumed to be disjoint. The set $\mathcal{M} = (M_1, M_2, \ldots, M_{|\mathcal{T}|}) \in \mathbb{R}_0^{+,|\mathcal{T}|}$ determines the real and non-negative amount of money each trader receives at the beginning of the Market. Similarly, the set $\mathcal{G} = (G_1, G_2, \ldots, G_{|\mathcal{T}|}) \in \mathbb{R}_0^{+,|\mathcal{T}|}$ specifies the real and non-negative amount of Good each trader obtains.

The demands $\mathcal{D} = (D_1, D_2, \ldots, D_{|\mathcal{B}|}) \in \mathbb{R}_0^{+,|\mathcal{B}|}$ describe the ideal real and non-negative amount of Good each buyer hopes to acquire during the trading. Function ϕ then implements the fairness mechanism, assigning real-valued, non-negative Rights $\mathcal{R} = (R_1, R_2, \ldots, R_{|\mathcal{B}|}) \in \mathbb{R}_0^{+,|\mathcal{B}|}$ to the individual buyers.

To allocate the Rights, the mechanism needs to know the amount of Good put up for trade. This amount is given by the parametric strategies of the sellers. For each seller $s \in \mathcal{S}$, the set Π_s contains the seller's all possible parametric strategies, represented as functions

$$\pi_s : \mathbb{R}_0^{+,|\mathcal{B}|} \times \mathbb{R}_0^{+,|\mathcal{B}|} \times \mathbb{R}_0^+ \to \mathbb{R}_0^{+,2},$$

and interpreted as $\pi_s(\mathcal{M}_{\mathcal{B}}, \mathcal{G}_{\mathcal{B}}, G_s) = (v_s^G, p_s^G)$, where $v_s^G \leq G_s$ is the amount of Good offered at price p_s^G. Each π_s is hence a function of the amount of Good and Money each buyer possesses as well as the amount of Good the seller has, but not of the amount of Good available to the other sellers. This corresponds to sellers investing in some market research[5]. A profile of one strategy per seller is denoted as $\pi_{\mathcal{S}} \in \Pi_{\mathcal{S}}$. The traders' behavior in a specific game $\mathbb{G}(\mathcal{M}, \mathcal{G})$ is then written as $\pi_{\mathcal{S}}(\mathcal{M}, \mathcal{G})$. We also refer to their offers at once as $v_{\mathcal{S}}^G$ and $p_{\mathcal{S}}^G$. The fairness mechanism is defined as:

DEFINITION 1. *For any sellers' strategy profile $\pi_{\mathcal{S}}$, the fairness mechanism is a function $\phi : \mathbb{R}_0^+ \times \mathbb{R}_0^{+,|\mathcal{B}|} \to \mathbb{R}_0^{+,|\mathcal{B}|}$ allocating Rights to each buyers, satisfying in each game $\mathbb{G}(\mathcal{M}, \mathcal{G})$*

$$\sum_{b \in \mathcal{B}} \phi_b(V, \mathcal{D}) = V \qquad \forall \mathcal{D} \in \mathbb{R}_0^{+,|\mathcal{B}|},$$

$$D_b = 0 \Rightarrow \phi_b(V, \mathcal{D}) = 0 \qquad \forall \mathcal{D} \in \mathbb{R}_0^{+,|\mathcal{B}|}, \forall b \in \mathcal{B}$$

$$\phi_b(V, \alpha(\mathcal{D})) = \phi_{\alpha^{-1}(b)}(V, \mathcal{D}) \qquad \forall \mathcal{D} \in \mathbb{R}_0^{+,|\mathcal{B}|}, \forall b \in \mathcal{B}, \forall \alpha,$$

where α is a permutation of $\mathcal{B}$ and $V(\pi_{\mathcal{S}}, \mathcal{M}, \mathcal{G}) = \sum_{v_s^G \in \pi_{\mathcal{S}}(\mathcal{M}, \mathcal{G})} v_s^G$.

In this work, we focus on a fairness mechanism implementing the *contested garment distribution* (CGD) [2] as a well-studied baseline fairness method, and refer to the assigned Rights further on in the text as $\mathcal{R} = \phi(V(\pi_s, \mathcal{M}, \mathcal{G}), \mathcal{D})$.

Function μ is of more importance to us, allocating the resources after the bidding phase. The bidding is determined by the strategies of the buyers, who act in two simultaneous-move phases. First, after observing the seller's offers as well as the amount of Good, Right and Money they have, each buyer declares the amount and price of Right (v_b^R, p_b^R) they are offering for sale. Second, after observing the offers of other buyers, each buyer declares the amount and price of Good and Right ($\bar{v}_b^R, \bar{p}_b^R, \bar{v}_b^G, \bar{p}_b^G$) they are willing to buy. We assume the announced prices can be interpreted as either maximum or average (if buying from multiple sellers) prices, and we construct

different μ mechanisms for each. A strategy of each seller can be then formulated in both stages at once as a function

$$\pi_b : \mathbb{R}_0^{+,|\mathcal{S}|} \times \mathbb{R}_0^{+,|\mathcal{S}|} \times \mathbb{R}_0^+ \times \mathbb{R}_0^+ \times \mathbb{R}_0^+ \times \mathbb{R}_0^{+,|\mathcal{B}|} \times \mathbb{R}_0^{+,|\mathcal{B}|} \to \mathbb{R}_0^{+,6},$$

written as $\pi_b(v_{\mathcal{S}}^G, p_{\mathcal{S}}^G, M_b, G_b, R_b, v_{\mathcal{B}}^R, p_{\mathcal{B}}^R) = (v_b^R, p_b^R, \bar{v}_b^R, \bar{p}_b^R, \bar{v}_b^G, \bar{p}_b^G)$. For each buyer $b \in \mathcal{B}$, their set of strategies is Π_b. A profile of one strategy per buyer is denoted as $\pi_{\mathcal{B}} \in \Pi_{\mathcal{B}}$. We write the Cartesian product $\Pi_{\mathcal{S}} \times \Pi_{\mathcal{B}}$ as Π and denote $\pi \in \Pi$ a profile of all traders' strategies. The market mechanism is then formally defined as:

DEFINITION 2. *For any traders' strategy profile π in a game $\mathbb{G}(\mathcal{M}, \mathcal{G})$, the market mechanism is a function $\mu : \Pi \times \mathbb{R}_0^{+,|\mathcal{T}|} \times \mathbb{R}_0^{+,|\mathcal{B}|} \times \mathbb{R}_0^{+,|\mathcal{T}|} \to \mathbb{R}_0^{+,|\mathcal{T}|} \times \mathbb{R}_0^{+,|\mathcal{B}|} \times \mathbb{R}_0^{+,|\mathcal{T}|}$ which returns a realization of trades, i.e., a reallocation of Good, Right and Money among the traders at the end of the Market. We further require that μ satisfies that*

(1) *no trader sells more Good or Right than they offer;*
(2) *no buyer buys more Good or Right than they declare;*
(3) *no trader sells Good or Right for a lower than the asking price;*
(4) *no buyer buys Good or Right for a higher (or higher on average) price than is their bidding price; and*
(5) *no buyer buys Rights from themselves.*

We abuse the notation a little and write $\mu^G(\cdot)$ and $\mu^M(\cdot)$ to refer to the restrictions to reallocated Goods and Money, respectively.

Note that the last condition ensures that the desired amount of Right is actually what a buyer would expect. Without it, the buyers can trade virtually with themselves and thus get a lower amount of Right from the Market, even if they could buy more. The choice of the market mechanism affects the strategizing of the traders to a great extent. We hence dedicate the entire next section to the study of multiple such mechanisms.

What remains is to define the utility functions $u = (u)_{t \in \mathcal{T}}$, $u_t : \Pi \times \mathbb{R}_0^{+,|\mathcal{T}|} \times \mathbb{R}_0^{+,|\mathcal{T}|} \to \mathbb{R}$ for each trader t. The sellers are motivated solely by profit. Thus, the utility they get from the Market is the amount of Money they receive. We refine this simple model by *adding negative utility* for the Good the seller has at the end of each Market. This penalty represents the societal desire for the sellers to sell most of the available critical Good, and may be implement, e.g., through the state penalties which are usually in place during crises[6]. Moreover, in case the Market terminates the Crisis, the sellers obtain also a small additional utility compensating for the Good they still keep stocked. Formally, for each $s \in \mathcal{S}$,

$$u_s(\pi, \mathcal{M}, \mathcal{G}) = \begin{cases} \mu_s^M(\pi, \mathcal{M}, \mathcal{R}, \mathcal{G}) + C_1 \mu_s^G(\pi, \mathcal{M}, \mathcal{R}, \mathcal{G}) & \text{NT,} \\ u_s\text{-NT} + C_2 \mu_s^G(\pi, \mathcal{M}, \mathcal{R}, \mathcal{G}) & \text{T,} \end{cases} \quad (1)$$

where NT/T denote non-terminal/terminal markets, and C_1 and C_2 are suitable constants. The utility of a buyer should incentivize them to keep a steady supply of Good throughout the Crisis. Therefore, after each trading period, they receive utility for the Good they have (up to their demand), which represents their regular consumption (e.g., per day). The buyers also receive some small utility C_3 per

---

unit of Money they have at the end the Crisis. Formally, for $b \in \mathcal{B}$,

$$u_b(\pi, \mathcal{M}, \mathcal{G}) = \begin{cases} \min\left\{D_b, \mu^G(\pi, \mathcal{M}, \mathcal{R}, \mathcal{G})\right\} & \text{NT}, \\ u_b\text{-NT} + C_3\mu^M(\pi, \mathcal{M}, \mathcal{R}, \mathcal{G}) & \text{T}. \end{cases} \quad (2)$$

The semantic meaning of the additional constants in terminal-state utilities is to compensate for the possible continuation of the Crisis.

Now let us reiterate and describe again the entire process of how sellers and buyers engage in trading in our two-sided market:

Sellers' stage:
(1) Each seller declares the amount and price of Good for sale.
(2) Each buyer is assigned Rights by the fairness mechanism ϕ.

Buyers' stage:
(3) Each buyer declares the amount of Right they are willing to sell along with the asking price.
(4) Each buyer, given the available amounts and asking prices of Good and Right, declares their *bidding price and desired amount* of Good and Right, separately.
(5) The bids are cleared using the market mechanism μ.
(6) The traders receive their utilities.

Note that an important aspect of our model is that the buyers can use the Money they obtained only in the next Market of the sequence. This gives the active buyers advantage of buying the critical Good earlier than the passive buyers; the price of this advantage is the cost of buying additional rights.

2.2 Sequence of Markets: the Crisis

We assume that trading takes place periodically, in a finite sequence of Markets, denoting, e.g., trading days. Formally, the Crisis is a partially observable stochastic game with continuous state space and continuous action space, denoted as $\mathbb{C} = (\mathcal{T}, \mathcal{D}, \phi, \mu, u, \mathcal{M}^1, \mathcal{G}^1, T, P)$. $\mathcal{T}, \mathcal{D}, \phi, \mu$ and u have the same meaning as in our definition of the Market game and are assumed to be fixed throughout the entire Crisis. Each state in this stochastic game corresponds to a Market game, with the initial game being $\mathbb{G}(\mathcal{M}^1, \mathcal{G}^1)$. We model a full-blown crisis[7], and assume all traders are stationary-Markovian, basing their strategies on their observations in the Crisis' current state, i.e., a game $\mathbb{G}$ with some parameters $(\mathcal{M}, \mathcal{G})$. The parametric strategies introduced in the previous section are hence employed in the entire Crisis. After each non-terminal trading, the sellers keep the unsold amount of Good and the buyers keep the Money and unconsumed amount of Good for the next Market. In contrast, the unused amount of Right is disposed of after each Market terminates. The traders also receive additional Goods or Money before the next Market starts. This process is described via the transition function $P : \Pi \times \mathbb{R}_0^{+,|\mathcal{T}|} \times \mathbb{R}_0^{+,|\mathcal{T}|} \to \mathbb{R}_0^{+,|\mathcal{T}|} \times \mathbb{R}_0^{+,|\mathcal{T}|}$. Using the same notation as for the market mechanism μ, we write

$$P_t^G(\pi, \mathcal{M}, \mathcal{G}) = \mu_t^G(\pi, \mathcal{M}, \mathcal{R}, \mathcal{G}) + \begin{cases} G_t^1 & \text{if } t \in \mathcal{S} \\ 0 & \text{if } t \in \mathcal{B}, \text{ and} \end{cases}$$

$$P_t^M(\pi, \mathcal{M}, \mathcal{G}) = \begin{cases} 0 & \text{if } t \in \mathcal{S} \\ \mu_t^M(\pi, \mathcal{M}, \mathcal{R}, \mathcal{G}) + M_t^1 & \text{if } t \in \mathcal{B}. \end{cases}$$

---
[7]We leave the study of boundary situations for future work.

The finite horizon is denoted T. As is usual, a situation in which no trader has an incentive to unilaterally change their strategies is called an *equilibrium*. We consider optimality under discounting γ.

DEFINITION 3. *Let $\pi \in \Pi$ induce a series of Markets $((\mathcal{M}^1, \mathcal{G}^1), \ldots, (\mathcal{M}^T, \mathcal{G}^T))$. We call π an equilibrium if for any unilaterally deviating profile $\hat{\pi} \in \Pi$ and the corresponding series of Markets $((\hat{\mathcal{M}}^1, \hat{\mathcal{G}}^1), \ldots, (\hat{\mathcal{M}}^T, \hat{\mathcal{G}}^T))$ it holds for all traders t that*

$$\sum_{i=1}^{T} \gamma^i u_t(\pi, \mathcal{M}^i, \mathcal{G}^i) \geq \sum_{i=1}^{T} \gamma^i u_t(\hat{\pi}, \hat{\mathcal{M}}^i, \hat{\mathcal{G}}^i).$$

2.3 Frustration and the Price of Anarchy

Traditionally, the efficiency of a designed system is measured via the notion of Price of Anarchy, which characterizes how much an equilibrial state reached by self-interested decision-makers differs from an optimal state. The optimality criterion the system's planner usually considers is the Pareto-optimal social-welfare, i.e., an optimal state maximizes the sum of individual players' utilities.

In our work, we borrow this idea, but consider a slightly different objective of the planner. In crises, instead of blindly maximizing all utilities, we are more interested in maintaining an egalitarian society. We hence study how the amount of Good acquired by buyers evolves for different Market mechanisms and compare it to the amount of Rights assigned to them. The resulting discrepancy describes the inherent inequality in the system, formally defined as *frustration*, and serves to measure (conceptually) the same quantity as in traditional Price of Anarchy: the price the system pays for freedom of agents to choose, instead of planning centrally.

DEFINITION 4. *The frustration of buyer b in Market $(\mathcal{M}, \mathcal{G})$ under profile $\pi \in \Pi$ and allocated rights $\mathcal{R} = \phi(V(\pi_S, \mathcal{M}, \mathcal{G}), \mathcal{D})$ is then*

$$f_b(\pi, \mathcal{M}, \mathcal{G}) = \max\left\{\frac{R_b - \mu_b^G(\pi, \mathcal{M}, \mathcal{R}, \mathcal{G})}{R_b}, 0\right\}.$$

Our Price of Anarchy in the system is then the normalized accumulated frustration the buyers experience in the sequence of $\tau \leq T$ Markets when the equilibrium π is reached and induces a series of Markets $((\mathcal{M}^1, \mathcal{G}^1), \ldots, (\mathcal{M}^\tau, \mathcal{G}^\tau))$, i.e.,

$$PoA^\tau = \frac{\sum_{i=1}^{\tau} \sum_{b \in \mathcal{B}} f_b(\pi, \mathcal{M}^i, \mathcal{G}^i)}{\tau|\mathcal{B}|}. \quad \text{(PoA)}$$

3 MARKET MECHANISMS

In this section, we study how to reallocate the resources in the Market by introducing four mechanisms that schedule individual trades based on the inputs (bids) of the traders. Formally, clearing constraints (compatibility of asking and bidding prices and possibly other constraints) will be represented by two bipartite graphs: $G_G = (\mathcal{B}, \mathcal{S}, E_G)$ which represents the compatibility for trading the Good and $G_R = (\mathcal{B}_S, \mathcal{B}_B, E_R)$ which represents the compatibility for trading the Right. Here, $\mathcal{B}_S$ and $\mathcal{B}_R$ are disjoint copies of $\mathcal{B}$, $\mathcal{B}_S$ represents the sellers of Right and $\mathcal{B}_R$ represents the buyers of Right. A trader of $\mathcal{B}$ can be both a seller and a buyer of Right, but G_R does not connect their representing vertices by an edge. Both G_G, G_R are equipped with a positive real *weight* $w_G : V_G \to \mathbb{R}$ and $w_R : V_R \to \mathbb{R}$. The weights of the vertices naturally represent

the individual amount (of Good or Right) offered for sale and the individual amount (of Good or Right) desired to buy.

We primarily focus on *absolute* mechanisms, i.e. those that prohibit *any* trades where the bidding price is larger than the asking price, i.e., $\overline{p}_b^G > p_s^G$ and/or $\overline{p}_b^R > p_{b'}^R$. At the end of this section, we introduce a mechanism which relaxes this condition to hold for *average* prices.

3.1 Random allocation

A simple random trading mechanism used for purchasing both Rights and Goods proceeds as follow. First, the buyers are randomly permuted. In this order, each buyer is given randomly permuted lists of offers of the traders for Good and Right, respectively. A buyer first trades Good with sellers in order given by the list, until he has no Right left. In the second stage, the buyer trades Good and Right in equal amount, again following the list. This continues until they buy in total their desirable volume, or there are no more offers. We also ensure at every step that the asking price is lower than their acceptable price, and the buyer purchases amount up to the amount offered by the other party.

This mechanism has a unsatisfactory property. Since the buyers are presented with offers in random order, they often do not buy the cheapest option. This can be realistic since no single buyer will be able to see all the offers and choose among them. However, if the trading proceeds sequentially, it is natural for the buyer to consider the cheapest offers first. This also gives incentive to the sellers and traders to make offers at a lower price.

3.2 Greedy allocation

This algorithm is a modification of the random allocation which aims to address the issues mentioned in the previous paragraph. At the beginning, the buyers are sorted by the acceptable price of Good $\overline{p}_b^G$ in descending order. The mechanism has again two stages for each buyer. In the first stage, a buyer uses the Rights allocated to them to buy Goods, starting with the cheapest offer. When they have no Right left, they buy the same amount of Rights and Goods, again starting with the cheapest offers for both. We proceed until there is no offer left, or the buyer bought their desirable volume and continue with the next buyer.

Note that the random and greedy allocations are heuristics that are easily implementable but do not necessarily lead to optimal allocation which clears a maximum amount of bids.

3.3 Maximum clearing using absolute prices

An allocation clearing maximum amount of bids where we also require that no Right is bought without buying equal amount of Good, can be obtained using network flows. We call a mechanism utilizing this approach *Maximum clearing*. Its advantage is that it works also for indivisible Good. Another advantage is that the result of the Maximum clearing allocation is the list of individual tradings with compatible asking and bidding prices. The final price of each individual trading may be chosen in various ways from this compatibility interval.

THEOREM 1. *Maximum clearing allocation can be found efficiently using a reduction to the Max Flow problem. As a consequence, a*

Algorithm 1 Equilibrium Learning Algorithm

1: $\mathcal{B} \leftarrow$ *set of buyers*, $\mathcal{S} \leftarrow$ *set of sellers*, $\mathbb{B} \leftarrow \{\}$
2: **for** *episode* $\in \{1, \dots N_{\text{sims}}\}$ **do**
3:     **for** $\tau \in \{1, \dots T\}$ **do**
4:         $G_s \leftarrow G_s + G_s^1, M_b \leftarrow M_b + M_b^1$
5:         $o_{\mathcal{S}} \leftarrow$ observation of sellers
6:         $\pi_{\mathcal{S}} \leftarrow \text{clip}(\Pi_{\mathcal{S}}(o_{\mathcal{S}}(s)), 0, 1)$
7:         Allocate Right according to ϕ
8:         $o_{\mathcal{B}} \leftarrow$ observation of buyers
9:         $\overline{\pi}_{\mathcal{B}} \leftarrow \text{clip}(\Pi_{\mathcal{B}}(o_{\mathcal{B}}(b), \pi_{\mathcal{S}}), 0, 1)$
10:        $\pi_{\mathcal{B}} \leftarrow \text{clip}(\Pi_{\mathcal{B}}(o_{\mathcal{B}}(b), \pi_{\mathcal{S}}, \overline{\pi}_{\mathcal{B}}), 0, 1)$
11:        Trade according to a market mechanism μ
12:        Compute utilities $u_{\mathcal{B}}, u_{\mathcal{S}}$
13:        $\mathbb{B} \leftarrow \mathbb{B} \cup \{o_{\mathcal{B}}, o_{\mathcal{S}}, \pi_{\mathcal{S}}, \pi_{\mathcal{B}}, u_{\mathcal{B}}, u_{\mathcal{S}}\}$
14:        $\mathcal{G}_{\mathcal{B}} \leftarrow \max(\mathcal{G}_{\mathcal{B}} - \mathcal{D}_{\mathcal{B}}, 0)$
15:        **if** τ mod N_{train} is zero **then**
16:            Sample *batch* $\sim \mathbb{B}$
17:            Train on *batch* using TD3
18:        **end if**
19:     **end for**
20:     Reset episode
21: **end for**

Maximum clearing allocation is polynomial for both divisible and indivisible Good.

PROOF SKETCH. Construct the network of the Max Flow problem combining G_G and G_R.[8]   □

3.4 Maximum clearing using average prices

In contrast to the Market mechanisms presented so far, in this variant of the maximum clearing mechanism, we view the prices $\overline{p}_b^G$ and $\overline{p}_b^R$ as maximum *average* prices b is willing to pay. This is a relaxation of the maximum clearing mechanism with absolute bids, as it allows for trades which were previously infeasible. Again, the solution can be computed efficiently.

THEOREM 2. *Maximum clearing allocation with average bids can be found efficiently using a linear program.*

PROOF SKETCH. The linear program maximizes the volume of Good sold, while ensuring the average selling price of Good and Right is at most the desired price. The linear program can be solved in polynomial time.[8]   □

Repeated buyers' stage. To further improve clearing during one iteration of a Market, we present a simple extension where the *buyers' stage* is repeated k-times using the same strategy profile π. In each round the Goods, Rights, and Money are reallocated according to a selected Market mechanism.

4 LEARNING THE GAME'S EQUILIBRIUM

In this section, we describe a reinforced learning algorithm we use to obtain an approximation of the equilibrium of the Crisis. We

---

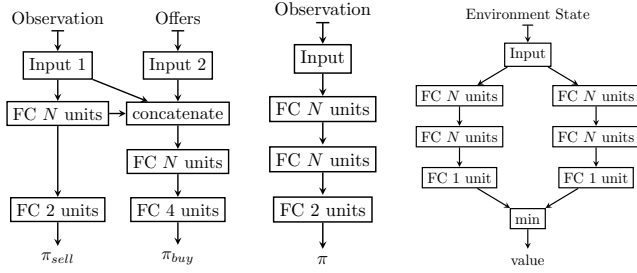[8] See full paper at https://arxiv.org/abs/2302.09959 for the entire proof.

Figure 2: Architectures of the used neural networks: (Left) sellers' actor, (Middle) buyers' actor, and (Right) the critic.

treat the entire interaction as a multi-agent reinforcement learning (MARL) problem as it is common in the literature [6, 10, 12, 13], with the assumption that the learning algorithm shall converge to a solution close to the equilibrium. We further verify the quality of the solution by computing its exploitability [9]. The sellers and buyers are represented as agents who interact in the environment described in sections 2 and 3. Each agent is trained to maximize their own expected discounted future utility in this environment.

4.1 Utilities in the learning environment

Next, we focus on the traders' utilities. To identify them, we need to specify constants C_i. For simplicity and to reduce the action space, we assume there exists a maximum price $\overline{P}$ the Good and Right can be offered at. Since the offered volume is bounded by the volume owned by a trader, the traders' actions fall in a closed interval.

Let us focus on the sellers first: their utility is given by two constants representing the price of storing the Good, and the expected future utility for the amount of Good in the terminal Market. We set the latter to be the market clearing price. This means the sellers expect to sell the Good for at least that price, which is a reasonable assumption during a crisis. The price of storing, C_1, may be chosen arbitrarily; however, it needs to be sufficiently high. If $|C_1|\frac{T}{2} < C_2$, it becomes beneficial for the sellers to keep the Good, and the selling price would thus be $\overline{P}$.

The buyers' utility is given in terms of the future expected utility for Money in the terminal Market. The relative penalty influences the mean utility a buyer obtains and again, it may be chosen arbitrarily. The future utility for Money is the utility for Good attainable with that Money, which is at least the utility for Good purchased at the maximum price $\overline{P}$. C_3 should hence inversely depend on $\overline{P}$.

4.2 Learning algorithm and model architecture

For training the agents' strategies we adopt an *actor-critic* algorithm called Twin-Delayed Deep Deterministic Policy Gradient (TD3) [7]. We modify it for the purpose of finding a solution in our scenario and depict the pseudocode in Algorithm 1. The policy Π_t of each trader t is a random variable with a Gaussian distribution with mean and standard deviation parameterized by a neural network.

The architecture of neural networks we employ is shown in Figure 2. The buyers' actor needs to process the offers of the sellers and consecutively offer the Right for sale before processing the offers of other buyers. To accomplish that, the output of the first

hidden layer is concatenated with the offers of the other buyers, and only the first hidden layer is used to predict the buyer's offer. In this way, the network can be used to obtain the buyer's offer without the offers of others influencing the result. Since the actions come from a bounded interval, the actors use a sigmoid activation function on the output layer on the means, which is then properly rescaled. The standard deviation uses the softplus activation.

Moreover, we enhance the vanilla TD3 algorithm with upgoing policy update [15] and reward clipping to $[-1, 1]$. To accelerate training, we allow the sellers to share the same replay buffer $\mathbb{B}$. This makes sellers' policies similar without using an identical actor.

5 EMPIRICAL EVALUATION

Finally, we demonstrate the properties of our hybrid system with fairness and market mechanisms, and the effectiveness of our learning algorithm, on practical examples. First, we assess the quality of the learned solutions using NashConv, a measure of exploitability. In the second part, we study how the learning algorithm scales with the number of traders. Lastly, we analyze to which degree the incorporation of Rights affects the Price of Anarchy of the approximated equilibrium throughout the entire crisis.

We evaluate systems combining three degrees of fairness with all four market mechanisms from Section 3. The variants of fairness we consider are systems with: (i) no distributed Rights (i.e., a free market), (ii) Rights and k=1 tradings; and (iii) Rights and k=2 tradings.

Experimental setting. All experiments were conducted on a computational cluster with AMD EPYC 7532 CPUs running at 2.40GHz. We utilized only 5 of its 16 cores and 3GB of RAM. The code was implemented in Python using tensorflow 2.6, tensorflow-probability 0.15, mip 1.14, and numpy 1.21. The open-source CBC solver carried out all LP computations. The complete list of all hyperparameters of Algorithm 1 can be found in the full version of the paper[9]. The code is available at github[10].

Experimental domain. We consider a sequence of $T = 10$ Markets with four buyers and four sellers. We choose a prototypical setting where three of the four buyers receive significantly more funds then the last buyer. At the same time, this last buyer suffers a large demand, in most cases exceeding the demands of the others. We refer to the first three buyers as *rich* and to the last buyer as *poor*. We generate the instances of this setting by sampling the demands and the earnings of the buyers uniformly randomly from given intervals. For the rich buyers, the demand D_b in drawn from $\mathcal{U}(1, 2)$ and the earning $M_b^1 \sim \mathcal{U}(4, 6)$. For a poor buyer, $D_b \sim \mathcal{U}(4, 6)$ and $M_b^1 \sim \mathcal{U}(1, 2)$. The set of demands and earnings is then normalized such that $\mathbb{E}_{b\sim\mathcal{B}}[D_b] = 1$ and $\mathbb{E}_{b\sim\mathcal{B}}[M_b^1] = 1/8$. To fix a scale[11], we set the maximum price as $\overline{P} = 1$. The constants in the utilities are then chosen as $C_1 = -1/8$, $C_2 = 1/2$, $C_3 = 1/\overline{P} = 1$ and $G_s^1 = 1/|\mathcal{S}|$.

---

[9]See https://arxiv.org/abs/2302.09959.
[10]The repository can be found at github.com/DavidSych.
[11]This corresponds to choosing a currency s.t. the price of a unit of Good is at most 1.
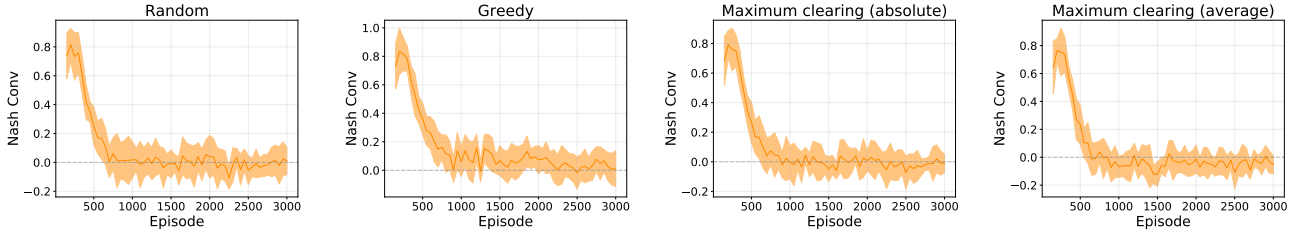
Figure 3: The exploitability of candidate solutions when learning the equilibria in systems with Rights and k=1 trading period for four different market mechanisms. The graphs show means and standard deviations over ten random seeds.
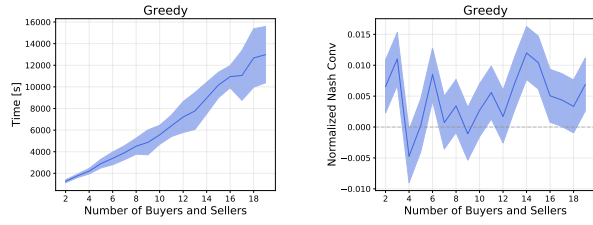

Figure 4: Computational time to learn equilibria and normalized final exploitability as functions of number of traders.

5.1 Exploitability

We measure the quality of a candidate solution from episode τ through its exploitability. For computing the exploitability we employ the notion of NashConv [9], given as

$$\sum_{i=1}^{T} \sum_{t \in \mathcal{T}} u_t(\pi_{-t}^{\tau} \cup \overline{BR}_t(\pi^{\tau}), \overline{\mathcal{M}}^{i_t}, \overline{\mathcal{G}}^{i_t}) - u_t(\pi^{\tau}, \mathcal{M}^i, \mathcal{G}^i).$$

Here, we denote by $\overline{BR}_t(\cdot)$ an approximate best response of trader t. The complement strategies are then written as π_{-t}. The corresponding sequences of Markets for π^{τ} and $\pi_{-t}^{\tau} \cup \overline{BR}_t(\pi^{\tau})$ are denoted as $(\mathcal{M}^i, \mathcal{G}^i)$ and $(\overline{\mathcal{M}}^{i_t}, \overline{\mathcal{G}}^{i_t})$, respectively. We train a best-response of each trader separately for 100 episodes, keeping the opponents' policies fixed, and starting from the policy of trader t in π^{τ}. In Figure 3 we present the results achieved with all four market mechanism in a system with Rights and $k = 1$, averaged over 10 runs with different random seeds. The results suggest the algorithm was able to reach a sufficiently close approximation of the equilibrium. Moreover, we verified the inclusion of Rights or the value of k do not have a significant effect on exploitability.

5.2 Scalability

In the second experiment we explore the computational time as a function of the number of participants. More specifically, we study Crises with N buyers and N sellers. In these simulations, we generate demands and earning as described earlier, but for more traders. Each buyer has a chance to be rich with probability 3/4 and poor otherwise. The computational time required for each N is depicted in Figure 4 on the left, for a system with the greedy market mechanism and fairness mechanism present, using $k = 1$.

For each value of N we sampled 10 instances. The algorithm exhibits near-linear time complexity in the number of buyers and sellers N.

To show that we reached an approximate equilibrium even with a large number of traders, we again evaluate the NashConv. However, computing the NashConv takes a significant portion of the computational time. To investigate purely the scalability of the algorithm we evaluate it once at the end of the training. The NashConv values, normalized by the number of buyers to facilitate comparison, are shown in Figure 4 on the right. The results suggest we reached close approximations of the equilibria.

5.3 Price of Anarchy over Markets

Here we present our main results: the empirical study of how the Price of Anarchy evolves in our hybrid system, in comparison to an intervention-free market. The results are depicted in Figure 5, and they show the prices the society pays for distributing the critical Goods through a (regulated) market instead of centrally. All results are averaged over 10 instances and show also the standard errors. The top row compares the Prices of Anarchy of systems with the three earlier described modes of fairness for the four introduced Market mechanisms. Note that the PoA is always lower in the systems with Rights. Moreover, introducing a second trading period further decreases it. Another noteworthy observation is that maximum clearing allocations offer lower PoA than the other two, more basic mechanisms.

The bottom row then shows the individual frustrations of the buyers. As expected, the poor buyer experiences the highest frustration. Otherwise the results observed with overall PoA clearly translate into the frustration of each buyer as well. Interestingly, the results suggest that introducing the fairness mechanism into the trading is beneficial not only for the poor buyer but for the rich buyers as well.

5.4 Price of Anarchy over buyers and sellers

We generate random instance of Crises with N buyers and N sellers for each $N \in \{2, \ldots, 19\}$ from the experimental domain described above. The PoA as a function of N in Figure 6. For each N we generate 10 instances and the graphs show also the standard errors. The same general trend we reported above clearly translate over all the mechanism here as well: (i) *the PoA is persistently the largest in the free market*, as opposed to the markets with Rights, and (ii) the market with $k = 2$ buyers' stages dominates the market with $k = 1$ stage. However, the results also indicate that different
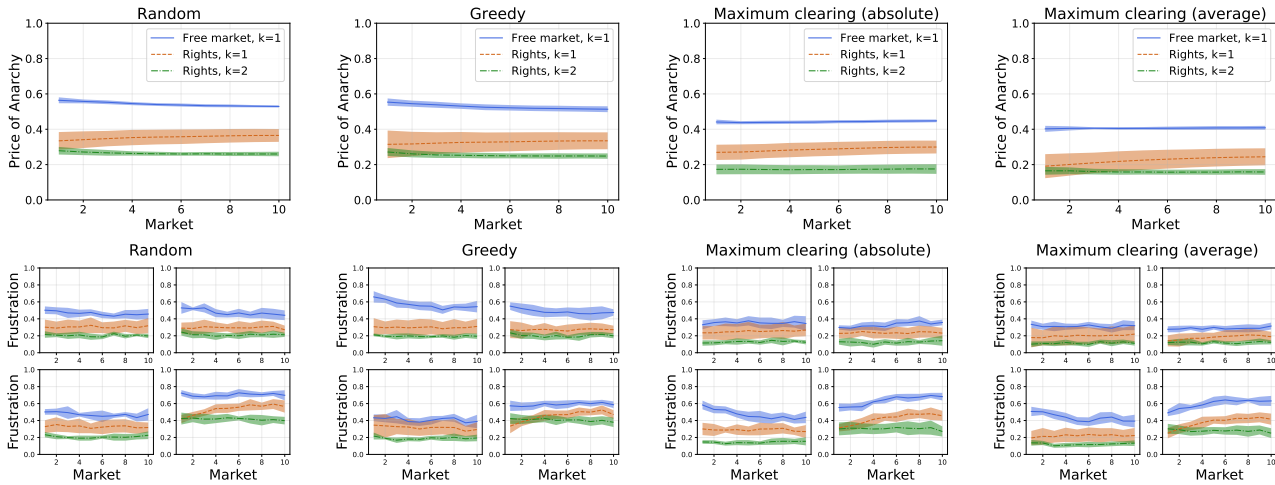
Figure 5: (Top) The Price of Anarchy and (Bottom) the individual frustrations of the four buyers in systems with three variants of fairness for four different market mechanisms. The poor buyer is always bottom right in the frustration graphs.
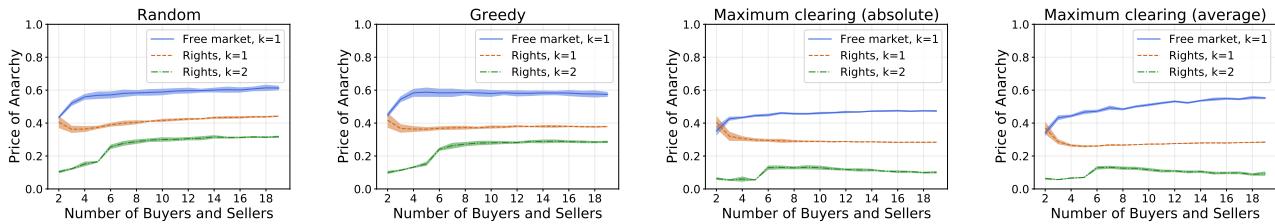

Figure 6: The Price of Anarchy as a function of the number of buyers and sellers in systems with three variants of fairness for four different market mechanisms.

market mechanisms perform vastly differently. With the **random** mechanism, the PoA seem to slowly yet gradually increase with the number of traders, while in the **greedy** mechanism the PoA reaches a constant. The most interesting cases are the **maximum clearing** mechanisms. With both absolute and average prices, the PoA in the free market keeps increasing, while the PoA in the system with Rights and $k = 1$ iteration of the buyers' stage mostly stagnates. In the redistribution system with $k = 2$ buyers' stages, the PoA after $N = 6$ consistently decreases, reaching a PoA value ≈ 5.93-times smaller than in the free market.

6 CONCLUSION

To the best of our knowledge, we are the first to introduce a system explicitly combining a *double-sided* market mechanism with a fairness mechanism allocating buying rights for more socially just redistribution of critical goods during the times of need. We adopted the contested garment distribution as a baseline method for fair allocation of rights and studied four separate market mechanism for trading: random, greedy, absolute-prices maximum clearing, and average-prices maximum clearing. Our two main theoretical results show that the last two allocations can be computed in polynomial time. We then defined an analogue of Price of Anarchy (PoA) in our system as a sum of so-called individual frustrations, which are

scaled differences between the amount of goods each trader was entitled to according to the fairness mechanism and the amount they were actually able to secure in the market. Furthermore, we developed a reinforcement-learning algorithm capable of approximating an equilibrium of the system in order to evaluate the PoA in practice. In the last part of our work, we show on a notorious example of a system with an underfunded and short-supplied buyer that introducing the buying rights may significantly decrease the frustrations, ergo, the PoA, especially for mechanisms prioritizing the amount of goods sold. Yet, it still remains an open question whether there exists a mechanism admitting zero PoA in the limit.

Future work. We see two major limitations of our work. First, we focused on the full-blown crises and assumed a small, but constant resupply of goods and money over many trading periods. We would like to study the consequences of more complex system dynamics akin to, e.g., the bullwhip effect caused by a steep increase in demand at the beginning of a crisis. Second, we restricted our fairness model to the contested garment distribution. Considering other models may change the system dynamics, and perhaps improve the PoA. Moreover, designing an optimal fairness rule may be done similarly as for voting mechanisms in the manner of [8].

ACKNOWLEDGMENTS

REFERENCES

[1] Mohammad Akbarpour, Piotr Dworczak, and Scott Duke Kominers. 2020. Redistributive allocation mechanisms. *Available at SSRN 3609182* (2020).

[2] R.J. Aumann and M. Maschler. 1985. Game Theoretic Analysis of a Bankruptcy Problem from the Talmud. *Journal of Economic Theory 36, 195-213* (1985).

[3] Daniele Condorelli. 2013. Market and non-market mechanisms for the optimal allocation of scarce resources. *Games and Economic Behavior* 82 (2013), 582–591.

[4] Martin W Cripps and Jeroen M Swinkels. 2006. Efficiency of large double auctions. *Econometrica* 74, 1 (2006), 47–92.

[5] Piotr Dworczak, Scott Duke Kominers, and Mohammad Akbarpour. 2021. Redistribution through markets. *Econometrica* 89, 4 (2021), 1665–1698.

[6] Haobo Fu, Weiming Liu, Shuang Wu, Yijia Wang, Tao Yang, Kai Li, Junliang Xing, Bin Li, Bo Ma, QIANG FU, and Yang Wei. 2022. Actor-Critic Policy Optimization in a Large-Scale Imperfect-Information Game. In *International Conference on Learning Representations*. https://openreview.net/forum?id=DTXZqTNV5nW

[7] Scott Fujimoto, Herke Hoof, and David Meger. 2018. Addressing function approximation error in actor-critic methods. (2018), 1587–1596.

[8] Raphael Koster, Jan Balaguer, Andrea Tacchetti, Ari Weinstein, Tina Zhu, Oliver Hauser, Duncan Williams, Lucy Campbell-Gillingham, Phoebe Thacker, Matthew Botvinick, and Christopher Summerfield. 2022. Human-centred mechanism design with Democratic AI. *Nature Human Behaviour* 6, 10 (01 Oct 2022), 1398–1407. https://doi.org/10.1038/s41562-022-01383-x

[9] Marc Lanctot, Vinicius Zambaldi, Audrunas Gruslys, Angeliki Lazaridou, Karl Tuyls, Julien Pérolat, David Silver, and Thore Graepel. 2017. A Unified Game-Theoretic Approach to Multiagent Reinforcement Learning. *arXiv* 1711.00832 (2017). https://doi.org/10.48550/ARXIV.2207.00898

[10] Siqi Liu, Luke Marris, Daniel Hennes, Josh Merel, Nicolas Heess, and Thore Graepel. 2022. NeuPL: Neural Population Learning. *arXiv preprint arXiv:2202.07415* (2022).

[11] John R Moroney and CAK Lovell. 1997. The relative efficiencies of market and planned economies. *Southern economic journal* (1997), 1084–1093.

[12] Paul Muller, Shayegan Omidshafiei, Mark Rowland, Karl Tuyls, Julien Pérolat, Siqi Liu, Daniel Hennes, Luke Marris, Marc Lanctot, Edward Hughes, Zhe Wang, Guy Lever, Nicolas Heess, Thore Graepel, and Rémi Munos. 2019. A Generalized Training Approach for Multiagent Learning. *CoRR* abs/1909.12823 (2019). arXiv:1909.12823 http://arxiv.org/abs/1909.12823

[13] Julien Perolat, Bart de Vylder, Daniel Hennes, Eugene Tarassov, Florian Strub, Vincent de Boer, Paul Muller, Jerome T Connor, Neil Burch, Thomas Anthony, et al. 2022. Mastering the Game of Stratego with Model-Free Multiagent Reinforcement Learning. *arXiv e-prints* (2022), arXiv–2206.

[14] J. Černý, A. Jedličková, M. Loebl, and D. Sychrovský. 2022. Critical Distribution System. *arXiv* 2207.00898 (2022). https://doi.org/10.48550/ARXIV.2207.00898

[15] Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, Junhyuk Oh, Dan Horgan, Manuel Kroiss, Ivo Danihelka, Aja Huang, L. Sifre, Trevor Cai, John P. Agapiou, Max Jaderberg, Alexander Sasha Vezhnevets, Rémi Leblond, Tobias Pohlen, Valentin Dalibard, David Budden, Yury Sulsky, James Molloy, Tom Le Paine, Caglar Gulcehre, Ziyun Wang, Tobias Pfaff, Yuhuai Wu, Roman Ring, Dani Yogatama, Dario Wünsch, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy P. Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps, and David Silver. 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* (2019), 1–5.