Price of Anarchy in a Double-Sided Critical Distribution System

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

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

Jakub Černý
Nanyang Technological University
Singapore
cerny@disroot.org

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
auction, fairness, price of anarchy

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 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.

In this context, we have in mind not just deciding on the price, but also storage, delivery, etc.
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 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 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.

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. \( \pi \) refers to the strategies and \( T \) checks the termination condition.

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. 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 imperfect-information

---

\(^1\)Our experimental results show that this fairness mechanism performs well in practice.

\(^2\)Our experimental results show that this fairness mechanism performs well in practice.

\(^3\)The results easily generalize to periodic trading of larger quantities of indivisible Good; see [14] for a more detailed explanation.

\(^4\)Why this market game and its strategies are parametric in \((M, g)\) will become apparent later during the definition of the sequence of markets constituting the crisis.
sequential game $G(M, G) = (T, M, G, D, \phi, \mu, \nu, k)$. We use calligraphic
to denote ordered sets and non-calligraphic
denote individual elements of these sets. The set of traders $T$
consists of buyers $B$ and sellers $S$; the sets of buyers and sellers
are assumed to be disjoint. The set $M = (M_1, M_2, \ldots, M_{|T|}) \in \mathbb{R}_0^{+ \times |T|}$
determines the real and non-negative amount of money each trader
receives at the beginning of the Market. Similarly, the set $G =
(G_1, G_2, \ldots, G_{|T|}) \in \mathbb{R}_0^{+ \times |T|}$ specifies the real and non-negative
amount of Good each trader obtains.

The demands $D = (D_1, D_2, \ldots, D_{|B|}) \in \mathbb{R}_0^{+ \times |B|}$ describe the
ideal real and non-negative amount of Good each buyer hopes
to acquire during the trading. Function $\phi$ then implements the
fairness mechanism, assigning real-valued, non-negative Rights
$R = (R_1, R_2, \ldots, R_{|B|}) \in \mathbb{R}_0^{+ \times |B|}$ 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 $s \in S$, the set $\Pi_s$ contains the
seller’s all possible parametric strategies, represented as functions
$\pi_s : \mathbb{R}_0^{+ \times |B|} \times \mathbb{R}_0^{+ \times |B|} \times \mathbb{R}_0^{+} \rightarrow \mathbb{R}_0^{+ \times 2}$,
and interpreted as $\pi_s(M, G, G_s) = (\phi^G, p^G_s)$, where $\phi^G \leq G_s$ is
the amount of Good offered at price $p^G_s$. Each $\pi_s$ 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 gives rise to sellers investing
in some market research. A profile of one strategy per seller is
denoted as $\pi_s \in \Pi_s$. The traders’ behavior in a specific game
$G(M, G)$ is then written as $\pi_s(M, G)$. We also refer to their offers
at once as $\phi^G_s$ and $p^G_s$. The fairness mechanism is defined as:

**Definition 1.** For any sellers’ strategy profile $\pi_s$, the fairness
mechanism is a function $\phi : \mathbb{R}_0^{+ \times |B|} \times \mathbb{R}_0^{+ \times |B|} \times \mathbb{R}_0^{+} \rightarrow \mathbb{R}_0^{+ \times 2}$ allocating Rights
to each buyers, satisfying in each game $G(M, G)$

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

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

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

where $\alpha$ is a permutation of $B$ and $V(\pi_s, M, G) = \sum_{s \in S} \phi^G_s \pi_s(M, G) \times p^G_s$.

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 $R = \phi(V(\pi_s, M, G), D)$.

Function $\mu$ 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 ($\phi^G_s, p^G_s$) they are offering for sale. Second, after observing
the offers of other buyers, each buyer declares the amount and price
of Good and Right ($\phi^G_s, p^G_s$) 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 $\mu$ mechanisms for each. A strategy of each seller can then
be formulated in both stages at once as a function

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

written as $\pi_b(v^G_s M_s, G_s, G_b, \psi B_b, \psi B_b, \psi B_b, \psi B_b, \psi B_b, \psi B_b, \psi B_b)$.

For each buyer $b \in B$, their set of strategies is $\Pi_b$. A profile of one
strategy per buyer is denoted as $\pi_b \in \Pi_b$. We write the Cartesian
product $\Pi_S \times \Pi_B$ as $\Pi$ 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 $\pi$ in a game $G(M, G)$, the
market mechanism is a function $\mu : \Pi \times \mathbb{R}_0^{+ \times |T|} \times \mathbb{R}_0^{+ \times |B|} \times \mathbb{R}_0^{+ \times |B|} \times \mathbb{R}_0^{+ \times |B|} \rightarrow \mathbb{R}_0^{+ \times |T|} \times \mathbb{R}_0^{+ \times |B|} \times \mathbb{R}_0^{+ \times |B|}$, 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 $\mu$ 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 $\mu_1^G(\cdot)$ and $\mu_2^M(\cdot)$ 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 $u = (u_t)_{t \in T}$, $u_t : \Pi \times \mathbb{R}_0^{+ \times |T|} \times \mathbb{R}_0^{+ \times |B|} \rightarrow \mathbb{R}$ 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 crises6.
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 $s \in S$,

$$u_s(\pi, M, G) = \begin{cases} 
\mu_1^M(\pi, M, R, G) + C_1 \mu_2^L(\pi, M, R, G) & \text{NT,} \\
\mu_2^M + C_2 \mu_2^L(\pi, M, R, G) & T, \end{cases}$$

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

---

6Without such penalty and if the Good is not perishable and the distribution crisis continues for a longer time, the strategic behavior of sellers would probably be to keep selling small amounts of the Good for very high prices.
unit of Money they have at the end the Crisis. Formally, for \( b \in B \),
\[
    u_b(\pi, M, G) = \begin{cases} 
        \min \{ D_b, \mu^G(\pi, M, R, G) \} & \text{NT}, \\
        u_b - NT + C_i \mu^M(\pi, M, R, G) & T.
    \end{cases}
\]
(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 seller is assigned Rights by the fairness mechanism \( \phi \).

**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 \( \mu \).
(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 \( C = (T, D, \phi, \mu, u, M^1, G^1, T, P) \). \( T, D, \phi, \mu, u \) 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 \( G(M^1, G^1) \). 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 \( G \) with some parameters \( (M, G) \). 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 \( P : \Pi \times R^{|T|} \times R^{|T|} \rightarrow R^{|T|} \times R^{|T|} \). Using the same notation as for the market mechanism \( \mu \), we write
\[
    \mu^G_t(\pi, M, G) = \mu^G_t(\pi, M, R, G) + G^1_t \quad \text{if } t \in S, \\
    0 \quad \text{if } t \in B, \text{ and}
\]
\[
    \mu^M_t(\pi, M, G) = \begin{cases} 
        \mu^M_t(\pi, M, R, G) + M^1_t & \text{if } t \in S, \\
        0 & \text{if } t \in B.
    \end{cases}
\]
\(^7\)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.

**Definition 3.** Let \( \pi \in \Pi \) induce a series of Markets \( (M^1, G^1), \ldots, (M^T, G^T) \). We call \( \pi \) an equilibrium if for any unilaterally deviating profile \( \hat{\pi} \in \Pi \) and the corresponding series of Markets \( (\hat{M}^1, \hat{G}^1), \ldots, (\hat{M}^T, \hat{G}^T) \) it holds for all traders \( t \) that
\[
    \sum_{t=1}^T y^t u_t(\pi, M^t, G^t) \geq \sum_{t=1}^T y^t u_t(\hat{\pi}, \hat{M}^t, \hat{G}^t).
\]

### 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 the 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 \( (M, G) \) under profile \( \pi \in \Pi \) and allocated rights \( R = \phi(V(\pi, S, M, G), D) \) is then
\[
    f_b(\pi, M, G) = \max \left\{ R_b - \mu^G_b(\pi, M, R, G), 0 \right\}.
\]

Our Price of Anarchy in the system is then the normalized accumulated frustration the buyer experiences in the sequence of \( \tau = T \) Markets when the equilibrium \( \pi \) is reached and induces a series of Markets \( (M^1, G^1), \ldots, (M^T, G^T) \), i.e.,
\[
    \text{PoA} = \frac{\sum_{t=1}^T \sum_{b \in B} f_b(\pi, M^t, G^t)}{\tau |B|}. \quad \text{(PoA)}
\]

### 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: \( G_C = (B, S, E_C) \) which represents the compatibility for trading the Good and \( G_R = (B_S, B_R, E_R) \) which represents the compatibility for trading the Right. Here, \( B_S \) and \( B_R \) are disjoint copies of \( B \). \( B_S \) represents the sellers of Right and \( B_R \) represents the buyers of Right. A trader of \( B \) can be both a seller and a buyer of Right, but \( G_R \) does not connect their representing vertices by an edge. Both \( G_C, G_R \) are equipped with a positive real weight \( w_C : V_C \rightarrow \mathbb{R} \) and \( w_R : V_R \rightarrow \mathbb{R} \). 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., \( p_b^G > p_b^G \) and/or \( p_b^R > p_b^R \). 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 assigned 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 \( p_b^G \) in descending order. The mechanism has 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 flow networks. 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 trades 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 Maximum clearing allocation is polynomial for both divisible and indivisible Good.

**Proof.** Given the disjoint copies of the graphs \( G_G, G_R \), we construct an instance of the Max Flow problem as follows:

1. introduce two new vertices \( s, t \);
2. join by an arc \((s, v)\) to each vertex \( v \) of \( G_G \). Let the capacity \( \text{cap}(s, v) \) of this arc be equal to the amount of the remaining rights \( v \), i.e., the assigned amount minus the amount intended to be sold. Clearly, each buyer desires to buy at least \( \text{cap}(s, v) \) of Good.
3. join by an arc \((s, v)\) to each vertex \( v \) of \( G_R \). Let the capacity of this arc be equal to \( w_R(v) \), i.e., the amount (possibly zero) of Right \( v \) intends to sell;
4. orient each edge of \( G_R \) towards \( B_R \), the capacity of \((x, y)\) being equal to \( w_R(y) \), i.e., the amount of Right \( y \) intends to buy;
5. orient each edge of \( G_G \) towards \( S \), the capacity of \((x, y)\) being equal to \( w_G(y) \), i.e., the amount of Good \( y \) intends to sell;
6. introduce a copy \( B' \) of \( B \) and join each vertex \( v \) of \( B \) by an arc \((v, v')\) to its copy \( v' \) of \( B' \), its capacity being \( w_B(v') \), i.e., the amount of Right \( y \) intends to buy;
7. join \((v, v')\) to \( S \) in the same way as its copy \( v \) is joined to \( S \) in \( G_G \); orient these new edges towards \( S \) and let the capacity of each such arc terminating in \( y \) in \( S \) be \( w_G(y) \), i.e., the amount of Good \( y \) intends to sell;
8. join each vertex \( y \) of \( S \) to \( t \) by the arc \((y, t)\), its capacity being \( w_G(y) \), i.e., the amount of Good \( y \) intends to sell.

This finishes the construction of the instance of the Max Flow problem. It is straightforward to see that max flow from \( s \) to \( t \) provides a clearing of bids with the maximum amount of the Good sold. Also, it is ensured that Right is bought along with the same amount of Goods. \( \Box \)

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.** We can find the maximum clearing allocation using the following linear program, where the variable \( r_{b, b'} \) represents the amount of Right sold to \( b' \) by \( b \), with \((b, b') \in E_G \) and the variable \( g_{s, b} \) represents the amount of Good sold to \( b \) by \( s \), with \((s, b) \in E_G \). We also introduce the variables \( m \) and \( M \), representing the minimal, resp maximal, amount of goods bought by a buyer. Furthermore, we define \( c = \epsilon + U \) where \( \epsilon \) is the desired sensibility of the objective function and \( U \) an upper bound on \((M - m)\):

\[
U = \max_{b \in B} \left\{ \min_{(s, b) \in E_G} \sum_{(s, b) \in E_G} w_G(s) \right\}; \quad c \text{ will be used to normalize (M - m) in order to not interfere with the rest of the objective}
\]
function. In our experiments, we used $c = \frac{1}{1000}$.

$$\max \sum_{s,b} g_{s,b} - c(M - m)$$

s.t.

$$\sum_{s,b} g_{s,b} \leq r_b + \sum_{(b',b) \in E_R} r_{b',b} - \delta^G_b \quad \forall b \in B$$

$$\sum_{s,b} g_{s,b} \leq \sigma^G_b \quad \forall b \in B$$

$$\sum_{s,b} r_{b,b'} \leq \delta^G_b \quad \forall b \in B$$

$$\sum_{s,b} r_{b,b'} \leq \delta^G_b \quad \forall b \in B$$

$$m \leq \sum_{s,b} g_{s,b} \quad \forall b \in B$$

$$M \geq \sum_{s,b} g_{s,b} \quad \forall b \in B$$

$$\sum_{s,b} g_{s,b} \geq 0 \quad \forall (s,b) \in E_G$$

$$r_{b,b'} \geq 0 \quad \forall (t,b) \in E_R$$

$$\forall (s,b) \in E_G$$

$$\forall (b',b) \in E_R$$

In this linear program, the objective function (1) maximizes the exchanges of goods, and spreads the distribution over the buyers. The constraints (2) and (3) then enforce that the buyers buy less good than they have rights, and the amount of good they buy does not exceed their demand $\delta^G_b$. The constraint (4) imposes a restriction on the amount of good the sellers may sell, ensuring it is at most $\sigma^G_b$, i.e., the amount they committed themselves to be willing to sell. Similarly, the constraint (5) imposes that the buyers selling good sell at most the amount they intend to sell $\delta^G_b$. The constraint (6), resp (7), assures that $m$ is lower, resp. higher, than the minimal, resp. maximal, amount of good bought by a buyer, and the sense of the objective function ensure that it will be exactly this quantity. The constraint (8) imposes that the buyers pay at most in average $p^G_b$ for the goods. The constraint (9) forces that the buyers pay at most in average $p^G_b$ for the rights. The constraint (10) is the budget constraint.

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 $\pi$. 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 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.

### Algorithm 1 Equilibrium Learning Algorithm

1. $B \leftarrow \text{set of buyers, } S \leftarrow \text{set of sellers, } \emptyset \leftarrow \{\}$
2. for episode $e \in \{1, ..., N_{\text{sims}}\}$ do
3. for $t \in \{1, ..., T\}$ do
4. $G_t \leftarrow G_t + G_t^1, M_t \leftarrow M_t + M_t^1$
5. $i_S \leftarrow \text{observation of sellers}$
6. $i_B \leftarrow \text{observation of buyers}$
7. Allocate Right according to $\phi$
8. Sample batch $\sim \emptyset$
9. $G_t \leftarrow \text{clip}(\Pi_B(i_B(b), i_S, \emptyset), 0, 1)$
10. $G_t \leftarrow \text{clip}(\Pi_B(i_B(b), i_S, G_t, \emptyset_B), 0, 1)$
11. Trade according to a market mechanism $\mu$
12. Compute utilities $u_B, u_S$
13. $B \leftarrow B \cup \{i_B, i_S, \Pi_B, G_t, \emptyset_B, u_B, u_S\}$
14. $G_t \leftarrow \text{clip}(G_t - D_G, 0)$
15. if $r \text{ mod } N_{\text{train}}$ is zero then
16. Sample batch $\sim B$
17. Train on batch using TD3
18. end if
19. end for
20. Reset episode
21. end for

4.1 Utilities in the learning environment

Next, we focus on the traders’ utilities. To identify them, we need to specify constants $C_t$. For simplicity and to reduce the action space, we assume there exists a maximum price $\bar{P}$ 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, $C_t$, may be chosen arbitrarily, however, it needs to be sufficiently high. If $|C_t| > C_2$, it becomes beneficial for the sellers to keep the Good, and the selling price would thus be $\bar{P}$.

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 $\bar{P}$. $C_3$ should hence inversely depend on $\bar{P}$.

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 $\Pi_t$ 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 \( R \). 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 approximation 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, mlp 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 Appendix A.

**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 than 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 \( D_b \) is 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 \( \mathbb{E}_{b \sim \mathcal{G}}[D_b] = 1 \) and \( \mathbb{E}_{b \sim \mathcal{G}}[M_b^1] = 1/8 \). To fix a scale, we set the maximum price as \( \bar{P} = 1 \). The constants in the utilities are then chosen as \( C_1 = -1/8, C_2 = 1/2, C_3 = 1/\bar{P} = 1 \) and \( G_1^i = 1/|S| \).

5.1 Exploitability

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

\[
\sum_{\tau=1}^{T} \sum_{t=1}^{T} u_t(\pi^\tau \cup \overline{BR}_t(\pi^\tau), \overline{M_1^t}, \overline{G^t}) - u_t(\pi^\tau, M_1^t, G^t).
\]

Here, we denote by \( \overline{BR}_t(\cdot) \) an approximate best response of trader \( t \). The complement strategies are then written as \( \pi_{-t}^\tau \). The corresponding sequences of Markets for \( \pi^\tau \) and \( \pi_{-t}^\tau \cup \overline{BR}_t(\pi^\tau) \) are denoted as \( (M_1^t, G^t) \) and \( (\overline{M_1^t}, \overline{G^t}) \), 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 \( \pi^\tau \). 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 earnings 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 thus do not compute the approximate best response during the entire simulation; instead, 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

Finally, 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
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.

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.

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].
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.

ACKNOWLEDGMENTS
This research was supported by the CRISDIS project of the Czech Ministry of the Interior no.VI04000107 and SIMTech-NTU Joint Laboratory on Complex Systems. Computational resources were supplied by the project "e-Infrastruktura CZ" (e-INFRA CZ LM2018140) supported by the Ministry of Education, Youth and Sports of the Czech Republic. The authors would also like to express their gratitude towards an anonymous reviewer who suggested to investigate the effects of the number of traders on the Price of Anarchy.

REFERENCES
A HYPERPARAMETERS
The experiments used the following values of parameters:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actor learning rate</td>
<td>3 \times 10^{-4}</td>
</tr>
<tr>
<td>Critic learning rate</td>
<td>10^{-3}</td>
</tr>
<tr>
<td>Actor hidden layer size</td>
<td>32</td>
</tr>
<tr>
<td>Critic hidden layer size</td>
<td>256</td>
</tr>
<tr>
<td>Batch size</td>
<td>512</td>
</tr>
<tr>
<td>L2 penalty</td>
<td>10^{-2}</td>
</tr>
<tr>
<td>Discount factor ( \gamma )</td>
<td>0.99</td>
</tr>
<tr>
<td>Target network update rate</td>
<td>0.002</td>
</tr>
<tr>
<td>Actor training frequency</td>
<td>3</td>
</tr>
<tr>
<td>Entropy penalty</td>
<td>3 \times 10^{-3}</td>
</tr>
<tr>
<td>Training episodes</td>
<td>3000</td>
</tr>
<tr>
<td>NashConv training episodes</td>
<td>100</td>
</tr>
</tbody>
</table>

B EXAMPLE OF A SIMPLE REDISTRIBUTION SYSTEM
Consider an example of a redistribution system with two buyers \( \{b, b'\} \) and a single seller \( s \). Let the buyer \( b \) have more funds and the other buyer \( b' \) have a larger demand, formalized as \( (M^G_{R^G_b}, M^G_{R^G_{b'}}) = (2, 1) \) and \( (D^G_{R^G_b}, D^G_{R^G_{b'}}) = (1, 3) \). In addition, let the seller have two units of Good for sale.

Assume the seller offers their entire stock \( (\omega^G = 2) \) for sale at price \( p^{G} = 1 \), resulting in a distribution of Rights \( (R^G_{b}, R^G_{b'}) = (0.5, 1.5) \) as dictated by the contested garment rule. After releasing the Rights to the buyers, the second buyer \( b' \) realizes they may directly buy a unit of Good with the unit of Money they have and puts the remaining half a unit of Right for sale. Assume they choose a price \( p^{G}_{b'} = 1 \). The other buyer \( b \) has no intention to offer their Right for sale as they prefer to use it for buying the Good and \( b' \) is unlikely to buy the Right anyway. Let the acceptable prices and volumes be declared as

\[
(\omega^{R}_{b'}, R^{R}_{b'}, \omega^{G}_{b'}, R^{G}_{b'}, \omega^{G}_{b}, R^{G}_{b}) = (0, 0, 0.5, 1, 1, 1),
\]

\[
(\omega^{R}_{b}, R^{R}_{b}, \omega^{G}_{b}, R^{G}_{b}, \omega^{G}_{b'}, R^{G}_{b'}) = (0.5, 1, 0, 1, 2).
\]

Next, the bids are cleared using a market mechanism. In this example we will show the greedy allocation defined in Subsection 3.2 because it is the most transparent market mechanism we consider in the paper.

- We start with buyer \( b' \) since \( p^{G}_{b'} > p^{G}_{b} \):
  1. Because the acceptable price \( p^{G}_{b} \) is greater than \( p^{G}_{b'} \), in the first stage the volume of min \( \min \left\{ \omega^{G}, \omega^{G}_{b'}, R^{G}_{b'} - \omega^{G}_{b}, \frac{M^{G}_{b'}}{\omega^{R}_{b'}} \right\} = 1 \) is sold to buyer \( b' \).
  2. In the second stage \( b' \) buys nothing since their desired volume and desired price of Right are zero and they already bought \( \omega^{G}_{b} \) of Good.
- We proceed with buyer \( b \):
  1. Since \( p^{G}_{b} \geq p^{G}_{b'} \), buyer \( b \) uses the Right they were allocated to buy min \( \min \left\{ \omega^{G}_{b} - 1, \omega^{G}_{b} - \omega^{G}_{b'}, \frac{M^{G}_{b}}{\omega^{R}_{b}} \right\} = 0.5 \) units of Good in the first stage.

(2) Finally, in their second stage \( b \) buys the Right and the Good in equal quantities from \( b' \) and \( s \), respectively, buying

\[
\min \left\{ \omega^{G} - 1.5, \omega^{G}_{b'} - 0.5, \frac{M^{G}_{b}}{\omega^{R}_{b}} \right\} = 0.5
\]

units of Good. The Money \( b \) pays is split between \( s \) and \( b' \) proportionally to \( p^{G} \) and \( p^{G}_{b'} \), which in this case is half and half.

After the market mechanism reallocates the resources in this manner, the buyers end up with \( (1, 1) \) units of Good and \( (0.5, 0.5) \) units of Money. The seller also sells all the Good they had for trade.

The redistribution system ends the trading iteration by using the transition function to subtract the buyers’ demands from their Good, simulating the consumption. The buyers hence have no Good left and they both receive utility one. The transition function also simulates the buyers’ external earnings and sellers’ resupply, providing them with extra Money and Good for the following Market. The model assumes the amounts of these periodic additional resources are held constant through the entire Crisis and are equal to the Money and Good the traders had at the beginning of the first Market. To sum up, this means that at the beginning of the second iteration of the redistribution system the buyers have \( (2.5, 1.5) \) units of Money and the seller has two units of Good.

In contrast to free market, the redistribution system with Rights hence guarantees buyer \( b' \) the extra 0.5 units of Money at the start of the second Market. This allows them to buy more Good in the future and decrease their accumulating frustration. How much the frustration decreases depends not only on the strategies of the traders but also on the market mechanism. If we employ the maximum clearing mechanisms, the result does not change in this simple case since there is no other way to clear the bids.

C PRICE OF ANARCHY AS A FUNCTION OF THE NUMBER OF BUYERS AND SELLERS
We consider the same scenario as described in Subsection 5.2 where we assessed scalability of our learning algorithm. We generate random instance of Crises with \( N \) buyers and \( N \) sellers for each \( N \in \{2, \ldots, 19\} \) and with the demands and earning same as explained at the beginning of Section 5. Again, each buyer has a chance to be rich with probability 3/4 and poor otherwise.

The results are depicted in Figure 6. For each \( N \) we generate 10 instances and the graphs show also the standard errors. The same general trend we reported in the main text 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 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
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.

\( N = 6 \) consistently decreases, reaching a PoA value \( \approx 5.93 \)-times smaller than in the free market.

Overall, the results clearly further substantiate our conclusions from the main text that the clearing mechanisms prioritizing the amount of goods sold are superior to the two alternatives we considered.