A Walrasian Mechanism with Markups for Nonconvex Economies

Paul Milgrom* and Mitchell Watt†

September 2024

Abstract

We introduce the Walrasian Markup equilbrium, an extension of Walrasian equilibrium that adds a markup to the prices that consumers pay to ensure existence even in nonconvex quasilinear economies. Markup equilibria are resource-feasible, incur no budget deficit, and require little more communication and computation than the Walrasian equilibrium. The Markup direct mechanism is asymptotically dominant-strategy incentive-compatible. Our Bound-Form First Welfare Theorem states that for any price vector, the welfare loss of any feasible allocation compared to the first-best is at most the sum of (i) the budget surplus and (ii) any rationing losses suffered by the participants. This implies that any Markup equilibrium with a small markup is nearly efficient.

Keywords: Approximate efficiency, Approximate incentive-compatibility, Market design, Nonconvexity, Prices, Rationing

JEL Codes: C620, D400, D440, D450, D470, D500, D510, D610.

*Stanford University and Auctionomics. Email: milgrom@stanford.edu

†Stanford University. Email: mwatt@stanford.edu.

Thank you to Mohammad Akbarpour, Martin Bichler, Robert Day, Ravi Jagadeesan, Fuhito Kojima, Shoshana Vasserman, and seminar participants at Stanford University, the University of Zurich, the NBER Market Design Working Group, the Simons Laufer Mathematical Sciences Institute and the 32nd Stony Brook International Conference on Game Theory, and the referees for helpful comments and suggestions relating to this project. An extended abstract of this paper appeared in *Proceedings of the 23rd ACM Conference on Economics and Computation (EC'22), July 11-15, 2022, Boulder, CO, USA* as "Linear Pricing Mechanisms for Markets without Convexity." An earlier draft of this paper was circulated as "Walrasian Mechanisms for Non-convex Economies and the Bound-Form First Welfare Theorem." Milgrom gratefully acknowledges support from the National Science Foundation (grant SES-1947514). Watt gratefully acknowledges the support of the Koret Fellowship, the Ric Weiland Graduate Fellowship and the Gale and Steve Kohlhagen Fellowship in Economics at Stanford University.

1 Introduction

Walrasian equilibrium has long been a standard benchmark for economic outcomes because of its many desirable properties: its allocations are feasible and efficient for producers and consumers, its balanced budget requires no subsidies from any third party, and it has the fairness property of offering the same prices to all participants. From an implementation perspective, it has two additional good properties. The first, emphasized by Hayek (1945), is that prices economize on communication: when a Walrasian equilibrium exists, its prices are the minimal information a planner would need to communicate to allow each agent to check whether its proposed allocation is part of an efficient plan.[1] The second concerns incentives: The Walrasian mechanism provides little opportunity for small market participants to profitably manipulate the plan.[2]

Mechanisms that aim to approximate Walrasian outcomes have sometimes been used in practice for multi-product markets, such as electric power, in which products are distinguished by location and time of day, and fishing rights, in which products are distinguished by location and species. However, any attempt to implement the Walrasian mechanism must overcome multiple challenges.

The first challenge is that the just-cited applications often entail fixed costs of production and other non-convexities, with the consequence that in some cases, Walrasian equilibrium may *fail to exist.* A mechanism allowing agents to report fixed costs needs a different procedure to determine prices and allocations in those cases. In this paper, we extend the Walrasian mechanism by including an extra pricing parameter. A *Walrasian mechanism with markups* or just *markup equilibrium* is a triple $(\mathbf{x}, p, \alpha)$ in which $\mathbf{x}$ is a feasible allocation for buyers and sellers, p is a price vector that determines payments to sellers and α is a markup, with buyers paying according to the price vector $(1 + \alpha)p$. For $(\mathbf{x}, p, \alpha)$ to be a markup equilibrium, all buyers and sellers must be assigned their most preferred bundles at the prices they face, and the payments received from buyers must weakly exceed the payments made to sellers. A *minimal markup equilibrium* is a markup equilibrium with the smallest value of α. If there are finite, nonzero choke prices for both supply and demand, then a minimal markup equilibrium always exists. Since $(\mathbf{x}, p)$ is a Walrasian equilibrium if and only if $(\mathbf{x}, p, 0)$ is a markup equilibrium, the minimal markup equilibria coincide with the Walrasian equilibria whenever the latter exist.

[1] See also Nisan and Segal (2006) and Segal (2007), who prove that any decentralized communication system that implements efficient allocations must communicate a vector of supporting prices to the agents.

[2] For incentive analysis of Walrasian equilibrium at the limiting case of a continuum of consumers, see Roberts and Postlewaite (1976) and Jackson (1992), while for analysis of the rates of convergence of incentives in finite markets, see Azevedo and Budish (2019) and Watt (2022).

Next are challenges related to *feasibility and efficiency*. A Walrasian allocation balances supply and demand, and its payments to sellers are equal to the payments from buyers, so it is always resource-feasible and budget-feasible, but notions of approximate Walrasian equilibrium including *pseudo-equilibrium* and *quasi-equilibrium* found in Starr (1969) can specify plans that are not fully feasible. In markup equilibrium, the allocations for each firm and consumer are individually feasible, the total production of each good weakly exceeds its consumption, and the total revenue collected from buyers weakly exceeds the payments made to sellers, so the markup outcome is fully feasible. Moreover, just as Walrasian allocations are efficient; markup allocations with small markups are nearly efficient.

A third set of challenges, emphasized by the mechanism design perspective, concerns *incentives* to participate and report truthfully. If producers have fixed costs, then the marginal cost prices of some Walrasian-like mechanisms can be insufficient to cover some producers' full costs, violating their participation constraints. Some Walrasian-like mechanisms seek to cover producers' full costs by adding so-called "uplift" payments, which can incentivize producers to exaggerate their fixed costs. In a markup mechanism, there are no uplift payments: the linear prices are sufficient to eliminate both the incentive and participation problems. In a markup equilibrium, production may strictly exceed consumption, but there is no budget deficit because the markup in consumer prices pays the cost of any excess production.

Fourth are challenges related to the *communications and computations*. For convex economies, if the planner announces the proposed Walrasian allocation and prices, then each participant can verify that its allocation is part of an efficient plan. For non-convex economies, if the planner announces the allocation, prices and one parameter more – the markup, then each participant can verify that its allocation is part of an approximately efficient plan. For the quasilinear cases we study in this paper, the Walrasian allocation and prices for convex economies can be computed by solving a convex optimization problem and its dual. The markup equilibrium allocation, prices and markup for nonconvex economies can be computed by solving a series of convex optimizations and their duals, as described below.

When the economist's task is to reform an existing market, another concern is to limit *disruption for current participants*. If a Walrasian-like mechanism is already being used in a nonconvex, multi-product marketplace, then limiting disruption may call for implementing another Walrasian-like mechanism, rather than, for example, adopting a Vickrey-Clark-Groves *pivot mechanism*. Changing to a pivot mechanism can be disruptive, requiring participants to adapt their processes and affecting the values of their past capital investments. The pivot mechanism may also be unacceptable because of its other serious disadvantages: it cannot guarantee that revenues weakly exceed costs, can pay higher compensation to firms that produce less output,[3] and can require impractical levels of communication and computation.[4]

In addition to introducing markup equilibrium, this paper also introduces an extension of the First Welfare Theorem for quasi-linear economies. We call this the *Bound-Form First Welfare Theorem* because it gives an upper bound on the welfare loss of any feasible allocation $\mathbf{x}$ using any price vector p, and delivers a bound of *zero* when the pair $(\mathbf{x}, p)$ is a Walrasian equilibrium.

To describe the theorem, we first need to define some terms. The *welfare* of any feasible allocation $\mathbf{x}$ is defined as the sum of the values enjoyed by consumers minus the sum of the costs incurred by firms; the *welfare loss* of $\mathbf{x}$ is defined to be the welfare of the first-best allocation minus the welfare of $\mathbf{x}$. Given an allocation and price vector $(\mathbf{x}, p)$, the *rationing loss* of firm f is the difference between its maximum profit at prices p and the profit from its assigned allocation x_f; the rationing loss for consumers is defined similarly. According to the theorem, for any price vector p and any feasible allocation $\mathbf{x}$, the welfare loss of $\mathbf{x}$ is bounded above by the sum of two terms: (1) the value using prices p of any excess of production over consumption plus (2) the sum of the rationing losses suffered by consumers and firms using prices p. For any Walrasian equilibrium $(\mathbf{x}, p)$, both terms are zero.

Given a markup equilibrium $(\mathbf{x}, p, \alpha)$, we can apply the Bound-Form First Welfare theorem to bound the welfare loss of the allocation $\mathbf{x}$ using price vector p by an amount proportional to the markup α. To see this, observe first that because firms produce their most preferred bundles at p, *producer rationing losses are exactly zero.* The rationing losses for consumers can be positive because, although each consumer n's bundle x_n is her most preferred one at the prices $(1 + \alpha)p$, the consumer may prefer a different bundle at

[3] For example, suppose firm A can produce up to two units at a cost of \$1 per unit and additional units at a cost of \$10 each, and firm B can produce any number of units at a cost of \$2 each. If the demand specification makes it efficient to produce exactly three units in total, then according to the pivot mechanism, firm A should produce 2 units and be paid \$4, and firm B should produce 1 unit and be paid \$10: the firm that produces more output receives the smaller payment.

[4] The VCG pivot mechanism can require very large amounts of computation time and resources, especially when applied to nonconvex markets. There are several reasons for this. First is the number of optimizations that the pivot mechanism can require: one possibly nonconvex optimization to determine the allocation and a second similarly difficult optimization *for each bidder* to compute its payment. Second, if participants are to use pivot prices to guide their investment decisions, then the mechanism needs to compute and communicate the pivot price for every participant and every potentially relevant bundle, multiplying the burdens of computation and communication. Third and perhaps most importantly, computing pivot mechanism prices with good accuracy requires that each of these optimizations achieve an unusual level of precision, because the pivot mechanism computes each participant's transfer as the *difference* between the maximum value when all participants are present and the maximum value when one participant is excluded.

To illustrate the impact of imprecise computations, suppose that there are 50 identical producers and that the optimization software initially finds the exact optimal allocation and the associated minimum cost. Further suppose that the exact minimum cost of production would be 1% higher if one of the 50 firms were excluded. In that case, the correct pivot price for any firm is about 3% of the total cost. If the additional cost minimizations were to overestimate the minimum total cost by 3% – which for many purposes would be good enough computational performance, then the calculated pivot price for each firm would be too high by 100%, doubling the planner's total payments to producers. Alternatively, if the computation error is 3% in the original allocation problem and zero in the extra minimization problems, then the computed pivot prices could be zero or a negative number. In contrast, the markup mechanism requires fewer and easier computations, as shown below.

prices p. However, by the envelope theorem (Milgrom and Segal, 2002), *each consumer's rationing loss is of a smaller order than* α. Finally, if consumer demand is strongly monotone, then *any excess production is proportional to* α. Adding up the three terms, the implied bound on the total welfare loss is proportional to α.

The incentives for truthful reporting in the markup mechanism are conceptually similar to those of the Walrasian mechanism in convex economies. In both mechanisms, participants expect to benefit from a false report only to the extent that their reports affect the prices used to compute their payments or receipts. We show that as the number of participants grows, with high probability, any one participant's effect on prices becomes vanishingly small.

To compute a markup equilibrium with a small α in a tractable way, we suggest an approach that begins with two changes to the standard Walrasian welfare maximization problem for convex economies: one to the constraints and one to the objective function. For constraints, our change requires that the total production of the firms must weakly exceed the total consumption *plus an operating reserve*, which is specified to depend on the largest nonconvexity but not on the numbers of producers and consumers. Second, for any given markup α, the objective to be maximized is $\frac{Utility}{1+\alpha} - Costs$. Both of these changes distinguish the markup program from the Walrasian one.

Next, if the specified objective is not concave, we replace it with its concavification to create a convex program. Solving the dual program yields the mark-up prices p; solving the primal problem yields an *approximate* mark-up allocation $\widehat{x}$. If the approximate allocation of some producer is not on its actual supply curve, then it is rounded to lie on the supply curve, giving us the markup allocation. In the markup allocation, supply always weakly exceeds demand because any supply reduction due to rounding can be replaced using the reserve. If the resulting plan is budget-feasible, meaning that total payments by consumers weakly exceed those to producers, then the triple $(\mathbf{x}, p, \alpha)$ is a markup equilibrium. We use a line search to find (approximately) the smallest α that is part of a budget-feasible plan, with each search candidate requiring a convex optimization and other steps as described above.

Following an earlier draft of this paper, Ahunbay, Bichler, Dobos, and Knörr (2024) performed a computational test to assess the potential of this computable markup mechanism for European wholesale spot electricity markets. That paper compared the markup equilibrium computations to those of a widely used mechanism, which optimizes allocations using mixed integer programming, computes prices using the dual of a relaxed version of the same optimization problem that omits the integer constraints, and pays additional "uplift" compensation to ensure that total payments cover producers' fixed costs. The paper found that (1) markup computations "are considerably faster for relevant problem sizes," (2) uplift compensation in the alternative mechanism results in substantial budget shortfalls, and (3) the markup allocation suffers only a small loss of welfare relative to the full optimum. This comparison assumes that both mechanisms have access to true reports, omitting the additional losses the alternative mechanism may suffer because of its incentives for false reporting.

The remainder of this paper is organized as follows. Section 1.1 contains a simple single product example to illustrate how markup equilibrium is computed and to highlight its properties, and Section 1.2 reviews the related literature. Section 2 introduces the quasilinear model and some preliminaries, including the measures of nonconvexity that we will use. Section 3 introduces the Bound-Form First Welfare Theorem. Section 4 introduces the markup equilibrium, including its computation, feasibility, incentive properties, and efficiency guarantee. Section 5 concludes.

1.1 A Single Product Example

We illustrate markup equilibrium with a single product example that features nonconvexity in production but not consumption.

On the supply side, each firm f can produce zero units at zero cost or any positive quantity up to its capacity K_f by incurring a fixed cost F_f. Its marginal cost of production is zero up to its capacity. If firm j produces at capacity, its average cost is $a_f = F_f/K_f$. Let $\overline{K} = \max_f K_f$ be the largest capacity among the firms. At any price p, firms will supply a total quantity of $S(p) = \sum_{\{f|a_f \le p\}} K_f$, which describes a discontinuous supply curve.

Total consumer values and demand in our example are described by a strictly concave value function $V(q)$ and an associated strictly downward-sloping demand function $D(p)$.

Walrasian equilibrium requires that there is some price p such that $D(p) = S(p)$. As illustrated in the example depicted in Figure 1, discontinuities in the supply function may imply that no Walrasian equilibrium exists. Indeed, considering a parameterized family of demand functions $b + D(p)$, there are intervals of the b parameter for which there is no solution to $b + D(p) = S(p)$, so there can be robust examples of demand functions for which Walrasian equilibrium does not exist.

Our extension of the Walrasian equilibrium is the markup equilibrium, which consists of a triple $(\mathbf{x}, p, \alpha)$, with $\mathbf{x}$ being the allocation, p the price per unit paid to producers, and $(1 + \alpha)p$ the price per unit paid by consumers. The price and markup are chosen to avoid rationing and budget deficits: in the example in

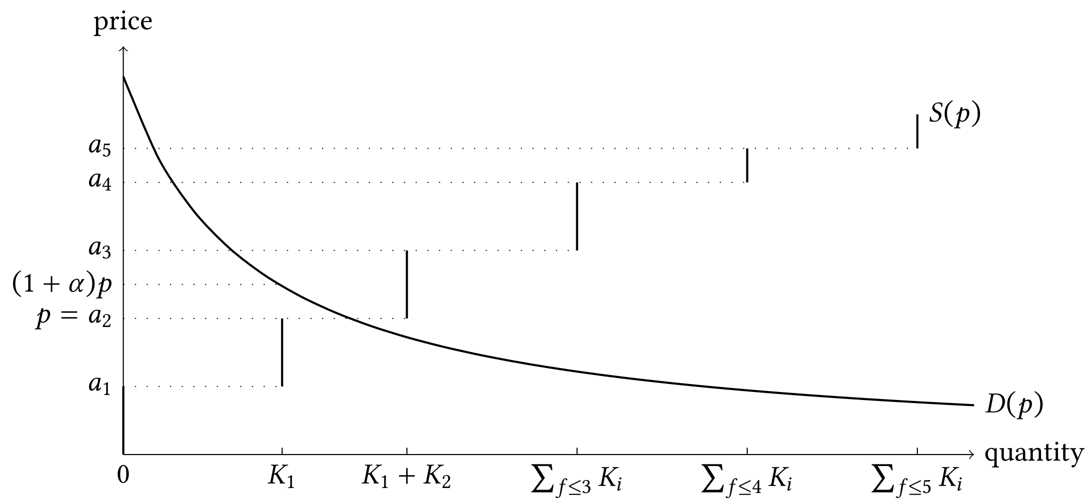

Figure 1: Five-firm example of Walrasian equilibrium nonexistence (here, firms are indexed in order of a_f). A markup equilibrium exists with total consumption and production equal to K_1, price $p = a_2$ paid to firms, and markup α on prices paid by consumers chosen so $D((1+\alpha)p) = K_1$.

Figure 1, one markup equilibrium involves production and consumption of the quantity K_1 with the price $p = a_2$ paid to producers and a markup α so that demand matches supply: $D((1+\alpha)p) = K_1$.

In this paper, we offer a computational approach to identify a particular markup equilibrium in large markets. The outer loop of the computation is a binary search for a suitable markup $\alpha \geq 0$.[5] In the inner loop, we adjust both the constraints and the objective of the usual welfare maximization problem. For the constraints, we specify that total production by firms must exceed the total allocation to consumers by at least $\overline{K}$. For the objective, on the side of the firms, we convexify each firm's cost function by setting $C_f(q) = a_f q$ for production up to capacity. It is convenient to let $\widehat{C}(q)$ denote the corresponding industry total cost and $\widehat{S}(p)$ the corresponding industry supply function. On the consumer side, we rescale the value function, setting $\widehat{V}(q) = V(q)/(1+\alpha)$. With these adjustments, the problem is to choose total industry production q to maximize $\widehat{V}(q - \overline{k}) - \widehat{C}(q)$. This is a convex optimization problem with its solution characterized by a first-order condition and a price p that equates supply and adjusted demand, as follows:

$$\widehat{S}(p) = D((1+\alpha)p) + \overline{K}$$

In the solution of this convex problem, each high-cost firm (ones with $a_f > p$) produces zero and each low-cost firm (ones with $a_f < p$) produces at capacity. There may also be one firm f' with $a_{f'} = p$ that produces a fraction of its capacity. For our computed markup equilibrium, the price is p, firm f'

[5] Since $\alpha \geq 0$, to ensure a compact search space, the binary search might search over $\beta := 1/(1+\alpha)$ which lies in $[0, 1]$.

produces zero, and all the other producers and consumers have the same allocation as was determined for the convexified problem. Because $K_{f'} \leq \overline{K}$, the total production of the low-cost firms is at least $D(1+\alpha)p$, that is, the markup plan is resource-feasible. To be a markup equilibrium, α must be chosen so that the plan is also be budget-feasible, which means that $p\widehat{S}(p) \leq (1+\alpha)D((1+\alpha)p)$.

In the example of this subsection, if there are many firms, then no single firm has much incentive to exaggerate its cost or understate its capacity in the markup mechanism, because such reports have only a limited effect on the firm's price. Moreover, misreporting is risky: if the firm reports a too-high fixed cost, its allocation will be zero. With suitable penalties for non-performance, firms are also deterred from overstating their capacities. At a markup equilibrium, there are no gains available from trade among consumers or among firms, because each group faces a single price. By the Envelope Theorem, there is little to be gained by adjusting the total output q. The welfare loss in markup equilibrium is mainly attributable to the unconsumed portion of production, which is some amount less than $\overline{k}$, independent of the number participants in the market. As a percentage of the trading volume, the total welfare loss decreases to zero as market participation grows.

Although this one-dimensional example is suggestive, it includes simplifying assumptions that need to be relaxed for the general theory. Nonconvexities may not always take the form of fixed costs, so we will need to work with more general measures of the nonconvexity of sets and determine the operational reserve accordingly. In the one-dimensional problem, rounding the solution just means rounding an output up or down, but that becomes subtler in higher dimensions. For example, a firm that can produce one unit of good 1 or good 2 could be allocated $(\frac{1}{2}, \frac{1}{2})$ in convexified optimization and to avoid rationing, the output of one good may need to be rounded up while the production of the other is rounded down.

1.2 Related literature

The problem of nonconvexity for the existence of competitive equilibrium was discussed in a series of papers by Farrell (1959), Rothenberg (1960), Koopmans (1961) and Bator (1961). Much of the subsequent classical literature on nonconvexity in general equilibrium theory focused on concepts of *approximate* equilibria which replace aggregate feasibility requirements with approximate feasibility, measured in terms of distance in the commodity space between the aggregate supply and demand, while maintaining the requirement that individual agents act optimally given the prices. Starr (1969) showed the existence of such an approximate equilibrium in nonconvex production economies, in which the maximum imbalance is proportional to the number of goods and a measure of nonconvexity. Heller (1972) proved a similar result with an alternative measure of nonconvexity. More recently, Nguyen and Vohra (2024) proved a bound for markets with indivisible goods that depends only on a measure of preference complementarity of agents. We build on some of these results (summarizing the key results we employ in Appendix B), but depart from this literature by requiring that any feasible mechanism must always specify a feasible outcome. Influenced by computer scientists' approaches to approximations in mechanism design, we will be interested in approximate efficiency and truthfulness, rather than approximate feasibility.[6]

A substantial literature has focused on identifying various conditions on preferences in markets with indivisibilities, under which competitive equilibria exist despite nonconvexities. Contributors include Bikhchandani and Mamer (1997), Gul and Stacchetti (1999), Danilov, Koshevoy, and Murota (2001), Sun and Yang (2006), Milgrom and Strulovici (2009), Hatfield, Kominers, Nichifor, Ostrovsky, and Westkamp (2013), Baldwin and Klemperer (2019) and Nguyen and Vohra (2024). Milgrom (2009) emphasizes the reporting language that agents use when goods are substitutes. None of these papers treat markets with fixed costs such as those described above, in which competitive equilibria do not generally exist. Our analysis seeks to develop practical mechanisms for those settings.

Goeree (2023) introduces an alternative equilibrium concept for nonconvex economies, "Yquilibrium," which involves computing an allocation and prices that minimize the difference between the economic welfare and its dual. Unlike our markup equilibrium, the "Yquilibrium" can lead to agents receiving allocations different from the ones demanded, which creates an additional incentive for agents to misreport.

An alternative approach to establishing equilibrium existence in nonconvex markets is to study the large market limit with a continuum of agents. Aumann (1966) showed that nonconvexities at the level of the individual firm or consumer are no barrier to equilibrium existence in an economy with a continuum of traders and divisible goods, while Azevedo, Weyl, and White (2013) demonstrated a similar result for quasi-linear economies with indivisibilities. In this paper, our markup mechanism exists and is resource-feasible and nearly efficient even in finite economies. Yet another approach is to allow nonlinear or personalized pricing rules, as explored by Wilson (1993), Chavas and Briec (2012), Azizan, Su, Dvijotham, and Wierman (2020) and others, but mechanisms that use anonymous linear prices may sometimes be preferred for other reasons, including ones related to communication and computation as well as familiarity and perceived fairness.

A two-price solution to equilibrium nonexistence has also been proposed in a contemporaneous contribution of Feldman, Shabtai, and Wolfenfeld (2021). The key difference between our approaches is the struc-

[6] Scarf (1967) also features an approximate efficiency objective.

ture and role of the two prices: Feldman et al. (2021) consider (one-sided) exchange economies in which buyers who are allocated a good face one price for the good and buyers who are not allocated a good face a different price for the same good. The role of the two prices in their mechanism is to prevent buyers from wanting to change their bundle of goods from the one allocated by the market designer. We restrict attention to mechanisms that use the same price vector for all buyers (and similarly all sellers) regardless of whether they are allocated a good, which makes achieving incentive-compatibility more difficult.

Our study is motivated by several important applications of linear pricing mechanisms with nonconvex production. In particular, we have taken inspiration from the novel market design for fisheries rights in New South Wales, Australia, introduced by Bichler, Fux, and Goeree (2018, 2019), in which the need to implement sustainable catches led to the exit of fishing boats, with an associated loss of fixed costs. Other sectors with nonconvexities that have used linear prices for electricity generation, with their large start-up and ramping costs, radio spectrum, where geographical complementarities can cause exposure problems.[7]

2 Model and preliminaries

2.1 Model

We employ a Walrasian model with a set of *buyers* N and a set of *firms* or *sellers* F, both finite. Together $A = N \cup F$ is the set of *agents*. There are L varieties of consumable *goods* and a numeraire, money.

Each buyer $n \in N$ chooses a consumption bundle in X, a compact subset of $\mathbb{R}^L_+$ containing 0, called the *consumption possibility set*. Buyer n has quasilinear preferences[8] over bundles in X with a continuous *valuation function* $u_n : X \to \mathbb{R}$, so that the buyer's *utility* associated with receiving allocation x_n and paying t is $U_n(x_n, t) = u_n(x_n) - t$. We suppose that the valuation functions are bounded, nondecreasing with respect to the partial order $\geq$ on $\mathbb{R}^L_+$ and normalized so that $u_n(0) = 0$. We let $\mathscr{U}$ be the space of possible valuation functions for the buyers, which we assume is admissible in the sense of Aumann (1963).[9]

Each seller $f \in F$ chooses a production bundle in the *production possibility set* Y, a compact subset of $\mathbb{R}^L_+$ containing 0. Seller f has a *cost function*[10] $c_f : Y \to \mathbb{R}_+$ which allows us to write f's *profit* from producing

7 See Liberopoulos and Andrianesis (2016) for a summary of pricing mechanisms used in electricity markets with nonconvexities, most of which include "uplift" (or side-payments) in addition to linear pricing, and Ausubel and Milgrom (2002) for a discussion of complementarities in spectrum auctions.

8 The quasilinearity assumption allows our analysis to abstract from income effects, as is usual in mechanism design analyses. For more discussion of the role of income effects see Morimoto and Serizawa (2015).

9 That is, it is possible to define a measure on $\mathscr{U}$, equipped with an appropriate σ-algebra. For example the set of bounded, continuous functions on a compact subset of $\mathbb{R}^L$ is admissible, as is the set of bounded functions with discontinuities of the first kind, or more generally, any subset of a Baire class (Aumann, 1963).

10 Note that sellers in this economy could equivalently be thought of as buyers with valuations $-c_f(y_f)$ and payments $-t$.

$y_f \in Y$ and receiving payment t as $\pi_f(y_f, t) = t - c_f(y_f)$. The cost functions are nondecreasing with respect to the partial order $\geq$ on $\mathbb{R}^L_+$ and normalized so that $c_f(0) = 0$. Let $\mathscr{C}$ be the space of sellers' cost functions, which we also assume to be admissible.

An *economy* $\mathscr{E}$ consists of buyers with their valuation functions and sellers with their cost functions, so that we may write $\mathscr{E} = \langle N, (u_n)_{n\in N}, F, (c_f)_{f\in F}\rangle$. When it is clear, we use the shorthand $\mathscr{E} = \langle N, F\rangle$. At times, it is also convenient to associate $\mathscr{E}$ with the normalized counting measures μ on $\mathscr{U}$ and ν on $\mathscr{C}$ defined by

$$\mu(u_n) = \frac{\text{\# of buyers in } \mathscr{E} \text{ with valuation function } u_n}{|N|},$$
$$\chi(c_f) = \frac{\text{\# of sellers in } \mathscr{E} \text{ with cost function } c_f}{|F|},$$

and to let $\phi = \frac{|F|}{|N|}$, so that $\langle N, \mu, \phi, \chi\rangle$ is an alternative specification of economy $\mathscr{E}$.

Throughout, we will suppose that agent types—that is, the valuation functions u_n of buyers and cost functions c_f of sellers—are private information, but $\mathscr{U}, \mathscr{C}$, $|N|$ and $|F|$ are common knowledge. In some results, we specialize to an *independent private valuations (IPV)* model, in which buyer types are drawn i.i.d. from a common knowledge distribution μ on $\mathscr{U}$ and seller types are drawn i.i.d. from common knowledge χ on $\mathscr{C}$.

Allocations and efficiency An *allocation* $\mathbf{x} = ((x_n)_{n\in N}, (y_f)_{f\in F})$ is an assignment of consumption bundles $x_n \in X$ to each buyer $n \in N$ and production bundles $y_f \in Y$ to each seller $f \in F$. An allocation is *feasible* if $\sum_{n\in N} x_n \leq \sum_{f\in F} y_f$. We denote by $\mathbf{X}$ the set of all feasible allocations.

We define the *surplus* $\mathcal{S}(\mathbf{x})$ associated with allocation $\mathbf{x} \in \mathbf{X}$ by

$$\mathcal{S}(\mathbf{x}) = \sum_{n\in N} u_n(x_n) - \sum_{f\in F} c_f(y_f).$$

The *efficient allocation* problem is to solve

$$\max_{\mathbf{x}\in\mathbf{X}} \mathcal{S}(\mathbf{x}), \tag{P}$$

with a solution denoted by $\mathbf{x}^* \in \arg\max_{\mathbf{x}\in\mathbf{X}} \mathcal{S}(\mathbf{x})$ and the resulting surplus $\mathcal{S}^* = \mathcal{S}(\mathbf{x}^*)$.

For any allocation $\mathbf{x} \in \mathbf{X}$, we will refer to $\mathcal{S}(\mathbf{x}) - \mathcal{S}^*$ as the *deadweight loss* of $\mathbf{x}$ and the ratio $\frac{\mathcal{S}(\mathbf{x})-\mathcal{S}^*}{\mathcal{S}^*}$ as the

However, we will be interested in mechanisms that may charge buyers and sellers different prices, and so it is convenient to distinguish the two groups in our notation.

percentage loss at $\mathbf{x}$.[11]

Pricing rules To prepare for our markup equilibrium, we allow two different price vectors $p^b, p^s \in \mathbb{R}^L_+$ for buyers and sellers such that if any buyer n purchases a bundle x, it makes a payment of $t = p^b \cdot x$ and if seller f supplies y, it receives a payment of $t = p^s \cdot y$.

Denote buyer n's utility by $U_n(x, p^b \cdot x) = u_n(x) - p^b \cdot x$. Its *demand correspondence* $D_n : \mathbb{R}^L_+ \rightrightarrows X$ maps each price vector p^b to the set of utility-maximizing bundles $D_n(p^b)$. Its *indirect utility* function is $\widehat{u}_n(p^b) = \max_{x \in X} u_n(x) - p^b \cdot x$. Similarly, denote seller f's profit by $\pi_f(y, p^s \cdot y)$ and its *supply correspondence* by $S_f : \mathbb{R}^L_+ \rightrightarrows Y$, which maps each price vector p^s to the set of profit-maximizing bundles $S_f(p^s)$. Its *indirect profit* function is $\widehat{\pi}_f(p^s) = \max_{y \in Y} p^s \cdot y - c_f(y)$.

2.2 Convex quasilinear economies

Convexity is defined with respect to the set of payoff-improving allocations for an agent in the economy. The $\bar{u}$-*upper contour set* of buyer $n \in N$ is defined by

$$UC_n^{\bar{u}} = \{(x, t) \in X \times \mathbb{R} \,:\, U_n(x, t) \geq \bar{u}\},$$

while the $\bar{\pi}$-*upper contour set* of seller $f \in F$ is given by

$$UC_f^{\bar{\pi}} = \{(y, t) \in Y \times \mathbb{R} \,:\, \pi_f(y, t) \geq \bar{\pi}\}.$$

We say that buyer n has *convex preferences* if the buyer's feasible set X is convex and the upper contour set $UC_n^{\bar{u}}$ is convex for all $\bar{u} \in \mathbb{R}$, which is equivalent to the quasiconcavity of U_n and the concavity of the valuation function u_n. Seller f has *convex technology* if Y is convex and $UC_f^{\bar{\pi}}$ is convex for all $\bar{\pi} \in \mathbb{R}$, which is equivalent to the quasiconcavity of π_f and the convexity of cost function c_f.

Under the assumption of quasilinearity and the convexity of agents' preferences and technologies, we have the following statement of the fundamental welfare theorems of Arrow (1951) and Debreu (1951). [12]

[11] Later we will make assumptions to rule out cases where $\mathcal{S}^* = 0$ so that this ratio is well-defined.

[12] The statement of Proposition 1 is stronger than the classic statements of the welfare theorems in the 'only if' direction, which is possible due to the quasilinear form of the utility and profit functions. Without quasilinearity or an alternative assumption, it may only be possible to find prices so that agents are expenditure-minimizing for a given level of utility or profit, that is, a price quasiequilibrium with transfers. With quasilinearity and convexity, the efficient allocation program is a convex program. Since there exists a feasible allocation, Slater's Theorem (see, for example, Boyd and Vandenberghe (2004)) implies strong duality. A solution p^* to the dual program, $\inf_{p \in \mathbb{R}^L_+} \sum_{n \in N} \widehat{u}_n(p) + \sum_{f \in F} \widehat{\pi}_f(p)$, and an efficient allocation $\mathbf{x}^*$ comprise a saddle point for the Lagrangian $\mathscr{L}(\mathbf{x}, p) = \sum_{n \in N} u_n(x_n) - \sum_{f \in F} c_f(y_f) - p \cdot \left(\sum_{n \in N} x_n - \sum_{f \in F} y_f\right)$, so that for

Proposition 1. *Suppose in (quasilinear) economy $\mathscr{E}$ that all buyers $n \in N$ have convex preferences and all sellers $f \in F$ have convex technologies. Then a feasible allocation $\mathbf{x} \in \mathbf{X}$ is efficient if and only if there exists $p \in \mathbb{R}^L_+$, $p \neq 0$ such that for all $n \in N$, $x_n \in D_n(p)$; for all $f \in F$, $y_f \in S_f(p)$; and $\sum_{n \in N} p \cdot x_n = \sum_{f \in F} p \cdot y_f$. The pair $(p, \mathbf{x})$ is a* competitive *or* Walrasian equilibrium.

2.3 Measures of nonconvexity and approximate equilibria

The nonconvexity of a set S can be measured in several ways by comparing S and its convex hull, $\mathrm{co}(S)$. We will work with the following measures of nonconvexity of a set:

- The *inner radius* of S is $r(S) = \sup_{x \in \mathrm{co}(S)} \inf_{T \subseteq S \,:\, x \in \mathrm{co}(T)} \mathrm{rad}(T)$.
- The *inner distance* of S is $\rho(S) = \sup_{x \in \mathrm{co}(S)} \inf_{y \in S} \|x - y\|$.

For a convex set S, both measures are zero: $r(S) = 0 = \rho(S)$. The two functions, which are illustrated in Figure 2, measure the size of the set of points in $\mathrm{co}(S)$ but missing from S.

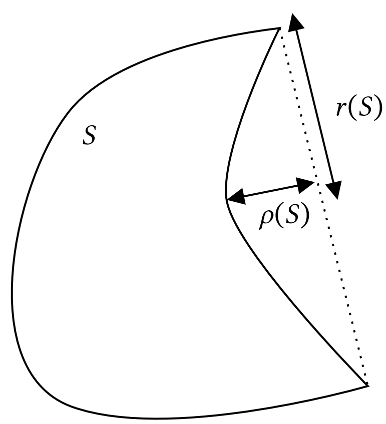

Figure 2: Measures of nonconvexity of a set

The nonconvexity of the preferences of buyer $n \in N$ may be measured by the largest inner radius or inner distance of their upper contour sets, that is $r_n = \sup_{\bar{u} \in \mathbb{R}} r(UC_n^{\bar{u}})$ or $\rho_n = \sup_{\bar{u} \in \mathbb{R}} \rho(UC_n^{\bar{u}})$. Similarly, the nonconvexity of the technology of seller $f \in F$ may be measured by $r_f = \sup_{\bar{\pi} \in \mathbb{R}} r(UC_f^{\bar{\pi}})$ or $\rho_f = \sup_{\bar{\pi} \in \mathbb{R}} \rho(UC_f^{\bar{\pi}})$. Let $r_{\mathscr{E}}$ and $\rho_{\mathscr{E}}$ denote the largest of such measures among all the buyers and sellers in economy $\mathscr{E}$.

When agents' upper contour sets are not convex, the second welfare theorem may not hold and there may be no competitive equilibrium. Proposition 2, the Shapley-Folkman Lemma, assists in identifying

any $\mathbf{x} \in \mathbf{X}$, $\mathscr{L}(\mathbf{x}, p^*) \leq \mathscr{L}(\mathbf{x}^*, p^*)$. Because the Lagrangian is separable across agents, the saddle point condition implies $x_n^* \in D_n(p^*)$ and $y_f^* \in S_f(p^*)$.

allocations which are nearly competitive equilibria.

Proposition 2 (Shapley-Folkman Lemma[13]). *Let $S_i \subseteq \mathbb{R}^L$ for $i = 1, ..., M$, and let $S = \bigoplus_{i=1}^M S_i$ be the Minkowski sum of those sets. Then for any $x \in \text{co}(S)$, $x = \sum_{i=1}^N x_i$ where $x_i \in \text{co}(S_i)$ and $|i : x_i \in \text{co}(S_i) \setminus S_i| \leq L' := \min(L, M)$. Moreover, there exists $y, y' \in S$ such that $\|x - y\| \leq (\max_i r(S_i))\sqrt{L'}$ and $\|x - y'\| \leq (\max_i \rho(S_i))L'$.*

Proposition 2 has been used to establish results about *approximate equilibria*, which are constructed as follows. First, consider a convexified version of the nonconvex economy in which the upper contour sets of all agents are replaced by their convex hulls. This is equivalent to replacing each buyer's valuation function u_n by its concave envelope $\text{cav}(u_n)$ and each seller's cost function c_f by its convex envelope, $\text{vex}(c_f)$. [14] The *convexified economy* is then $\widehat{\mathscr{E}} = \langle N, (\text{cav}(u_n))_{n\in N}, F, (\text{vex}(c_f))_{f\in F}\rangle$.

By Proposition 1, the convexified economy has a competitive equilibrium which is efficient (according to the concavified valuation functions and convexified cost functions). Since the convexified economy's efficient allocation problem is a relaxation[15] of the same problem for the original economy, the efficient surplus of the convexified economy is an upper bound on the efficient surplus of the original economy. We call the resulting price-allocation pair $(p, \mathbf{x})$ a *pseudoequilibrium* of the actual economy $\mathscr{E}$. Proposition 2 implies that $\mathbf{x}$ can be chosen so at most L' agents in $\mathscr{E}$ are not utility- or profit-maximizing at $\mathbf{x}$ given prices p and that there is a nearby allocation $\mathbf{x}'$ such that *all* agents are maximizing given prices p,[16] but markets may not exactly clear at $\mathbf{x}'$. The price-allocation pair $(p, \mathbf{x}')$ is called an *approximate equilibrium*.

Pseudoequilibria and approximate equilibria describe allocations rather than mechanisms. These allocations can be infeasible or may impose large losses on some agents. This may make them inappropriate for use as mechanisms for practical market designs. We utilize these ideas to devise mechanisms that are

[13] It is perhaps most accurate to refer only to the result in the second sentence of Proposition 2 as the Shapley-Folkman Lemma, althought it was first reported by Starr (1969) as a result of private communication with Lloyd Shapley and Jon Folkman. Starr (1969) then proved the first half of sentence three of Proposition 2, while Heller (1972) proved the second half. For simplicity, we will refer to the whole of Proposition 2 as the Shapley-Folkman Lemma. Budish and Reny (2020) provide an improved bound for the Shapley-Folkman Lemma involving a different measure of nonconvexity of the set, which could also be applied in our setting to improve the constant but not the asymptotic rate of convergence in several of our results.

[14] Recall that the concave envelope of a function is the pointwise smallest concave function everywhere above that function, while the convex envelope of a function is the pointwise largest convex function everywhere below that function.

[15] That is, the constraint set is weakly larger than the original constraint set and the objective function is pointwise weakly larger than the original objective.

[16] To see this, note that if a buyer is assigned a bundle x_n in $\mathbf{x}$ that is not utility-maximizing at p, then x_n must be the convex combination of bundles (x_n') in X which are *exposed points* in u_n (i.e. where $\text{cav}(u_n) = u_n$), and that the agents in the convexified economy must be indifferent between x_n and these bundles. That is, the concavified portions of buyers' utility functions consist of (patches of) hyperplanes, and if an agent is assigned a bundle on such a patch, then the price vector must be normal to that hyperplane. This implies that the original buyer must be maximizing at bundles in (x_n'), which are on the relative boundaries of the patch of hyperplane.

computable, select feasible allocations, and have the other desirable properties that we listed earlier.

3 Bound-Form First Welfare Theorem

When competitive equilibrium does not exist, no feasible allocation is supported by a single anonymous price vector that is the same for buyers and sellers, but resource-feasibility can be restored by varying the price vectors for the two sides of the market or by rationing some agents, requiring them to accept bundles that are not their most preferred ones at the specified prices. Given an allocation and prices, we can characterize the welfare effects of rationing on buyers and sellers in terms of *rationing losses*, which are defined as the excess of the payoff an agent would obtain from its most preferred bundle given the prices compared to the payoff it receives in the prescribed allocation.

Definition 3.1. The *rationing loss* $\mathscr{R}_n(p, x)$ *of buyer* n at price p and allocation x is

$$\mathscr{R}_n(p, x) = \hat{u}_n(p) - U_n(x, p \cdot x).$$

The *rationing loss* $\mathscr{R}_f(p, y)$ *of seller* f is

$$\mathscr{R}_f(p, y) = \hat{\pi}_f(p) - \pi_f(y, p \cdot y).$$

The *rationing losses* of allocation $\mathbf{x} = ((x_n)_{n\in N}, (y_f)_{f\in F})$ at price p is defined by

$$\mathscr{R}(p, \mathbf{x}) = \sum_{n\in N} \mathscr{R}_n(p, x_n) + \sum_{f\in F} \mathscr{R}_f(p, y_f).$$

If competitive equilibrium does not exist, any price-allocation pair must entail rationing or wasted supply (and thus budget deficit) or both. In our first main result, we show that the extent of such rationing and budget losses fully characterize the efficiency of the allocation.

Theorem 1 (Bound-Form First Welfare Theorem)**.** *Let* $p \in \mathbb{R}^L_+$ *be a price vector and* $\mathbf{x} = ((x_n)_{n\in N}, (y_f)_{f\in F})$ *be any allocation. Then, the deadweight loss of allocation* $\mathbf{x}$ *satisfies*

$$\underbrace{\mathcal{S}^* - \mathcal{S}(\mathbf{x})}_{\text{deadweight loss}} \leq \underbrace{\mathscr{R}(p, \mathbf{x})}_{\text{rationing loss}} + \underbrace{p \cdot \left(\sum_{f\in F} y_f - \sum_{n\in N} x_n\right)}_{\text{budget deficit}}.$$

Proof. Fix any efficient allocation $\mathbf{x}^*$. By the definitions of indirect utility and consumer surplus, the following must hold for any prices:

$$\hat{u}_n(p) \geq u_n(x_n^*) - p \cdot x_n^*$$
$$\hat{\pi}_f(p) \geq p \cdot y_f^* - c_f(y_f^*).$$

Summing these inequalities, we obtain

$$\sum_{n \in N} \hat{u}_n(p) + \sum_{f \in F} \hat{\pi}_f(p) + p \cdot \left(\sum_{n \in N} x_n^* - \sum_{f \in F} y_f^* \right) \geq \sum_{n \in N} u_n(x_n^*) - \sum_{f \in F} c_f(y_f^*) = \mathcal{S}^*.$$

Since $\mathbf{x}^*$ is feasible, the third term on the left side is nonpositive, so

$$\sum_{n \in N} \hat{u}_n(p) + \sum_{f \in F} \hat{\pi}_f(p) \geq \mathcal{S}^*$$

Subtracting $\mathcal{S}(\mathbf{x})$ and applying the definitions,

$$\mathcal{S}^* - \mathcal{S}(\mathbf{x}) \leq \sum_{n \in N} \hat{u}_n(p) + \sum_{f \in F} \hat{\pi}_f(p) - \mathcal{S}(\mathbf{x}) = \mathcal{R}(p, \mathbf{x}) + p \cdot \left(\sum_{f \in F} y_f - \sum_{n \in N} x_n \right),$$

which is what we sought to prove. □

The Bound-Form First Welfare Theorem extends the First Welfare Theorem for quasilinear economies by applying to any price-allocation pair $(p, \mathbf{x})$ rather than just to competitive equilibria. If $(p, \mathbf{x})$ is a Walrasian equilibrium, then both the budget deficit and the rationing losses are zero, so the theorem asserts that the welfare loss is zero, or equivalently that any Walrasian equilibrium is efficient.

One interpretation of the First Welfare Theorem is that prices act as a "certificate of optimality": given some allocation, if supporting prices exist for it, then that allocation is efficient. As Scarf (1994) lamented, in the absence of convexity, there is, in general, no such optimality test. However, one interpretation of Theorem 1 is as an *approximate* optimality test: if we can identify a price-allocation pair for which the rationed surplus plus the net budget deficit is small, then the welfare loss is small as well. Theorem 1 also begins to link incentives to efficiency: since for any fixed price, agents would prefer not to be rationed, Theorem 1 suggests that a pricing mechanism with little rationing and in which individual agents have little influence over prices will have both good incentive properties and small deadweight losses. These observations are key to our extensions below.

4 The markup mechanism

4.1 Pricing mechanisms and approximate mechanism design

In this section, we study pricing mechanisms that map profiles of reports of sellers' cost functions $(c_f)_{f\in F}$ and buyers' value functions $(u_n)_{n\in N}$ to a feasible allocation $\mathbf{x} \in \mathbf{X}$ and anonymous prices for buyers and sellers, $\boldsymbol{p} = (p^b, p^s)$.[17] We require that the mechanism specify outcomes that are both *resource feasible* and *budget-feasible* so that for all report profiles, $\sum_{f\in F} y_f \geq \sum_{n\in N} x_n$ and $p^s \cdot \sum_{f\in F} y_f \leq p^b \cdot \sum_{n\in N} x_n$. We do not delve into the important question of how agents communicate their potentially complicated costs and values to the mechanism.[18]

For convenience, we restate some familiar definitions. A pricing mechanism is

(a) *efficient* if the output allocation $\mathbf{x}$ is an efficient allocation given the reported value and cost functions,

(b) *ex post incentive-compatible (EPIC)* if truthful reporting is an *ex post* Nash equilibrium of the reporting game induced by the mechanism,

(c) *interim incentive-compatible (IIC)* if truthful reporting maximizes each agents' expected payoffs under the mechanism, and

(d) *individually-rational (IR)* if, given reported value and cost functions, the allocation and prices determined by the mechanism delivers each agent a payoff (utility or profit) no worse than non-participation (here, 0).

It is typically impossible for a mechanism to exactly satisfy these desirable properties; instead, we seek mechanisms satisfying appropriate approximations to these goals. A pricing mechanism is

(a) *ε-efficient* if the deadweight loss of $\mathbf{x}$ is bounded by ε given the reports,

(b) *ε-EPIC* if truthful reporting is an *ε-ex post* Nash equilibrium,

(c) *ε-IIC* if, for each agent, the expected payoff associated with any report in the mechanism is no more than ε greater than that of the truthful report, and

[17] We choose not to consider randomized mechanisms, both because these are unnecessary to achieve our objectives and because they raise daunting practical issues, including most importantly very high trust requirements in the mechanism designer and the possible failure of ex post individual rationality for ex ante individually rational lotteries.

[18] The design of reporting languages to report complex preferences for economic mechanisms has been studied by Milgrom (2009), Bichler, Goeree, Mayer, and Shabalin (2014), Bichler, Milgrom, and Schwarz (2022) and others.

(d) ε*-individually-rational* if each agent obtains a payoff of at least $-\varepsilon$ given its report.

The *Walrasian mechanism* inputs reports of value and cost functions and outputs a Walrasian equilibrium price and allocation. In case there are multiple Walrasian equilibria, we suppose that the mechanism has a predetermined rule for selecting among them (and we make a similar assumption when multiplicity arises in the other mechanisms discussed in this paper). Walrasian mechanisms are efficient, individually rational and have good large-market incentive properties (discussed further in Section 4.3) but can only be applied in settings in which Walrasian equilibria are guaranteed to exist.

We are interested in extensions of the Walrasian mechanism that can be applied in settings where Walrasian equilibria may not exist. We seek mechanisms that perform well in large markets, with approximations to efficiency that depend on $|N|$. Let $\mathscr{E}_t = \langle N_t, \mu_t, \phi_t, \chi_t \rangle$ be a sequence of economies indexed by $t = 1, 2, ...$ We make the following additional assumptions as the economy grows large.

Assumption 1 (Existence of limit economy)**.** *As $t \to \infty$, $|N_t| \to \infty$ and $\phi_t \to \phi \in (0, 1)$. Furthermore, μ_t converges weakly to probability measure μ_∞ on $\mathscr{U}$ and χ_t converges weakly to measure χ_∞ on $\mathscr{C}$.*

Assumption 2 (Individual nonconvexities are bounded)**.** *There exists $R > 0$ with $R_{\mathscr{E}_t} < R$ for all t.*

Assumption 3 (Growing gains from trade)**.** *As $t \to \infty$, the efficient surplus $\mathcal{S}_t^*$ grows at least as quickly as $|N_t|$ asymptotically, or in Knuth's (1976) asymptotic notation, $\mathcal{S}_t^* = \Omega(|N_t|)$.*[19]

Assumption 4 (Prices are bounded)**.** *There exists $M > 0$ such that the minimum distance between demand and supply, $d_H\left(\sum_{n\in N_t} D_n(p), \sum_{f\in F_t} S_f(p)\right)$,*[20] *is at least R for sufficiently large t and $p \in \mathbb{R}_+^L$ such that $\|p\| < \frac{1}{M}$ or $\|p\| > M$.*

Assumption 1 asserts that for large t, the important variation among economies is their scale: the proportions of various types converge to a limit. Assumption 2 asserts that there is a uniform bound on the size of any nonconvexity across the sequence of markets, limiting the impact of any single firm's or consumer's nonconvexity in a large economy.[21] Assumption 3 is the condition that the efficient surplus per participant is bounded away from zero. Assumption 4 implies that there is a compact set of possible prices for which the aggregate demand and supply are nearly equal.

[19] Recall that $f(x) = \Omega(g(x))$ if $\liminf_{x\to\infty} |f(x)|/g(x) > 0$.

[20] Here, d_H is the Hausdorff distance, where $d_H(S, S')$ is defined by letting $\mathrm{dist}(x, S) = \inf_{y\in S} \|x - y\|$ and $d_H(S, S') = \max\{\sup_{x\in S'} \mathrm{dist}(x, S), \sup_{x\in S} \mathrm{dist}(x, S')\}$.

[21] It suffices here to assume that the consumption possibility set X does not grow with t, since $R \le \mathrm{rad}(X)$.

4.2 Markup mechanisms

We now introduce *markup mechanisms*, which are designed to maintain a no-rationing property similar to that of the Walrasian mechanism. If agents are not rationed, it may be impossible to find prices such that all supply is demanded by buyers. In order to pay for any unwanted supply provided by the firms, we use a markup on the prices paid by buyers.

Definition 4.1 (Markup equilibrium)**.** A *markup equilibrium* is a triple $(\alpha, p, \mathbf{x})$ consisting of a markup parameter $\alpha \geq 0$, a price $p \in \mathbb{R}^L_+$ and a resource-feasible allocation $\mathbf{x}$ such that:

(a) payments for sellers are determined by price vector p and sellers are not rationed given these prices, so $y_f \in S_f(p)$;

(b) payments for buyers are determined by price vector $(1+\alpha)p$ and buyers are not rationed given these prices, so $x_n \in D_n((1+\alpha)p)$; and

(c) budgets are at least weakly balanced: $\sum_{n \in N}(1+\alpha)p \cdot x_n - \sum_{f \in F} p \cdot y_f \geq 0$.

A *markup mechanism* is a mechanism that inputs reports of cost and value functions and outputs a markup equilibrium. We are especially interested in markup mechanisms that select α close to zero and leave few goods unallocated, because their allocations are nearly efficient. This follows by applying Theorem 1 at the price-allocation pair $(p, \mathbf{x})$: if few goods are unallocated, the budget deficit at price p is small, while if prices p and $(1+\alpha)p$ are close, the rationing losses for each buyer at price p are small. This latter claim follows by an envelope theorem argument, which formalized in Proposition 3.

Proposition 3. *Let bundle x be demanded by buyer n at price p, so $x \in D_n(p)$. Consider some other price $p' \neq p$. Then the rationing loss of buyer n at allocation x given price p', $\mathcal{R}_n(x, p')$ is $O(\|p - p'\|)$.*[22]

If computational challenges were not a concern, a market designer may seek to identify a markup mechanism with the smallest loss, which we call a *minimum markup mechanism*. The pair (p, x) is a Walrasian equilibrium if and only if the triple $(x, p, 0)$ is a markup equilibrium, and in that case the latter is also the minimum markup equilibrium. In nonconvex economies, however, solving for the minimum markup mechanism is difficult. We now show that an $O(1)$-efficient markup mechanism—that is, one for which the percentage loss in welfare is at most inversely proportional to the number of agents in the economy—can be identified using only *convex* optimization problems and a one-dimensional binary search. Before providing a technical description of this pricing mechanism and its analysis, we sketch intuitively the steps

[22] Recall Knuth's (1976) big O notation: $f(x) = O(g(x))$ if $\limsup_{x \to \infty} |f(x)|/g(x) < \infty$.

of our approach, echoing the example in the introduction.

For a fixed α, we select $(p, \mathbf{x})$ to be the equilibrium price-and-allocation pair of a related economy with three changes from the actual economy: (1) every buyer's value function is replaced by the smaller function $u_n/(1+\alpha)$, (2) all values and costs are then replaced by their concave or convex hulls, respectively,[23] and (3) we add an *operational reserve* for each good, which is a quantity demanded by the auctioneer in the amount of $R := \min\{r_{\mathscr{G}}\sqrt{L}, \rho_{\mathscr{G}} L\}$. Step (1) in this construction leads to prices and allocations such that buyers would demand exactly the same allocations using their actual value functions u_n and their marked-up prices of $(1+\alpha)p$. We apply the Shapley-Folkman Lemma (Proposition 2) to round $\mathbf{x}$ to one of the demanded allocations for each agent while changing the net demand for each good by at most R units. To maintain feasibility after that change, we balance with an offsetting change from the reserve allocated to the auctioneer in step (3).

Excluding the auctioneer's demand, the resulting final allocation always has supply greater than or equal to demand—it is resource-feasible—and its excess supply is no more than $2R$ units of each good. Any excess supply can result in a loss of efficiency (units allocated to the auctioneer are wasted), but the quantity allocated to the auctioneer is bounded by an amount that is independent of the size of the market.[24]

Since the price vector and the excess supply of goods in a markup equilibrium are bounded, the budget imbalance at price p is also bounded. As trade increases with the size of the economy, the markup, α, needed to guarantee budget balance is inversely related to market size: it is $O(1/N)$. Thus, the total welfare loss from the markup mechanism is bounded by a constant plus a term that is inversely proportional to market size.

Definition 4.2 (simple markup mechanism)**.** The *simple markup mechanism* is the markup mechanism with parameters $(\alpha^*, p^*, \mathbf{x}^*)$ determined as follows. If all reported values are concave and all reported costs are convex, set $\alpha^* = 0$ and choose $(p^*, \mathbf{x}^*)$ to be some Walrasian equilibrium. Otherwise, for each $\alpha > 0$, consider the following convex program:

[23] This need not be computationally expensive. For example, if the value and cost functions are reported to the mechanism using a mixed integer program, the mechanism may simply convert integer variables to real variables to obtain the convex hulls in the form of linear or quadratic programs.

[24] The choice of R units of each good as a set-aside for the auctioneer in step (3) is a theoretical guarantee. It might be possible to allocate fewer units to the auctioneer in step (3) and arrive at a more efficient feasible allocation using the same approach. We note that an alternative approach could be to start by checking for feasible allocations with zero units set aside (these would correspond to competitive equilibria) and then increase the set-aside intelligently until a budget-feasible markup mechanism is identified, but we leave such details for future research.

$$\min_{p\in\mathbb{R}^L_+}\max_{x_n\in\operatorname{co}(X),y_f\in\operatorname{co}(Y)}\sum_{n\in N}\frac{\operatorname{cav}(u_n)(x_n)}{1+\alpha}-\sum_{f\in F}\operatorname{vex}(c_f)(y_f)-p\cdot\left(\sum_{n\in N}x_n+R1_L-\sum_{f\in F}y_f\right),$$

where 1_L is the vector of ones in $\mathbb{R}^L$. Let $(p^\alpha, \tilde{\mathbf{x}}^\alpha)$ denote any solution to this program.

From $\tilde{\mathbf{x}}^\alpha$, we obtain, via Proposition 2, an allocation $\mathbf{x}^\alpha$ with $\|\mathbf{x}^\alpha-\tilde{\mathbf{x}}^\alpha\|\le R$ such that $x_n^\alpha\in\max_{x\in X}\frac{1}{1+\alpha}u_n(x)-p\cdot x$ for each $n\in N$ and $y_f^\alpha\in S_f(p)$. By construction, this $\mathbf{x}^\alpha$ will be feasible in $\mathscr{E}$. Let

$$\alpha^*=\min\left\{\alpha\,\middle|\,\sum_{n\in N}(1+\alpha)p^\alpha\cdot x_n^\alpha-\sum_{f\in F}p^\alpha\cdot y_f^\alpha\ge 0\right\},\tag{A}$$

and define $p^*=p^{\alpha^*}$ and $\mathbf{x}^*=\mathbf{x}^{\alpha^*}$.

In Theorem 2, we show that the simple markup mechanism is well-defined (that is, the minimum in (A) exists) and the resulting mechanism is $O(1/|N_t|)$-efficient. Before doing so, we note some other important properties of the markup mechanism that follow directly from the construction. First, the equilibrium is resource-feasible and budget-feasible. Second, the equilibrium allocation and payments are individually rational for each agent. For sellers, this follows because their profits are identical to those in the pseudoequilibrium used in the construction. For buyers, the pseudoequilibrium price p^α and consumption allocation in $\tilde{\mathbf{x}}^\alpha$ satisfy

$$\frac{1}{1+\alpha}u_n(\tilde{x}_n^\alpha)-p^\alpha\cdot\tilde{x}_n^\alpha=\frac{1}{1+\alpha}u_n(x_n^\alpha)-p^\alpha\cdot x_n^\alpha\ge 0$$

so that $u_n(x_n^\alpha)-(1+\alpha)p^\alpha\cdot x_n^\alpha\ge 0$ as well.

Theorem 2. *Let $\mathscr{E}_t$ be a sequence of economies satisfying Assumptions 1–4. Then*

(a) the simple markup mechanism is well-defined (that, is the minimum in (A) is attained),

(b) the simple markup mechanism's markup $\alpha^*\le O(1/|N_t|)$, and

(c) the deadweight loss of the simple markup mechanism's allocation is $O(1)$, so the percentage loss is $O\left(1/|N_t|\right)$.

Although the rates of convergence in Theorem 2 are stated in terms of $|N_t|$, by Assumption 1, the same asymptotic rate of convergence holds with respect to $|F_t|$ or $|A_t|$.

4.3 Incentives

In the Walrasian and markup mechanisms, both buyers and sellers receive their optimal bundles given their prices, so an agent can profit from false reports only to the extent that it can influence its prices. Moreover, because the prices in the simple markup mechanism are Walrasian equilibrium prices of a related convex economy, the limited ability of agents to manipulate Walrasian prices in large economies implies a similar difficulty for the markup mechanism.

We now briefly discuss the most relevant literature related to the agents' ability to influence prices in large markets. Roberts and Postlewaite (1976) initiated the formal literature of this subject in a study of a sequence of exchange economies with the number of agents going to infinity. They represented the sequence of economies by measures μ_t on $\mathscr{U}$ with $\lim_{t\to\infty} \mu_t = \mu_\infty$, showing that if the Walrasian price correspondence is continuous at μ_∞, then each agent's influence on the price goes to zero as t increases. Jackson (1992) shows in the same model that an agent's optimal reported demand converges in the L^∞ norm to the demand associated with that agent's true preferences.

Watt (2022) studies the rate of this convergence, showing that a condition on the demand and supply correspondences, called *strong monotonicity*, ensures fast convergence of incentives in the Walrasian mechanism.

Definition 4.3 (Strong monotonicity)**.**

(a) A buyer n is *active* at price p if $D_n(p) \neq \{0\}$.

(b) Buyer n's demand correspondence $D_n : \mathbb{R}^L_+ \rightrightarrows X$ is *strongly monotone* if there exists some $m > 0$ such that for all p, p' at which buyer n is active and for all $d \in D_n(p)$, $d' \in D_n(p')$,

$$(p - p') \cdot (d' - d) \geq m\|p - p'\|^2.$$

(c) A seller f is *active* at price p if $S_f(p) \neq \{0\}$ and there is some $\beta > 1$ such that $S_f(\beta p) \neq S_f(p)$.

(d) Seller f's supply correspondence $S_f : \mathbb{R}^L_+ \rightrightarrows Y$ is *strongly monotone* if for all p, p' at which seller f is active and for all $s \in S(p)$, $s' \in S(p')$,

$$(p - p') \cdot (s - s') \geq m\|p - p'\|^2.$$

Strong monotonicity is a condition on how quickly demand or supply changes in response to price changes:

in settings with one good, strong monotonicity is equivalent to a lower bound on the absolute value of the slope of the firm's supply curve and the buyer's demand curve. If each buyer has strongly monotone demand and each seller has strongly monotone supply, Watt (2022) shows that the resulting sequence of economies is *perturbation-proof*, which implies that the maximum influence of any one agent on Walrasian prices is $O\left(1/|N_t^a|\right)$, where N_t^a is the set of active agents at the Walrasian price. Furthermore, if each buyer is drawn independently at random from some distribution μ over $\mathscr{U}$ and each seller is drawn independently at random from χ over $\mathscr{C}$ for which the *expected* demand and supply correspondences are strongly monotone,[25] the maximum *ex post* benefit of misreporting under the Walrasian mechanism is $O_P\left(1/|N_t|^{1-\varepsilon}\right)$ for all $\varepsilon > 0$, which means that it is nearly $O\left(1/|N_t^a|\right)$ with probability approaching 1.

While the above results considered ex post incentives, Azevedo and Budish (2019) studied interim incentives and show that the Walrasian mechanism (with a finite set of buyer and seller types) is *strategy-proof in the large*, which implies that the benefit to any agent of misreporting against any full-support, independent and identically-distributed distribution of agent types tends to zero at $O(1/|N_t|^{\frac{1}{2}-\varepsilon})$ for all $\varepsilon > 0$.

Theorem 3, proved in Appendix A.3, adapts these results to the markup mechanism.

Theorem 3. *Let $(\mathscr{E}_t)_{t\in\mathbb{N}} = (\langle N_t, \mu_t, \phi_t, \chi_t\rangle)_{t\in\mathbb{N}}$ be a sequence of economies satisfying Assumptions 1–4 and suppose that a markup mechanism with markups α_t is applied to $\mathscr{E}_t$.*[26]

(a) Suppose $(\mathscr{E}_t)_{t\in\mathbb{N}}$ are economies with values and costs drawn according to independent full-support probability distributions μ and χ defined on finite *type spaces $\mathscr{U}$ and $\mathscr{C}$. Then the markup mechanism is strategy-proof-in-the-large and $O(1/|N_t|^{\frac{1}{2}-\varepsilon})$-IIC for any $\varepsilon > 0$.*

Now suppose that the markup satisfies $\alpha_t \leq O(1/|N_t|)$ as in the minimal and simple markup mechanisms.

(b) Suppose that each buyer in each $\mathscr{E}_t$ has strongly monotone demand and that each seller has strongly monotone supply. Then the markup mechanism is $O(1/|N_t^a|)$-EPIC, where N_t^a is the number of active buyers and sellers at the mechanism's respective prices.

(c) Suppose that $(\mathscr{E}_t)_{t\in\mathbb{N}}$ are economies with values and costs drawn independently according to distributions μ and χ for which the expected demand and supply correspondences are strongly monotone. Then, the maximum ex post benefit of misreporting for any agent is $O_P\left(1/|N_t|^{1-\varepsilon}\right)$, for any $\varepsilon > 0$, and the mechanism is $O(1/|N_t|^{1-\varepsilon})$-IIC.

[25] Here the expectation is with respect to the measure over economies induced by draws of agents from μ and χ. We clarify the meaning of the "expected demand and supply correspondences" in Appendix A.3.

[26] The markups may be determined endogenously, as in the minimal and simple markup mechanisms.

4.4 Computational properties

While equilibrium computation is hard in general,[27] computing the Walrasian equilibrium in concave quasilinear economies reduces to solving a convex optimization problem and its dual. A wide class of such optimization problems are efficiently solvable, including problems with self-concordant or strongly convex objectives. For example, Walrasian prices in economies with strongly monotone supply and demand may be efficiently computed via tâtonnement (Watt, 2022).

In contrast, finding efficient allocations in many nonconvex economies is computationally complex even with quasilinear preferences. For example, the problem of identifying an optimal allocation in the fisheries market of Bichler et al. (2018) involved solving a large integer programming problem. Because integer programming is NP-hard, no efficient optimization algorithm is known for all instances of such problems, although heuristics and approximations are sometimes useful.

The approach in this paper is based on approximation. Conditional on α, our simple markup mechanisms require solving only a convex optimization problem. Identifying the optimal markup α^* in the simple markup mechanism is more challenging, although a binary search algorithm for α might be employed, in practice, to identify a small markup that ensures weak budget balance. Since the welfare loss of the simple markup mechanism is $O(\alpha)$, small markups are associated with small losses. Ahunbay et al. (2024) provide further details on how to adapt our markup mechanism for practical computations, focusing on an application to European wholesale spot electricity markets.

4.5 A related mechanism

Our markup mechanisms are related to a linear pricing mechanism which was proposed for allocating commercial fisheries licenses in New South Wales: the *maximum surplus anonymous pricing* mechanism described by Bichler et al. (2018).[28] That mechanism solves the usual surplus optimization problem

$$\max_{\mathbf{x} \in \mathbf{X}} \sum_{n \in N} u_n(x_n) - \sum_{f \in F} c_f(y_f),$$

but subject to the constraint that there exist prices p^b and p^s satisfying

(a) Individual rationality: for all $n \in N$, $f \in F$, $u_n(x_n) - p^b \cdot x_n \geq 0$ and $p^s \cdot y_f - c_f(y_f) \geq 0$.

(b) Budget balance: $\sum_{n \in N} p^b \cdot x_n \geq \sum_{f \in F} p^s \cdot y_f$.

---

27 See, for example, Chen, Dai, Du, and Teng (2009) and Daskalakis, Goldberg, and Papadimitriou (2009).

28 A variant of this mechanism was later implemented in New South Wales.

A corollary of Theorem 2 is that this mechanism has a deadweight loss that is bounded by a constant independently of the market size $|N|$. We have introduced and analyzed markup mechanisms rather than studying the just-described alternative mechanism for two main reasons. *First*, the alternative mechanism is hard to scale up because it requires solving nonconvex optimization problems, while the simple markup mechanism can be implemented by solving convex optimization problems (plus a binary search). *Second*, the alternative mechanisms may involve rationing at the prevailing prices, which can give agents an additional incentive for false reporting that is avoided by the markup mechanism.

5 Discussion

In some regulated sectors of the economy, with nonconvexities in production or consumption and multiple closely interrelated products, Walrasian-like mechanisms are already in use. For example, in wholesale electricity markets with products distinguished by time and location, producers often incur fixed costs to start their plants, ramping production up or down to deliver power at different times, and large power users rely on prices to guide consumption choices. In markets for fishing permits with products distinguished by species and location, participants pay fixed costs to staff a boat and send it to sea. For these markets with their nonconvexities, Walrasian equilibrium may not exist, so the Walrasian mechanism cannot be implemented unchanged.

This paper adopts a market-design perspective to the problem of extending the Walrasian mechanism to apply in non-convex economies. Its proposed extension draws on two older branches of economics research: one that evaluates reporting incentives in the Walrasian mechanism in convex economies and another that identifies price-allocation pairs that, while not necessarily feasible, are close to Walrasian in other ways. Our market design approach requires that any recommended allocation be exactly resource-feasible – not just approximately so – and may also require budget-feasibility, so that no external funds are needed to operate the system. Our proposed markup equilibrium, which is feasible in both these ways, has no close antecedent in either tradition.

We have shown that markup equilibria exist and have properties similar to those of Walrasian equilibrium. Because markup equilibrium uses linear prices for producers and consumers, it economizes on communication and computation and supports a robust system that many market participants are likely to find familiar and fair. When markups are small, its allocations are nearly efficient. We have shown how small-markup equilibrium can be computed in practice with welfare losses bounded by an amount that is proportional to the largest relevant nonconvexity. Incentives in the markup mechanism depend on paricipants' abilities to manipulate their prices, leading to results resembling those known to apply for the Walrasian mechanism in the case of convex economies.

The markup mechanism that we have studied suffers inefficiency almost exclusively from overproduction. One might wonder: can another mechanism with similar properties do better? Instead of setting different prices for the two sides of the market, an alternative approach would use rationing, forcing participants on one side of the market, say consumers, to consume a bundle that is different from their most preferred ones at the prices they must pay. Our Bound-Form First Welfare Theorem provides a tool to evaluate the welfare losses of any such mechanism, although a separate analysis would be required to assess the incentive of any such alternative mechanism.

We have adopted the market design approach in this paper to be able to treat concerns that are hard to quantify, such as ones about communications, computation, and familiarity. This approach takes a different view about "hard" and "soft" constraints compared to many older theories. Many extensions of Walrasian equilibrium to non-convex economies treat resource constraints and budget constraints as soft ones that should be satisfied only approximately, but our approach requires a mechanism that satisfies those constraints exactly. While many mechanism design analyses require that participants should be unable to gain from misreporting, our market design approach imposes a softer constraint, leading to mechanisms in which the potential gains to misreporting are vanishingly small, but not necessarily zero, in large economies. For some important applications including the two described above, the markup mechanism that satisfies constraints in this way provides an appealing way to confront the real-life problems of market design.

References

AHUNBAY, M. Ş., M. BICHLER, T. DOBOS, AND J. KNÖRR (2024): "Solving large-scale electricity market pricing problems in polynomial time," *European Journal of Operational Research*.

ARROW, K. J. (1951): "An extension of the basic theorems of classical welfare economics," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, The Regents of the University of California.

AUMANN, R. J. (1963): "On choosing a function at random," in *Ergodic Theory*, ed. by F. Wright, Academic Press.

——— (1965): "Integrals of set-valued functions," *Journal of Mathematical Analysis and Applications*, 12, 1–12.

——— (1966): "Existence of competitive equilibria in markets with a continuum of traders," *Econometrica*, 34, 1–17.

AUSUBEL, L. M. AND P. R. MILGROM (2002): "Ascending auctions with package bidding," *Advances in Theoretical Economics*, 1.

AZEVEDO, E. M. AND E. BUDISH (2019): "Strategy-proofness in the large," *Review of Economic Studies*, 86, 81–116.

AZEVEDO, E. M., E. G. WEYL, AND A. WHITE (2013): "Walrasian equilibrium in large, quasilinear markets," *Theoretical Economics*, 8, 281–290.

AZIZAN, N., Y. SU, K. DVIJOTHAM, AND A. WIERMAN (2020): "Optimal Pricing in Markets with Nonconvex Costs," *Operations Research*, 68, 480–496.

BALDWIN, E. AND P. KLEMPERER (2019): "Understanding preferences: "Demand types", and the existence of equilibrium with indivisibilities," *Econometrica*, 87, 867–932.

BATOR, F. M. (1961): "On convexity, efficiency, and markets," *Journal of Political Economy*, 69, 480–483.

BICHLER, M., V. FUX, AND J. GOEREE (2018): "A matter of equality: Linear pricing in combinatorial exchanges," *Information Systems Research*, 29, 1024–1043.

BICHLER, M., V. FUX, AND J. K. GOEREE (2019): "Designing combinatorial exchanges for the reallocation of resource rights," *Proceedings of the National Academy of Sciences*, 116, 786–791.

BICHLER, M., J. GOEREE, S. MAYER, AND P. SHABALIN (2014): "Spectrum auction design: Simple auctions for complex sales," *Telecommunications Policy*, 38, 613–622.

BICHLER, M., P. MILGROM, AND G. SCHWARZ (2022): "Taming the communication and computation complexity of combinatorial auctions: The FUEL bid language," *Management Science*.

BIKHCHANDANI, S. AND J. W. MAMER (1997): "Competitive equilibrium in an exchange economy with indivisibilities," *Journal of Economic Theory*, 74, 385–413.

Boyd, S. P. and L. Vandenberghe (2004): *Convex optimization*, Cambridge University Press.

Budish, E. and P. J. Reny (2020): "An improved bound for the Shapley–Folkman theorem," *Journal of Mathematical Economics*, 89, 48–52.

Chavas, J.-P. and W. Briec (2012): "On economic efficiency under non-convexity," *Economic Theory*, 50, 671–701.

Chen, X., D. Dai, Y. Du, and S.-H. Teng (2009): "Settling the complexity of Arrow-Debreu equilibria in markets with additively separable utilities," in *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, IEEE, 273–282.

Danilov, V., G. Koshevoy, and K. Murota (2001): "Discrete convexity and equilibria in economies with indivisible goods and money," *Mathematical Social Sciences*, 41, 251–273.

Daskalakis, C., P. W. Goldberg, and C. H. Papadimitriou (2009): "The complexity of computing a Nash equilibrium," *SIAM Journal on Computing*, 39, 195–259.

Debreu, G. (1951): "The coefficient of resource utilization," *Econometrica*, 19, 273–292.

Farrell, M. J. (1959): "The convexity assumption in the theory of competitive markets," *Journal of Political Economy*, 67, 377–391.

Feldman, M., G. Shabtai, and A. Wolfenfeld (2021): "Two-Price Equilibrium," *arXiv preprint arXiv:2112.08215*.

Goeree, J. K. (2023): "Yquilibrium: A Theory for (Non-) Convex Economies," *Working paper*.

Gul, F. and E. Stacchetti (1999): "Walrasian equilibrium with gross substitutes," *Journal of Economic Theory*, 87, 95–124.

Hatfield, J. W., S. D. Kominers, A. Nichifor, M. Ostrovsky, and A. Westkamp (2013): "Stability and competitive equilibrium in trading networks," *Journal of Political Economy*, 121, 966–1005.

Hayek, F. (1945): "The Use of Knowledge in Society," *The American Economic Review*, 35, 519–530.

Heller, W. P. (1972): "Transactions with set-up costs," *Journal of Economic Theory*, 4, 465 – 478.

Jackson, M. O. (1992): "Incentive compatibility and competitive allocations," *Economics Letters*, 40, 299–302.

Knuth, D. E. (1976): "Big omicron and big omega and big theta," *ACM Sigact News*, 8, 18–24.

Koopmans, T. C. (1961): "Convexity assumptions, allocative efficiency, and competitive equilibrium," *Journal of Political Economy*, 69, 478–479.

Liberopoulos, G. and P. Andrianesis (2016): "Critical review of pricing schemes in markets with non-convex costs," *Operations Research*, 64, 17–31.

Milgrom, P. (2009): "Assignment messages and exchanges," *American Economic Journal: Microeconomics*, 1, 95–113.

Milgrom, P. and I. Segal (2002): "Envelope theorems for arbitrary choice sets," *Econometrica*, 70, 583–601.

Milgrom, P. and B. Strulovici (2009): "Substitute goods, auctions, and equilibrium," *Journal of Economic Theory*, 144, 212–247.

Morimoto, S. and S. Serizawa (2015): "Strategy-proofness and efficiency with non-quasi-linear preferences: A characterization of minimum price Walrasian rule," *Theoretical Economics*, 10, 445–487.

Nguyen, T. and R. Vohra (2024): "(Near) Substitute Preferences and Equilibria with Indivisibilities," *Forthcoming at Journal of Political Economy.*

Nisan, N. and I. Segal (2006): "The communication requirements of efficient allocations and supporting prices," *Journal of Economic Theory*, 129, 192–224.

Roberts, D. J. and A. Postlewaite (1976): "The incentives for price-taking behavior in large exchange economies," *Econometrica*, 44, 115–127.

Rockafellar, R. T. and R. J.-B. Wets (2009): *Variational analysis*, vol. 317, Springer Science & Business Media.

Rothenberg, J. (1960): "Non-convexity, aggregation, and Pareto optimality," *Journal of Political Economy*, 68, 435–468.

Scarf, H. (1967): "The approximation of fixed points of a continuous mapping," *SIAM Journal on Applied Mathematics*, 15, 1328–1343.

——— (1994): "The allocation of resources in the presence of indivisibilities," *Journal of Economic Perspectives*, 8, 111–128.

Segal, I. (2007): "The communication requirements of social choice rules and supporting budget sets," *Journal of Economic Theory*, 136, 341–378.

Starr, R. M. (1969): "Quasi-equilibria in markets with non-convex preferences," *Econometrica*, 37, 25–38.

Sun, N. and Z. Yang (2006): "Equilibria and indivisibilities: Gross substitutes and complements," *Econometrica*, 74, 1385–1402.

Watt, M. (2022): "Strong monotonicity and perturbation-proofness of exchange economies," *Working paper.*

Wilson, R. B. (1993): *Nonlinear pricing*, Oxford University Press.

A Proofs omitted from the main text

A.1 Proof of Proposition 3

Proof. We offer two proofs of this claim: the first directly from the definitions and the second highlights the relationship with the envelope theorem.

For the first proof, let $x' \in D_n(p')$. We have that

$$\mathscr{R}_n(p', x) = u_n(x') - p' \cdot x' - (u_n(x) - p' \cdot x).$$

But since $x \in D_n(p)$, we have $u(x') - p \cdot x' \leq u(x) - p \cdot x$, so that

$$\mathscr{R}_n(p', x) \leq p \cdot x' - p \cdot x + p' \cdot x - p' \cdot x = (p - p') \cdot (x' - x),$$

which is $O(\|p - p'\|)$ since $x', x \in X$, a compact set.

For the second proof, write

$$\begin{aligned}\mathscr{R}_n(p', x) &= \hat{u}_n(p') - (u_n(x) - p' \cdot x) \\ &= \hat{u}_n(p') - (u_n(x) - p \cdot x) - p \cdot x + p' \cdot x \\ &= \hat{u}_n(p') - \hat{u}_n(p) + (p' - p) \cdot x.\end{aligned}$$

Now let $p(t) = (1 - t)p + tp'$ for $t \in [0, 1]$ and apply the Milgrom and Segal (2002) envelope theorem to the parametrized utility maximization problem

$$\hat{u}_n(p(t)) = \max_{x \in X} u_n(x) - p(t) \cdot x,$$

to give

$$\hat{u}_n(p') = \hat{u}_n(p) - \int_0^1 (p' - p) \cdot d(t) dt.$$

for selections $d(t) \in D_n(p(t))$. Substituting into the expression for $\mathscr{R}_n(p', x)$ above, we obtain

$$\mathscr{R}_n(p', x) = -\int_0^1 (p' - p) \cdot d(t) dt + (p' - p) \cdot x = \int_0^1 (p' - p) \cdot (x - d(t)) dt$$

which is bounded above by $(p' - p) \cdot (x - x')$, the same expression as before, since $(p' - p) \cdot (x - d(t))$ is increasing in t by the law of demand. □

A.2 Proof of Theorem 2

Part (a) Fix some $\mathscr{E}$ and consider any sequence $\alpha_i \to \alpha$ and selections R_i of revenues $\sum_{n \in N}(1+\alpha_i)p^{\alpha_i} \cdot x_n^{\alpha_i} - \sum_{f \in F} p^{\alpha_i} \cdot y_f^{\alpha_i}$ associated with some markup mechanisms $(\alpha_i, p^{\alpha_i}, \mathbf{x}^{\alpha_i})$ constructed as in Definition 4.2. We will show that $\lim_i R_i$ is obtainable as the revenue of some markup mechanism $(\alpha, p^\alpha, \mathbf{x}^\alpha)$ so that the infimum in equation (A) is attained (and thus the minimum exists).

By the saddle point condition associated with the objective in Definition 4.2, we have that there are some $\tilde{\mathbf{x}}^{\alpha_i}$ maximizing over $\mathrm{co}(\mathbf{X})$ the objective $\sum_{n \in N} \frac{\mathrm{cav}(u_n)(x_n)}{1+\alpha_i} - \sum_{f \in F} \mathrm{vex}(c_f)(y_f)$. As $\alpha_i \to \alpha$, this objective hypo-converges[29] (since it is continuous and bounded) to $\sum_{n \in N} \frac{\mathrm{cav}(u_n)(x_n)}{1+\alpha} - \sum_{f \in F} \mathrm{vex}(c_f)(y_f)$, so that $\tilde{\mathbf{x}}^{\alpha_i} \to \tilde{\mathbf{x}}^\alpha$ for some $\tilde{\mathbf{x}}^\alpha$ that maximizes this latter objective. By optimality, each p^{α_i} lies in the superdifferential ∂^* of the concavified valuation functions of each buyer and the subdifferential of the convexified cost functions of each seller at $\tilde{\mathbf{x}}^{\alpha_i}$. For these concave / convex functions, the super- and subdifferential correspondences are upper hemicontinuous, so that the sequence of p^{α_i} must converge to some p^α in the super- and subdifferentials at $\tilde{\mathbf{x}}^\alpha$. Finally, since the demand and supply correspondences are upper hemicontinuous, the convergence of prices implies that $\mathbf{x}^{\alpha_i}$ must approach some $\mathbf{x}^\alpha$ such that $x_n^\alpha \in D_n((1+\alpha)p^\alpha)$ and $y_f^\alpha \in S_f(p^\alpha)$. Thus the limit of R_i is attained as the revenue of some markup mechanism $(\alpha, p^\alpha, \mathbf{x}^\alpha)$ as required.

Part (b) For notational simplicity, we drop the index for t in the prices, premiums and allocations.

The construction in Definition 4.2 ensures, via Proposition 2, that $\sum_{f \in F} y_f^\alpha - \sum_{n \in N} x_n^\alpha \leq (2R)1_L$. Thus, it suffices to show that for sufficiently large $|N_t|$, there is an α such that

$$\alpha \sum_{n \in N} p^\alpha \cdot x_n^\alpha \geq (2R)p^\alpha \cdot 1_L.$$

Moreover, if this $\alpha = O\left(\frac{1}{|N_t|}\right)$, then since $\alpha^* < \alpha$, (b) will follow. To arrive at this result, we will show that for fixed $\alpha > 0$ close enough to zero, $\sum_{n \in N} p^\alpha \cdot x_n^\alpha$ is $\Omega(|N_t|)$, while $p^\alpha \cdot 1_L$ is $O(1)$.

Let $\mathcal{S}^\alpha$ be the value of the saddle point problem

$$\min_{p \in \mathbb{R}_+^L} \max_{\mathbf{x} \in \mathbf{X}} \sum_{n \in N} \frac{\mathrm{cav}(u_n)(x_n)}{1+\alpha} - \sum_{f \in F} \mathrm{vex}(c_f)(y_f) - p \cdot \left(\sum_{n \in N} x_n + R1_L - \sum_{f \in F} y_F \right).$$

[29] See Rockafellar and Wets (2009), Section 7.B.

First, note that $\mathcal{S}^\alpha$ is $\Theta(|N_t|)$ for sufficiently small α by Assumption 3.[30] To see this, denote by $f(|N_t|) = \sum_{n \in N_t} u_n(x_n^*)$ and $g(|N_t|) = \sum_{f \in F_t} c_f(y_f^*)$. Assumption 3 and the boundedness of utilities and costs implies that the efficient surplus is $\Theta(|N_t|)$. As a result, $\liminf_{N\to\infty} \frac{f(N)}{N} = u > 0$, say, and $\limsup_{N\to\infty} \frac{g(N)}{N} = c > 0$, with $u - c > 0$. Then $\mathcal{S}^\alpha \geq \liminf_{N\to\infty} \frac{f(N)}{(1+\alpha)N} - g(N) = \frac{u}{1+\alpha} - c$, which is positive for sufficiently small α.

Now we show that this implies $\sum_{n\in N} p^\alpha \cdot x_n^\alpha$ is $\Omega(|N_t|)$ for small, fixed α. To see this, note that since $\sum_{f\in F_t} c_f(y_f^\alpha)$ is $\Omega(|N_t|)$, individual rationality of the sellers (in the perturbed economy) implies that $\sum_{f\in F_t} p^\alpha \cdot y_f^\alpha$ is $\Omega(|N_t|)$. But then by complementary slackness $\sum_{n\in N_t} p^\alpha \cdot x_n^\alpha = \sum_{f\in F_t} p^\alpha \cdot y_f^\alpha - Rp^\alpha \cdot 1_L$, and since Assumption 4 implies $\|p\| \leq M$, we must have that $\sum_{n\in N} p^\alpha \cdot x_n^\alpha$ is $\Omega(|N_t|)$.

Since for α near zero, $\sum_{n\in N} p^\alpha \cdot x_n^\alpha$ is $\Omega(|N_t|)$ and $(2R)p^\alpha \cdot 1_L$ is $O(1)$ (where R is $O(1)$ by assumption 2), for sufficiently large $|N_t|$, there is some α (and thus some least α by (a)) such that

$$\alpha \sum_{n\in N} p^\alpha \cdot x_n^\alpha \geq (2R)p^\alpha \cdot 1_L,$$

and furthermore, this α is $O\left(\frac{1}{|N_t|}\right)$. Since $\alpha^* < \alpha$, we have that α^* is $O\left(\frac{1}{|N_t|}\right)$, as required.

Part (c) We now apply the First Welfare Bound to show that the allocation $\mathbf{x}^{\alpha^*}$ is approximately efficient. In order to satisfy the assumptions on prices in Theorem 1, we imagine $\mathbf{x}^{\alpha^*}$ was implemented with a single price vector p^{α^*} and (therefore) a budget deficit. Theorem 1 tells us that

$$\mathcal{S}(\mathbf{x}^*) - \mathcal{S}(\mathbf{x}^{\alpha^*}) \leq \mathcal{R}(p^{\alpha^*}, \mathbf{x}^{\alpha^*}) + p^{\alpha^*} \cdot \left(\sum_{f\in F_t} y_f^{\alpha^*} - \sum_{n\in N_t} x_n^{\alpha^*} \right). \tag{1}$$

By construction, in $\mathbf{x}^{\alpha^*}$ at prices p^{α^*}, no sellers are rationed, while at prices $(1+\alpha^*)p^{\alpha^*}$, no buyers are rationed. But (1) requires the rationing of buyers at price p^{α^*}. For this, we use Proposition 3: since α^* is $O\left(\frac{1}{|N_t|}\right)$ and $\|p^{\alpha^*}\|$ is bounded (by Assumption 4), this implies that $\mathcal{R}_n(p^{\alpha^*}, x_n^{\alpha^*})$ is $O\left(\frac{1}{|N_t|}\right)$, and so $\mathcal{R}(p^{\alpha^*}, \mathbf{x}^{\alpha^*}) = \sum_{n\in N_t} \mathcal{R}_n(p^{\alpha^*}, x_n^{\alpha^*})$ is $O(1)$.

Finally, we note that the budget deficit (the second term on the right of Equation (1)) is $O(1)$ since the excess supply is bounded by construction and each component of p^{α^*} is $O(1)$. Thus Theorem 2 follows.

[30] Recall that $f(x) = \Theta(g(x))$ if $f(x) = O(g(x))$ and $f(x) = \Omega(g(x))$

A.3 Proof of Theorem 3

Part (a) This follows by simply noting that all markup mechanisms are envy-free, and so Theorem 1 of Azevedo and Budish (2019) implies the result.

Parts (b) and (c) First, we formally define the expected supply and demand correspondences. Given distribution ν on $\mathscr{V}$, The expected indirect utility function is defined pointwise for $p \in \mathscr{P}$ by

$$\mathbb{E}_\nu[\hat{u}(p)] = \int_{\mathscr{V}} \hat{u}_n(p) d\nu(u_n),$$

and similarly the expected indirect profit function is

$$\mathbb{E}_\chi[\hat{\pi}(p)] = \int_{\mathscr{C}} \hat{\pi}_f(p) d\chi(\pi_f).$$

The expected demand correspondence is then $\mathbb{E}_\nu[D(p)] = -\partial \mathbb{E}_\nu[u(p)]$ and the expected supply correspondence is $\mathbb{E}_\chi[S(p)] = \partial \mathbb{E}_\chi[\pi(p)]$.[31]

Parts (b) and (c) follow directly from the corresponding theorems for the Walrasian mechanism—namely, Theorems 1, 3 and 4 adapted as in Appendix C of Watt (2022). To see this, we show that the objective for the markup mechanism—both under truthful reporting and after misreporting by a single agent—differs from the objective of the Walrasian mechanism for the convexified economy by a $O(1)$−Lipschitz convex function, which constitutes a perturbation under the definition in Watt (2022). Suppose that under truthful reporting, the mechanism chooses some markup $\alpha \geq 0$ and that under a misreport by an agent, the mechanism chooses markup $\alpha' \geq 0$. The markup mechanism (seller) prices p^α under truthful reporting minimize the dual objective

$$\frac{1}{1+\alpha} \sum_{n \in N} \hat{u}_n(p) + \sum_{f \in F} \hat{\pi}_f(p) - p \cdot R1_L,$$

while under the misreport, the (seller) price vector $p^{\alpha'}$ minimizes

$$\frac{1}{1+\alpha'} \sum_{n \in N} \hat{u}_n(p) + \sum_{f \in F} \hat{\pi}_f(p) - p \cdot R1_L.$$

Note that these objectives do not have cav or vex in them since the indirect utility and profit functions are the same for the original valuations and costs as their concavified/convexified versions.

[31] The expected demand and supply correspondences can also be defined using the set-valued integral of Aumann (1965), but we refer the reader to Watt (2022) for details on this construction.

The Walrasian mechanism for the convexified economy minimizes

$$\sum_{n\in N} \hat{u}_n(p) + \sum_{f\in F} \hat{\pi}_f(p).$$

This objective and the $\alpha-$objective differ by

$$\frac{\alpha}{1+\alpha} \sum_{n\in N} \hat{u}_n(p) - p \cdot R1_L.$$

Since $\alpha \leq O(1/N)$, we have $\frac{\alpha}{1+\alpha} \leq O(1/N)$ as well, while $\sum_{n\in N} \hat{u}_n(p)$ is $O(N)-$Lipschitz since its subdifferential is total demand at p which is $O(N)$ (and the Lipschitz constant is the largest selection from the subdifferential). This implies that the perturbation above is $O(1)$-Lipschitz. A similar analysis applies for the α'-objective.

B Nonconvexity and approximate equilibria

We begin with a slightly stronger statement of the Shapley-Folkman Lemma that is used in general equilibrium theorem with nonconvexities.

Proposition 4. *Let $S_i \subseteq \mathbb{R}^L$ for $i = 1, ..., N$, $S = \bigoplus_{i=1}^N S_i$ and $L' = \min(L, N)$. Then for any $x \in \mathrm{co}(S)$:*

(a) (Shapley-Folkman Lemma) $x = \sum_{i=1}^N x_i$ where $x_i \in \mathrm{co}(S_i)$ and $|i : x_i \in \mathrm{co}(S_i) \setminus S_i| \leq L'$.

(b) (Starr, 1969) If S_i is ordered so that $r(S_i)$ is nonincreasing in i, then there is $y \in S$ such that $|x - y| \leq \sqrt{\sum_{i=1}^{L'} r(S_i)^2}$.

(c) (Heller, 1972) If S_i is ordered so that $\rho(S_i)$ is nonincreasing in i, then there is $y \in S$ such that $|x - y| \leq \sqrt{\sum_{i=1}^{L'} \rho(S_i)^2}$.

These results have been used in the general equilibrium context to obtain *approximate equilibria*, which are price-allocation pairs $(p, \mathbf{x})$ such that $x_n \in D_n(p)$ for all n, $y_f \in S_f(p)$ for all f but $\left|\sum_{n\in N} x_n - \sum_{f\in F} y_f\right| \leq s$ for some $s > 0$. In particular, the allocation associated with an approximate equilibrium may have excess demand and therefore be infeasible. The approximate equilibrium is obtained by identifying the competitive equilibrium associated with a convexified version of the economy (in which each agent's upper contour set is replaced by its convex hull) and applying the results of Proposition 4 to the resulting allocation. The approximate equilibrium analogues of Proposition 4 are contained in Proposition 5 below.

Proposition 5. *For economy $\mathscr{E} = (N, F)$:*

(a) *([Starr, 1969]) There is $\mathbf{x} \in \operatorname{co}(\mathbf{X})$ and $p \in \mathbb{R}^L_+, p > 0$ such that $|n : x_n \in \operatorname{co}(X)| + |f : y_f \in \operatorname{co}(Y)| \leq L$ and for all other agents, $x_n \in D_n(p)$ and $y_f \in S_f(p)$.*

(b) *Let $r_n = \sup_{\bar{u} \in \mathbb{R}} r(UC^{\bar{u}}_n)$ and $r_f = \sup_{\bar{\pi} \in \mathbb{R}} r(UC^{\bar{\pi}}_f)$. Let $\delta \geq 0$ satisfy $r_n \leq \delta$ for all $n \in N$ and $r_f \leq \delta$ for all $f \in F$. Then there exists $p \in \mathbb{R}^L_+, p > 0$, $x_n \in X$ and $y_f \in Y$ such that $x_n \in D_n(p)$ for all $n \in N$, $y_f \in S_f(p)$ for all $f \in F$ and $\left|\sum_{n \in N} x_n - \sum_{f \in F} y_f\right| \leq \delta\sqrt{L}$.*

(c) *Let $\rho_n = \sup_{\bar{u} \in \mathbb{R}} \rho(UC^{\bar{u}}_n)$ and $\rho_f = \sup_{\bar{\pi} \in \mathbb{R}} \rho(UC^{\bar{\pi}}_f)$. Let $\delta' \geq 0$ satisfy $\rho_n \leq \delta'$ for all $n \in N$ and $\rho_f \leq \delta'$ for all $f \in F$. Then there exists $p \in \mathbb{R}^L_+, p > 0$, $x_n \in X$ and $y_f \in Y$ such that $x_n \in D_n(p)$ for all $n \in N$, $y_f \in S_f(p)$ for all $f \in F$ and $\left|\sum_{n \in N} x_n - \sum_{f \in F} y_f\right| \leq \delta' L$.*

The statement of Proposition 5(a) is standard, but the statements of Proposition 5(b) and (c) are stronger than the classical statement due to Starr (1969) and Heller (1972). Again, the quasilinearity of agent preferences allows us to conclude that agents are utility- and profit-maximizing, rather than just expenditure-minimizing.

Finally, we introduce a more general class of quasilinear preferences to which many of our results also apply and which offer a meaningful interpretation in terms of perceived complementarity and substitutability of goods. Nguyen and Vohra (2024) introduced the concept of the *generalized* Δ*-single improvement property*, which is a generalization of the well-known single improvement property.

Definition B.1. The preferences of buyer $n \in N$ satisfies the generalized Δ-single improvement property (or satisfy Δ−substitutes) for some $\Delta > 0$ if for any price vector $p > 0$, any two bundles $x, y \in D_n(p)$ and any price change δp such that $\delta p \cdot x > \delta p \cdot y$, there exist $a \leq (x - y)^+$ and $b \leq (y - x)^+$ such that:

(a) $|a| + |b| \leq \Delta$

(b) $\delta p \cdot a > \delta p \cdot b$, and

(c) $x - a + b \in D_n(p)$.

Here $(x - y)^+$ denotes the vector whose ℓ^{th} component is $\max(x_\ell - y_\ell, 0)$.

The Δ in this definition captures a measure of the substitutability and complementarity between goods. Preferences with the single improvement property of Gul and Stacchetti (1999) are contained in the class with $\Delta = 2$.

By our assumption on the compactness of X and Y, all preferences and technologies satisfy the general

Δ-improvement property for some Δ (as noted by Nguyen and Vohra (2024)). But the following stronger relationship between the inner radii of preferences and the Δ-single improvement property also holds.

Proposition 6. *Let $r_n = \sup_{\bar{u}\in\mathbb{R}} r(UC_n^{\bar{u}})$. Then the preferences of buyer $n \in N$ satisfy the generalized $\Delta-$single improvement property for all $\Delta > 2\sqrt{2}r_n$.*

Proof. Let the preferences of buyer n satisfy $r_n = \sup_{\bar{u}\in\mathbb{R}} r(UC_n^{\bar{u}})$. Let $x, y \in X$ and $p \in \mathbb{R}_+^L$ be given such that $x, y \in D_n(p)$. Suppose $|(x-y)^+| + |(y-x)^+| \geq 2r_n$ (else the preferences immediately satisfy the Δ improvement property for $\Delta = 2r_n$).

For any $\epsilon > 0$, let $z \in \mathbb{R}_+^L$ be the unique convex combination of x and y such that $|x - z| = r_n + \epsilon$ and write $z = \lambda x + (1-\lambda)y$. By construction $(z, p\cdot z) \in \mathrm{co}(UC_n^{u_n(x)-p\cdot x})$. Then by the bound on the nonconvexity of the preferences, there is a set $T \subseteq UC_n^{u(x)-p\cdot x}$ with $\mathrm{rad}(T) \leq r_n$ such that $(z, p\cdot z) = \sum_{(x',t')\in T} \alpha_{(x',t')}(x', t')$ where $\sum_{(x',t')\in T} \alpha_{(x',t')} = 1$.

We now argue that for all $(x', t') \in T$, $x' \in D_n(p)$ and $t' = p\cdot x'$. To see this, note that $x \in D_n(p)$ implies $u_n(x') - p\cdot x' \leq u_n(x) - p\cdot x$. Summing, we have

$$\begin{aligned} u_n(x) - p\cdot x &\geq \sum_{(x',t')\in T} \alpha_{(x',t')}[u_n(x') - p\cdot x'] \\ &= \sum_{(x',t')\in T} \alpha_{(x',t')} u_n(x') - p\cdot z \\ &= \sum_{(x',t')\in T} \alpha_{(x',t')}[u_n(x') - t'] \end{aligned}$$

On the other hand, since $(x', t') \in UC_n^{u(x)-p\cdot x}$ we have $u_n(x') - t' \geq u(x) - p\cdot x$. The only way these can simultaneously hold is if $u_n(x') - t' = u(x) - p\cdot x$ for all $(x', t') \in T$.

However, we then have $\sum_{(x',t')\in T} \alpha_{(x',t')}[u_n(x') - p\cdot x'] = u(x) - p\cdot x$. This implies that at least one of $u_n(x') - p\cdot x' \geq u_n(x) - p\cdot x$. But then $x \in D_n(p)$ implies that $u_n(x') - p\cdot x' = u_n(x) - p\cdot x$ for all x', so $x' \in D_n(p)$.

By construction, $|x - x'| \leq 2r_n + \epsilon$. But then $||x - x'||_1 \leq 2\sqrt{2}r_n + \epsilon$ as well. $\square$

Clearly, the generalized Δ-single property can be readily extended to sellers, by replacing the expressions for utility with those for profits, and an analogue of Proposition 6 also holds.

Nguyen and Vohra (2024) demonstrate the following approximate equilibrium result in a setting with indivisibilities (so that $X \subseteq \mathbb{Z}^L_+$ and $Y \subseteq \mathbb{Z}^L_+$).

Proposition 7. *Suppose all buyers' preferences and sellers' technologies satisfy the generalized Δ-improvement property and that $X \subseteq \mathbb{Z}^L_+$ and $Y \subseteq \mathbb{Z}^L_+$. Then there exists $p \in \mathbb{R}^L_+, p > 0$, $x_n \in X$ and $y_f \in Y$ such that $x_n \in D_n(p)$ for all $n \in N$, $y_f \in S_n(p)$ and for each $\ell \in L$, $\left|\sum_{n \in N} x_{nl} - \sum_{f \in F} y_{fl}\right| \leq \Delta - 1$.*

Note that the concept of approximate equilibrium in this result is somewhat stronger than the previous results since the maximum imbalance in supply and demand is bounded good-by-good, rather than in terms of Euclidean distance in commodity space. However, depending on the relative size of Δ, the inner radii of nonconvexity and the breadths of nonconvexity of preferences, any of the approximate equilibrium bounds in Proposition 5(b), 5(c) or 7 may be strongest for our purposes.