Do Prices Coordinate Markets?

Justin Hsu[*]   Jamie Morgenstern[†]   Ryan Rogers[‡]   Aaron Roth[§]   Rakesh Vohra[¶]

November 11, 2015

Abstract

Walrasian equilibrium prices have a remarkable property: they allow each buyer to purchase a bundle of goods that she finds the most desirable, while guaranteeing that the induced allocation over all buyers will globally maximize social welfare. However, this clean story has two important caveats:

- First, the prices may induce indifferences—in fact, the minimal equilibrium prices *necessarily* induce indifferences. In general, buyers may need to coordinate with one another to resolve these indifferences, so the prices alone are *not* sufficient to coordinate the market.

- Second, although we know natural procedures which converge to Walrasian equilibrium prices on a fixed population, in practice buyers typically observe prices without participating in a tâtonnement process. These prices cannot be perfect Walrasian equilibrium prices, but instead somehow reflect distributional information about the market.

To better understand the performance of Walrasian prices when facing these two problems, we give results of two sorts. First, we show a mild genericity condition on valuations under which the minimal Walrasian equilibrium prices induce allocations which result in low over-demand for arbitrary (even adversarial) tie-breaking by buyers. In fact, our results show that the over-demand of any good can be bounded by 1, which is the best possible. We demonstrate our results in the unit demand setting and give an extension to the class of Matroid Based Valuations (MBV), conjectured to be equal to the class of Gross Substitute valuations (GS).

Second, we use techniques from learning theory to argue that the over-demand and welfare induced by a price vector converges to its expectation uniformly over the class of all price vectors, with sample complexity only linear in the number of goods in the market in the former case and quadratic in the number of goods in the latter case. These results make no assumption on the form of the valuation functions of the buyers.

Combining these two results implies that under a mild genericity condition, the exact Walrasian equilibrium prices computed in a market are guaranteed to induce both low over-demand and high welfare when used in a new market, in which agents are sampled independently from the same distribution, whenever the number of agents is larger than the number of commodities in the market.

---

[*]Department of Computer and Information Science, University of Pennsylvania. Email: `justhsu@cis.upenn.edu`

[†]Warren Center Postdoctoral Fellow, Departments of Computer and Information Science and Economics, University of Pennsylvania. Email: `jamiemmt@cs.cmu.edu`

[‡]Department of Applied Mathematics and Computational Science, University of Pennsylvania. Email: `ryrogers@sas.upenn.edu`

[§]Department of Computer and Information Sciences, University of Pennsylvania. Email: `aaroth@cis.upenn.edu` Supported in part by an NSF grant CCF-1101389 and an Alfred P. Sloan Research Fellowship.

[¶]Economics Department, University of Pennsylvania. Email: `rvohra@seas.upenn.edu`

1 Introduction

The existence of Walrasian equilibrium is often expressed as a pithy slogan: prices coordinate markets. However, this is not exactly true—a Walrasian equilibrium specifies a price for each good *and* an assignment of goods to buyers. The assignment is just as important as the prices, since there can be multiple bundles of goods which maximize a buyer's utility at given prices. If buyers select arbitrarily among these bundles, they may over-demand some goods.

One way to avoid this problem is by assuming that each buyer's valuation function is strictly concave, ensuring that the utility maximizing bundle of goods is unique at any set of prices. However, this idea does not apply in many economic settings; for instance, when goods are indivisible. Here, prices (and, in particular, Walrasian equilibrium prices) can induce indifferences.

For another possible solution, one might hope that a market maker could select Walrasian prices that eliminate the coordination problem. For example, with a single good and distinct valuations, a price *strictly* between the largest and second largest buyer valuation will eliminate indifferences and over-demand. However, relying on a principal with full knowledge of the market to solve the coordination problem seems to defeat the purpose of having markets in the first place.

In this work, we consider a third approach: natural assumptions on buyers to simplify the coordination problem. We start our investigation from *minimal* Walrasian equilibrium prices, the prices that result from many natural market dynamics[1] (also known as tâtonnment procedures) [28]. We consider allocating bundles of m types of indivisible goods g, each with some supply s_g, to n buyers who have matroid based valuations[2] and quasi-linear preferences; the assignment ("unit demand") model is an important special case. We present details of our model in Section 2.

We begin our technical results in Section 3, by showing that indifferences at the minimal Walrasian prices can be a serious problem—goods can be in the demand correspondence of *every* buyer. Clever tie-breaking does not help; for any tie-breaking rule, the induced over-demand can be as large as $\Omega(n)$. Even worse, we observe that the minimal Walrasian equilibrium prices induce indifferences in *any* market, for any set of buyers.

Hence, we cannot hope to rule out all over-demand. But, we give a "genericity" condition on buyer valuations that is in some sense the next-best thing: the over-demand for each good g will be at most 1, independent of its supply s_g and the tie-breaking strategy used by buyers. Therefore, as the supply grows large, worst-case over-demand becomes negligible. We warm up with the assignment model in Section 4, where our condition is simple to state: buyer valuations for goods should be linearly independent over the coefficients $\{-1, 0, 1\}$. We generalize our techniques to the matroid based valuations case in Section 5; the situation is significantly more complicated, but the core genericity definition is in the same spirit.

After we show that exact minimal Walrasian prices "generically" induce low over-demand, a natural question is whether this property still holds when the same prices are used on new buyers; we turn to this question in Section 6. More formally, imagine a sample N_1 of n buyers is drawn from an unknown distribution Π of buyer valuations. The goods are priced using the minimal Walrasian equilibrium prices, computed from the valuations of buyers in N_1. Now, keeping these prices fixed, we draw a fresh sample N_2 of n buyers from Π, who each choose some bundle from their demand correspondence at the given prices (breaking ties arbitrarily). Will the over-demand and welfare

---

[1]Minimal Walrasian prices are focal in other ways: e.g., in matching markets they correspond to VCG prices.

[2]"Matroid Based Valuations" are a structured subclass of *Gross Substitutes* valuations. In fact, Ostrovsky and Paes Leme [37] conjecture that the class of matroid based valuations is *equal* to the class of gross substitutes valuations, the largest class of valuations for which Walrasian equilibrium prices are guaranteed to exist.

on N_2 be close to the over-demand and welfare on N_1? If the supply s_g of a good g is small, the difference in over-demand between N_1 and N_2 may be large when compared to s_g. However, we show that if $s_g \geq \widetilde{O}(m/\epsilon^2)$, the demand for any good g on sample N_2 will be within a $1 \pm \epsilon$ factor of the supply s_g of good g. Note that the supply requirement is independent of the market size n. Similarly, if the optimal welfare for N_1, N_2 is large enough, the welfare of the two markets at these prices will be within a $1 \pm \epsilon$ of one another (and within a $1 - \epsilon$ factor of the optimal welfare for N_2).

Notably, we are able to prove these bounds *without* any assumption on the structure of the valuation functions. This lack of structure makes it difficult to argue directly about notions of combinatorial dimension like VC-dimension, and so we take a different approach which may be of independent interest. Using a recent compression argument of Daniely and Shalev-Shwartz [13], we show that, assuming a fixed but unknown set of prices, the class of functions *predicting* a buyer's demanded bundle at those prices is learnable using $\widetilde{O}(m/\epsilon^2)$ many samples. Because this is a multi-class learning problem, learning does not imply uniform convergence. However, the binary problem of predicting demand for a *particular* good is a 1-dimensional projection of the bundle prediction problem, and hence is also learnable with the same number of samples. Finally, by a classical result of Ehrenfeucht et al. [19], learning and uniform convergence have the same sample complexity in binary prediction problems. So, we can bound the VC dimension, and thus the sample complexity needed to obtain uniform convergence for demand. Moreover, our bound is tight—even for unit demand buyers, the VC-dimension of the class of demand predictors is $\Omega(m)$.

Welfare, unlike demand, corresponds to a real-valued prediction problem, and so the sample complexity needed for uniform convergence cannot be bounded by bounding the sample complexity of learning. Instead, we directly bound the pseudo-dimension of the class of welfare predictors by $\widetilde{O}(m^2)$, again without making any assumptions about the form of the valuation functions. We show that if the optimal welfare is $\widetilde{\Omega}\left(m^4/\epsilon^2\right)$, the welfare induced by the Walrasian prices p for N_1 when applied to N_2 is within a $1 - \epsilon$ factor of optimal.

Related work. This paper follows a long line of work on understanding how markets behave under limited coordination. If buyers' valuations are strictly concave and items are divisible, Arrow and Debreu [2] show there exist item prices p such that each agent has a unique utility-maximizing bundle at p, and that when each agent selects her unique utility-maximizing bundle, the market clears. With indivisible items, anonymous equilibrium item pricings may not exist (and finding such prices when they do is NP-hard [14]). Mount and Reiter [32] consider the size of the message space needed to compute Walrasian equilibria, and Nisan and Segal [36] show that polynomial communication is sufficient. This stands in contrast to the situation for submodular buyers, where exponential communication is needed to compute prices which support an efficient allocation [36].

Our work is also related to the blossoming area of learning for mechanism design, where a mechanism is selected from some class of mechanisms as a function of sampled buyers. Recent work has measured the sample complexity of revenue maximization in single-parameter settings [3–6, 9–12, 17, 20, 22, 26, 27, 30, 31, 40] and multi-parameter settings [15, 18].

In particular, sample complexity results for pricing are known. In the unlimited supply setting (in which "over-demand" cannot arise), Balcan et al. [6] show how to learn approximately revenue-optimal prices with polynomial sample complexity using a covering argument. In Balcan et al. [5], the authors show how to extend this work to large but limited supply settings and to welfare maximization, although their guarantees only hold for finite classes of mechanisms. Devanur and Hayes [16] also consider prior-independent pricing for revenue maximization with unit demand

buyers in an online setting. These last two papers consider buyers that make decisions sequentially, avoiding the issue of over-demand from uncoordinated resolution of indifferences. Finally, we use compression schemes to derive several uniform convergence results; this compression tool was first used in the context of game theory by Balcan et al. [7], who use it directly to upper bound the PAC complexity of a learning problem rather than to imply uniform convergence over a class.

We are aware of less work related to our genericity results, but we point out that our swap graph construction bears a resemblance to the exchangability graph of Murota [34]. The exchangability graph also has nodes defined by goods, but is defined without considering a Walrasian allocation and pricing. In contrast, our swap graph is designed to model indifferences at equilibrium.

2 Model

We consider a market with m indivisible goods, where good g has supply $s_g \geq 1$. We will write the set of goods as $[m] = \{1, \ldots, m\}$ and denote the bundles of the m goods as $\mathcal{G} = 2^{[m]}$. The market will also have a set N of n buyers, where each buyer demands at most one copy of each good. For simplicity, we consider valuations defined over subsets of goods, rather than arbitrary sets of copies. See the extended version for a formal treatment. For each buyer $q \in N$, let $v_q : \mathcal{G} \to [0, H]$ denote q's valuation function over bundles of goods. We will assume v_q is monotone and $v_q(\emptyset) = 0$. We now specify the set of feasible assignments of goods to buyers, called allocations.

Definition 2.1 (Allocation). An *allocation* $\mu : N \to \mathcal{G}$ is a function that assigns each buyer a bundle such that the whole assignment is feasible: $\sum_{q \in N} \mathbb{1}\{g \in \mu(q)\} \leq s_g$ for all $g \in [m]$.

As is typical, we consider quasi-linear utility functions $u_q : \mathcal{G} \times \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$, defined by $u_q(S; p) = v_q(S) - \sum_{g \in S} p_g$ for all $q \in N$ and $S \in \mathcal{G}$. We will consider prices defined over goods—each *copy* of a good has the same price. We will write prices as vectors $p = (p_g)_{g \in [m]} \in \mathbb{R}^m_{\geq 0}$ or as functions over bundles $p : \mathcal{G} \to \mathbb{R}_{\geq 0}$ such that $p(S) = \sum_{g \in S} p_g$. Buyers demand bundles maximizing utility.

Definition 2.2 (Demand correspondence). The *demand correspondence* for buyer $q \in N$ at prices p is $\mathcal{D}_q(p) = \operatorname{argmax}_{S \in \mathcal{G}}\{u_q(S; p)\}$. We call bundles $S \in \mathcal{D}_q(p)$ *demand* bundles. Note that $\mathcal{D}_q(p)$ contains only bundles with non-negative utility, since $u_q(\emptyset; p) = 0$ for every p.

We focus our investigation on *Walrasian equilibria*, defined by a pricing and an allocation.

Definition 2.3 (Walrasian equilibrium). For valuations $\{v_q\}_{q \in N}$, we say that a pair (p, μ) of prices p and allocation $\mu = (\mu_q)_{q \in N}$ is a *Walrasian equilibrium* (WE) if both: 1) $\mu_q \in \mathcal{D}_q(p)$ for all $q \in N$; and 2) for every good g, $\sum_{q \in N} \mathbb{1}\{g \in \mu(q)\} < s_g$ only if $p_g = 0$.

We call the price vector p a *Walrasian equilibrium price vector*. Note that there may be many distinct Walrasian equilibrium price vectors: in fact, the set of all Walrasian prices forms a lattice. The *minimum* Walrasian equilibrium price vector p is the Walrasian equilibrium price vector that is coordinate-wise minimal amongst all Walrasian equilibrium price vectors.

Likewise, we call the allocation μ a *Walrasian allocation*. While there may be multiple distinct Walrasian allocations, it is known that all such allocations must maximize welfare.

In general, Walrasian equilibrium prices p are not sufficient to coordinate a corresponding allocation μ, because buyers might have indifferences ($|\mathcal{D}_q(p)| > 1$); if buyers choose one bundle arbitrarily, the resulting allocation can violate supply constraints. To reason about the degree of this violation, we make the following two natural definitions.

Definition 2.4 (Demanders and over-demand)**.** The set $U(g; p)$ of *demanders* for a good $g \in [m]$ at price p is the set of buyers that demand g: $U(g; p) = \{q \in N : \exists D \in \mathcal{D}_q(p) \quad \text{where } g \in D\}$.

Then, the *over-demand* $OD(g; p)$ for g at prices p is the number of demanders beyond the supply of a good: $OD(g; p) = \max\{|U(g; p)| - s_g, 0\}$. Read another way, the over-demand is the worst-case excess demand if bidders break ties in their demand correspondence arbitrarily.

To build intuition, we focus the first part of our paper on unit demand bidders, where $v_q(S) = \text{argmax}_{g \in S}\{v_q(g)\}$ for all $S \in \mathcal{G}$, and non-empty bundles $S \in \mathcal{D}_q(p)$ are *demand goods*.

3   Lower bound

To build intuition for why tie-breaking at equilibrium can lead to infeasibility, we give an example of a market with n buyers in which the over-demand of a good can be $\Omega(n)$, for any tie-breaking rules buyers use, so long as they cannot coordinate with one another after seeing the market instance.

To see why lack of coordination can be a problem, consider the following instance.

Lemma 3.1. *There exist unit demand valuations such that at the minimal Walrasian prices p, there exists a good with over-demand $\Omega(n)$.*

Proof. Consider a market with n unit demand buyers $N = [n]$ and $m = n$ distinct goods. For $g \in [m]$ a distinguished good, every buyer q has valuation $v_q(q) = v_q(g) = 1$ and $v_q(g') = 0$ for all $g' \notin \{q, g\}$. The minimal Walrasian equilibrium prices are $p = \mathbf{0}$, and the unique max-welfare allocation is $\mu_q = q$. At these prices, g is demanded by every buyer. Hence, $OD(g; p) = n - 1$. $\square$

Even if bidders are allowed to break ties in a clever way, the problem remains. Given a set of prices p, buyers N and tie breaking rule e_q^p for buyer q, let the demanders of g be $U^{e^p}(g; p) = \{q \in N : g \in \mathcal{D}_q(p) \text{ and } g = e_q^p(\mathcal{D}_q(p))\}$, and the *tie-breaking over-demand* with respect to e^p be $OD^{e^p}(g; p) = \max\{|\{q \in U^{e^p}(g; p)\}| - s_g, 0\}$.

Lemma 3.2. *There exists a distribution over unit demand valuations such that for any set of tie-breaking rules, the expected tie-breaking over-demand from n buyers is $\Omega(n)$.*

While this result does demonstrate that over-demand can be high without coordination, it is in some sense pathological—the buyers' valuations are highly similar. In the following section we give a simple and natural condition which rules out this example, and more generally ensures that the over-demand for any good is at most 1 at the minimal Walrasian prices regardless of tie-breaking. This bound is the best possible: for *any* buyer valuations, minimal Walrasian prices always induce over-demand of at least 1 for every good with positive price.

Lemma 3.3. *Fix any set of buyer valuations $\{v_q : q \in [n]\}$, and let p be a minimal Walrasian equilibrium price vector. For any good g with positive price $p(g) > 0$, we have $OD(g; p) \geq 1$.*

4   Unit demand

Now that we have seen how buyer indifferences at equilibrium can lead to allocations that exceed market supply, we consider whether the over-demand is large for "typical" instances. To build intuition, we start with the special case of unit demand valuations, defined by $v_q(S) =$

$\operatorname{argmax}_{g \in S}\{v_q(g)\}$. For each buyer $q \in N$, such valuations can be encoded with m real numbers $v_q(g)$ for each $g \in [m]$. Thus, we can restrict our attention to allocations which are a many-to-one matchings between buyers and goods—each buyer should be matched to at most 1 good, and each good g should be matched to at most s_g buyers.

We now give conditions on unit demand valuations which ensure that at the minimal Walrasian prices, when buyers buy *arbitrary* items in their demand sets, the resulting selection of goods results in high welfare and low over-demand. Accordingly, we need to reason in precise ways about how the equilibrium prices depend on the valuations. Getting our hands on this relation is surprisingly tricky—typical characterizations of Walrasian equilibrium prices are not enough for our needs.

For instance, two standard characterizations show that unit demand Walrasian prices i) are dual variables to a particular linear program (the "many-to-one matching linear program"), and ii) are computed from ascending price auction dynamics. While the former observation is often useful, it does not provide fine-grained information about how the prices depend on the valuations. While the latter understanding is useful for computing prices, the auction may proceed in a complicated manner, obscuring the relationship between the prices and valuations.

Swap graph. To reason about how prices depend on valuations, we define a graph called the *swap graph* of a Walrasian equilibrium (p, μ). This graph directly encodes buyer indifferences induced by the prices, which is what we ultimately want to reason about. Furthermore, the swap graph allows us to read off equations involving the prices and valuations. We define the swap graph as follows.

Definition 4.1 (Swap graph for (p, μ)). The *swap graph* $G = (V, E)$ defined with respect to a Walrasian equilibrium (p, μ) has a node for each good g and an additional null node $\bot$ representing the empty allocation: $V = [m] \cup \{\bot\}$. There is a directed edge $(a, b) \in E$ for $a \neq b$ from a to $b \neq \bot$ for each buyer q that receives good a in μ but also demands b, i.e. if $\mu_q = a$ and $b \in \mathcal{D}_q(p)$ for some buyer $q \in N$. Note that while there may be parallel edges—representing the same indifferences by different buyers—there are no self loops since edges run between *distinct* goods.

Since we will work extensively with the swap graph in the remainder of the section, it will be convenient to rephrase Theorem 3.3 in terms of the swap graph.

Corollary 4.2. *For buyers with unit demand valuations and a Walrasian equilibrium (p, μ) with p the minimal Walrasian prices, any node in the swap graph G with in-degree zero has price zero.*

Almost by definition, the over-demand of a good g is its in-degree in the swap graph. The proof of this lemma can be found in the full version.

Lemma 4.3. *Let G be the swap graph corresponding to any Walrasian equilibrium. If a node g in G has in-degree d then $OD(g; p) \leq d$.*

Thus, to bound the maximum over-demand of any good, it suffices to bound the in-degree for every good in the swap graph. Before turning to this task, we first introduce simple conditions on valuations that will rule out pathological market instances with high over-demand.

Generic valuations. Recall that in Section 3 we showed that over-demand can be high at the minimal Walrasian equilibrium prices. To provide a better bound on over-demand, we need an additional condition on valuations (ideally, a condition that will hold "typically", say with high probability under a perturbation of the valuations). In the lower bound instance for Theorem 3.2,

over-demand is large because the buyers have valuations that are too similar. Indeed, consider a market with two goods a and b where all buyers have the same *difference* in valuations between a and b. If *some* buyer is indifferent between a and b—by Theorem 3.3, this must be the case in equilibrium—*all* buyers are indifferent. This observation motivates our genericity condition.

Definition 4.4 (Generic valuations)**.** A set of valuations $\{v_q(g) \in \mathbb{R} : q \in N, g \in [m]\}$ is *generic* if they are linearly independent over $\{-1, 0, 1\}$, i.e.

$$\sum_{q \in N} \sum_{g \in [m]} \alpha_{q,g} v_q(g) = 0 \qquad \text{for } \alpha_{q,g} \in \{-1, 0, 1\} \quad \text{implies } \alpha_{q,g} = 0 \text{ for all } q \in N, g \in [m].$$

Remark 4.5. Note that this condition holds with probability 1 given any continuous perturbation of a profile of valuation functions, and so for many natural distributions, a profile of valuation functions will "generically" (i.e., with high probability) satisfy our condition. We also show how to discretely perturb a fixed set of valuations to satisfy our condition in the extended version.

Over-demand. Now, we are ready to present the main technical result of this section. Subject to our genericity condition, when buyers select an *arbitrary* good in their demand correspondence given minimal Walrasian equilibrium prices, over-demand is low and welfare is high. Our first goal will be to show that the *in-degree* for any node in our swap graph is at most 1. This will imply that no good has over-demand more than 1, regardless of its supply. The proof of all lemmas in this section can be found in the full version. We proceed via a series of properties of the swap graph.

Lemma 4.6. *The swap graph G, defined with respect to Walrasian equilibrium (p, μ) and generic valuations $\{v_q(g) : q \in N, g \in [m]\}$, is acyclic.*

Because the swap graph G is acyclic, we can choose a partial order of the nodes so that all edges go from nodes earlier in the ordering to nodes later in the ordering (e.g., by topologically sorting the graph). For the remainder of the argument, we rename our nodes according to such an ordering (i.e. we now have for every edge $(a_i, a_j) \in E$ implies $i < j$). Now, the price of a good can be written in terms of the valuations for goods that come earlier in this ordering.

Lemma 4.7. *For every good g, the price $p(g)$ can be written as a linear combination of valuations $v_q(j)$ over $\{-1, 0, 1\}$ for $q \in N$ and $j < g$. Specifically, for every good g, and for every set of goods $g_1 < g_2 < \ldots < g_k < g$ such that $g_1 \to \ldots \to g_k \to g$ forms a path in the graph and g_1 has in-degree 0, there is a set of buyers $q_1, \ldots, q_k$ such that $\mu_{q_i} = g_i$ and $p(g) = \sum_{i=1}^{k} (v_{q_i}(g_{i+1}) - v_{q_i}(g_i))$ where $g_{k+1} = g$. If the first node $g_1 = \perp$, we define $v_{q_1}(\perp) = 0$ by convention.*

Finally, we can show that no node in the swap graph has in-degree greater than 1.

Lemma 4.8. *For generic buyer valuations, every node in the the swap graph defined with respect to a Walrasian equilibrium (p, μ) with minimal Walrasian prices has in-degree at most 1.*

Finally, by Theorem 4.3 the over-demand for any good is at most its in-degree in the swap graph, so the lemmas immediately imply the following theorem.

Theorem 4.9. *For any set of unit demand buyers with generic valuations, and for p the minimal Walrasian equilibrium price vector computed with respect to these buyers, the over-demand for any good $g \in [m]$ is at most 1, independent of its supply s_g: $OD(g; p) \le 1$ for all goods $g \in [m]$.*

As a result, when generic buyers face minimal[3] prices p and buy a good in their demand set—resolving indifferences in arbitrary or even adversarial ways—the excess demand of any good is at most 1. In particular, this holds when buyers break ties in favor of non-empty allocations. This tie-breaking rule will be necessary to obtain good welfare properties, which we turn to now.

Welfare. When buyers choose arbitrary goods in their demand correspondence, the over-demand of each good is at most 1. We also lower bound the welfare achieved when buyers choose arbitrarily amongst the goods in their demand correspondence. In our calculations of welfare, Welfare$_N(B)$, for some (possibly infeasible) allocation B, we assume over-demand is resolved in a worst-case way—i.e. we assume goods are allocated to buyers that demand them in the way that minimizes welfare, while obeying the supply constraints. For the welfare argument only, we assume that a buyer chooses a *nonempty* demand bundle whenever possible. Subject to this restriction, buyers can break ties however they like. Then, we can lower bound the welfare of these buyers N in terms of the optimal welfare Opt-Welfare$_N$.

Theorem 4.10. *Consider any set of buyers N with generic unit demand valuations bounded in $[0, H]$ and the minimal Walrasian equilibrium prices p. For each buyer q, let $b_q \in \mathcal{D}_q(p)$ be some arbitrary set in his demand correspondence, assuming only that $b_q \neq \emptyset$ whenever $|\mathcal{D}_q(p)| > 1$. Then the welfare obtained by the resulting allocation is always near-optimal: Welfare$_N(b_1, b_2, \ldots, b_n) \geq$ Opt-Welfare$_N - 2 \cdot m \cdot H$.*

5   Towards gross substitutes

In this section, we generalize our results beyond unit demand buyers. Ideally, we would like to extend our results to buyers who have *gross substitutes* valuations. It is known that if all buyers have GS valuations, there always exist Walrasian equilibria, which can be found by following a tatônnement procedure [28]. Gul and Stacchetti [24] show that the class GS is in some sense the most general class of valuations that are guaranteed to have WE.

While GS is an attractive class to target, its definition is axiomatic rather than constructive. This poses a problem for defining genericity: it is not obvious which, if any, valuations in GS satisfy some proposed definition of genericity. Ultimately, we will prove our results for buyers with *matroid based valuations* (MBV), which are a subclass of gross substitutes valuations, but conjectured to be equal to all of gross substitutes [37]. Such valuations can be explicitly constructed from a set of numeric weights and a matroid, giving us a path to define and construct generic valuations, as well as certifying that generic valuations lie in MBV and are contained in GS.

However, there are still several obstacles to be overcome, and the arguments are significantly more involved than for unit demand. Roughly, the overarching difficulty is establishing a connection between valuations on bundles, and valuations on the items in the bundle, since in some sense we will require genericity on values for single goods. While this connection is immediate in the case of unit demand, the situation for MBV is more delicate—some bundles may contain "irrelevant" goods, which don't contribute at all to the valuation, for instance. We address this issue with a more complex swap graph, designed to capture the richer structure of GS valuations. To prevent buyers from taking bundles that contain goods that do not add value, we focus on *non-degenerate*

---

[3]We show in the full version that even with generic valuations, *non-minimal* Walrasian prices can still induce high over-demand, further justifying our focus on minimal Walrasian prices.

bundles which ensures that buyers only purchase most demanded bundles that do not contain goods that impart zero marginal value. We formally define non-degenerate bundles in the full version.

Despite the complications, our high-level argument remains the same as in unit demand. We first define a swap graph and connect it to over-demand. Then, we define a generic version of MBV (GMBV) and prove properties about the swap graph for valuations in GMBV.

Swap graph with GS valuations. To define the swap graph for gross substitutes, the core idea is to have an edge (a, b) represent a single *swap* of good a for good b in some larger bundle, naturally generalizing our construction for unit demand. The main challenge is ensuring that we faithfully model buyer indifferences—which are between *bundles* of goods—via indifferences of single swaps. More precisely, in order to bound over-demand with arbitrary tie-breaking, we must ensure that if a buyer is indifferent between her Walrasian allocation and some other bundle B, then there must be one incoming edge to every good in B. If the swap graph satisfies this property, we can describe potential demand for goods in terms of in-degree of nodes, like we did for unit demand.

For the first step, we want that if a buyer is indifferent between bundles B_1 and B_2, every good in B_1 can be exchanged for a good in B_2 while preserving utility. While this is not true for general bundles, it is true for the "smallest" bundles in a buyer's demand correspondence.

Definition 5.1 (Minimal demand correspondence). For price vector p and buyer $q \in N$, the *minimum demand correspondence* is $\mathcal{D}_q^*(p) = \{S \in \mathcal{D}_q(p) : T \notin \mathcal{D}_q(p) \text{ for all } T \subsetneq S\}$.

Crucially, the minimum demand correspondence for GS valuations forms the bases of a matroid [8]. By standard facts from matroid theory [38], we have the exchange property: for B_1, B_2 in the basis set $\mathcal{B}$ of a matroid and for every $b \in B_1 \setminus B_2$, there exists $b' \in B_2 \setminus B_1$ such that $B_1 \cup b' \setminus b \in \mathcal{B}$ and $B_2 \cup b \setminus b' \in \mathcal{B}$. If all bundles in a buyer's demand correspondence have the same size, then we can take $\mathcal{B}$ to be $\mathcal{D}_q^*(p)$, and we have exactly what we need: for two bundles in the correspondence, every good can be swapped while remaining in the correspondence. Of course, buyers may prefer bundles of different sizes. We return to this issue after we define the swap graph.

Definition 5.2 (Swap graph). Let buyers have GS valuations, (p, μ) be a WE, and for each buyer $q \in N$ fix a minimum demand bundle $M_q \in \mathcal{D}_q^*(p)$ where $M_q \subseteq \mu_q$. Define the *swap graph* $G(p, \mu, (M_q)_{q \in N})$ to have a node for every good $g \in [m]$ and an additional node which we refer to as the *null node* $\bot$. There is a directed edge from (a, b) for every buyer $q \in N$ such that $a \in M_q$, $b \notin \mu_q$, and there exists $B \in \mathcal{D}_q^*(p)$ with $b \in B$ where $M_q \cup b \setminus a \in \mathcal{D}_q^*(p)$. Further, the graph contains the edge $(\bot, b)$ if b has strictly positive price, and there exists some buyer $q \in N$ with $b \notin \mu(q)$ and $B \in \mathcal{D}_q(p)$ where $b \in B$ and $B \setminus b \in \mathcal{D}_q^*(p)$.

We have two conditions in our definition of which edges exist from the null good: the first involves demand bundles, while the second involves positive price. These two conditions solve two distinct problems. The first condition models cases where the demand correspondence has bundles of different sizes, say there is some bundle B with cardinality strictly larger than the minimum cardinality demand bundle. We need all goods in B to have an incoming edge, to reflect over-demand if a buyer selects B. We cannot ensure these edges via matroid properties, since B is not a minimum cardinality demand bundle and so it is not a matroid basis. However, there is a minimum bundle B^{min} contained in B, which *is* a matroid basis, so we have incoming edges to B^{min}. We can show that for every good $g \in B \setminus B^{min}$ the bundle $B^{min} \cup g$ is also a demand bundle, so the swap graph will also have edges from $\bot$ to g, covering all goods in B as desired.

Up to this point, we have argued informally that if B is a bundle in $\mathcal{D}_q^*(p)$, then all goods in $B \setminus M_q$ will have an incoming edge. We plan to bound the over-demand by the in-degree, but there is an important corner case: goods g with price 0. Such goods can be freely added to *any* buyer's bundle, ruining the bound on over-demand. However, all is not lost: the problem stems from buyers with zero marginal valuation for a good in their bundle, which we can rule out.

Definition 5.3 (Non-degenerate)**.** A bundle S is *non-degenerate with respect to S'* if goods in S' have non-zero marginal value: $v(S \setminus g) < v(S)$ for each $g \in S \cap S'$. When $S' = [m]$, we will say S is *non-degenerate*. The *non-degenerate correspondence* for buyer $q \in N$ at price p is the set of demand bundles defined by $\mathcal{D}_q^\bullet(p) = \{S \in \mathcal{D}_q(p) : v_q(S \setminus g) < v_q(S) \ \forall g \in S\}$.

If we only consider buyers who purchase bundles in $\mathcal{D}_q^\bullet(p)$, there is no problem with zero-price goods g: for buyers with non-degenerate bundles that contain g, that good must automatically be in any minimum bundle for those buyers. So, we only need edges in the swap graph from $\perp$ to goods with positive price.

We are almost ready to formally connect the swap graph with over-demand, but there is one last wrinkle. Since a buyer is assigned a bundle rather than a single good, she may have two different swaps to the same good: perhaps $M_q \setminus a_1 \cup b$ and $M_q \setminus a_2 \cup b$ are both demanded. We want to count this buyer as causing demand 1 rather than 2, since she will select a bundle of distinct goods.

Definition 5.4 (Buyer in-degree)**.** Consider buyers with GS valuations $\{v_q\}_{q \in N}$, the minimal Walrasian prices p and allocation μ, and the corresponding swap graph G. The *buyer in-degree* of node b is the number of distinct buyers with an edge directed to b.

We can also define the over-demand $OD^\bullet(g; p)$ of a good when buyers take only non-degenerate bundles. Like the unit demand case, this is simply the worst-case over-demand assuming that buyers choose an arbitrary non-degenerate bundle from their demand correspondence, without any assumption on how they break ties. Finally, we can relate the swap graph to over-demand.

Lemma 5.5. *Let buyers have GS valuations and let (p, μ) be a Walrasian equilibrium with minimal prices. If a node $g \in [m]$ in the swap graph G has buyer in-degree at most d, then $OD^\bullet(g; p) \le d$.*

Bounding the buyer in-degree. To conclude, we first define a generic class of valuations (GMBV). Then, we show properties of the swap graph, similar to the unit demand case. Then, the following theorem bounds buyer in-degree and hence over-demand by Theorem 5.5.

Theorem 5.6. *For GMBV buyers and Walrasian equilibrium (p, μ) with minimal Walrasian prices, each node in the swap graph G has buyer in-degree at most 1.*

6 Welfare and over-demand generalization

In this section, we show that Walrasian prices *generalize*: the equilibrium prices for a market N of buyers induces similar behavior when presented to a new sample N' of buyers, both in terms of the demand for each good and in terms of welfare. More precisely, a set of prices which minimizes over-demand of each good *and* maximizes welfare when each buyer purchases her most-preferred bundle retains these properties (approximately) on a *new* market N' when each buyer purchases her most-preferred bundle, if buyers in N and N' are drawn independently from the same distribution.

Generalization results for arbitrary valuations. To speak precisely about how over-demand and welfare induced by prices p vary between two markets, we first fix a tie-breaking rule e that buyers use to select amongst demanded bundles relative to a valuation from set $\mathcal{V}$. We define classes of *functions* parameterized by pricings mapping valuations to (a) bundles purchased, (b) whether or not a particular good g is purchased, and (c) the value a buyer gets for her purchased bundle.

Let $e : 2^{\mathcal{G}} \times \mathcal{V} \to \mathcal{G}$ denote some tie-breaking function, which satisfies some minor technical conditions defined in the full version. We call the bundle that e selects from $\mathcal{D}_q(p)$ the *canonical* bundle for q at p, denoted as $B_q^*(p) = e\left(\mathcal{D}_q(p), v_q\right)$. Then, for each good g, let $h_{g,p}(v_q) = \mathbb{1}[g \in B_q^*(p)]$ indicate whether or not g is in q's canonical bundle at prices p. Let $C^e(g; p; N) = \sum_{q \in N} h_{g,p}(v_q)$, i.e. the number of buyers in N whose canonical bundles at p contain g. For a sample of n buyers $\{v_q\} \overset{i.i.d.}{\sim} \Pi$, let $C^e(g; p; \Pi)$ represent the *expected* number of copies of g demanded at prices p if buyers demand canonical bundles. Similarly, let the *welfare* of a pricing p on a market N be $\text{Welfare}_N(p) = \sum_{q \in N} v_q(\widehat{B_q^*}(p))$ for $\widehat{B_q^*}(p)$ a worst-case resolution of the over-demand for each good.

Before presenting the technical details, we state two main results showing how the behavior induced by prices generalizes for buyers with arbitrary valuations. Our first theorem bounds the over-demand when the Walrasian prices computed for a market are applied to a new market.

Theorem 6.1. *Fix a pricing p and two sampled markets N, N' of buyers with arbitrary valuations, with $|N| = |N'| = n$. For good g, suppose $C^e(g; p; N) \leq s_g + 1$. Then, for any $\alpha \in (0, \frac{4}{5})$, if $s_g = \Omega\left(\frac{m \cdot \ln \frac{1}{\delta}}{\alpha^2}\right)$, with probability $1 - \delta$, $C^e(g; p; N') \leq (1 + \alpha) s_g$.*

Our second theorem is an analagous generalization result, for welfare instead of over-demand.

Theorem 6.2. *Fix two markets $N, N' \sim \Pi$ for which $|N| = |N'| = n$. With probability $1 - \delta$, for any $\alpha \in (0, 4/5)$, if p are welfare-optimal for N and*

$$Opt\text{-}Welfare_N = \Omega\left(\frac{H^3 n^{0.5} m^4 \ln^2(m) \ln^2 \frac{1}{\delta}}{\alpha^2}\right)$$

then $Welfare_{N'}(p) \geq (1 - \alpha) Opt\text{-}Welfare_{N'}$.

These proofs can be found in the full version. Theorem 6.1 relies on bounding the VC dimension of the class of good g's demand indicator functions (the class contains a function for each pricing which labels a valuation v positive if and only if g is in the canonical bundle for v at those prices). This is done by first proving the class of *bundle* predictors (the class contains one function for each pricing p, which maps v to v's canonical bundles at p) is *linearly separable* in a space of $m + 1$ dimensions. Using a recent result by Daniely and Shalev-Shwartz [13], this class admits a compression scheme of size $m + 1$, and can therefore be (ϵ, δ)-PAC learned with $O\left(\frac{m}{\epsilon} \ln \frac{1}{\delta}\right)$ samples. Since item predictor functions are simply projections of bundle prediction functions, their PAC-complexity cannot be greater. Thus, since item predictors are binary valued, we can apply the classical result of [19] which shows the equivalence of learning and uniform convergence for binary prediction problems, and can upper bound the class's VC dimension as a function of its PAC sample complexity.

We prove Theorem 6.2 by bounding the pseudo-dimension of the class of welfare predictor functions (containing, for each pricing a function which map valuations to the *value* of a buyer purchasing her canonical bundle at these prices). The argument uses the existence of a compression scheme to upper bound the number of possible distinct labelings of valuations by bundles. Fixing

a bundle labeling of a valuation also fixes the welfare of that valuation; thus, one can upper-bound the size of a "shatterable" set and the pseudo-dimension of the class of welfare predictors.

References

[1] Martin Anthony and Peter L. Bartlett. *Neural Network Learning: Theoretical Foundations*. Cambridge University Press, NY, NY, USA, 1999.

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

[3] Moshe Babaioff, Liad Blumrosen, Shaddin Dughmi, and Yaron Singer. Posting prices with unknown distributions. In *Innovations in Computer Science (ICS)*. Tsinghua University Press, January 2011.

[4] Maria-Florina Balcan, Avrim Blum, Jason D. Hartline, and Yishay Mansour. Mechanism design via machine learning. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 605–614, Pittsburgh, PA, USA, 2005. IEEE.

[5] Maria-Florina Balcan, Nikhil Devanur, Jason D. Hartline, and Kunal Talwar. Random Sampling Auctions for Limited Supply. Technical report, Carnegie Mellon University, September 2007.

[6] Maria-Florina Balcan, Avrim Blum, Jason D Hartline, and Yishay Mansour. Reducing mechanism design to algorithm design via machine learning. *Jour. of Comp. and System Sciences*, 74(8):1245–1270, 2008.

[7] Maria-Florina Balcan, Amit Daniely, Ruta Mehta, Ruth Urner, and Vijay V Vazirani. Learning economic parameters from revealed preferences. In *Web and Internet Economics*, pages 338–353. Springer, 2014.

[8] Oren Ben-Zwi, Ron Lavi, and Ilan Newman. Ascending auctions and walrasian equilibrium. *CoRR*, abs/1301.1153, 2013.

[9] Avrim Blum and Jason D. Hartline. Near-optimal online auctions. In *SODA*, pages 1156–1163, Vacouver, BC, Canada., 2005. SIAM.

[10] Avrim Blum, Vijay Kumar, Atri Rudra, and Felix Wu. Online learning in online auctions. In *Symposium on Discrete Algorithms*, Baltimore, MD, USA., 2003. ACM.

[11] Nicol Cesa-Bianchi, Claudio Gentile, and Yishay Mansour. Regret minimization for reserve prices in second-price auctions. In *SODA*, pages 1190–1204, NY, NY, USA., 2013. SIAM.

[12] Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 243–252, NY, NY, USA, 2014. SIAM.

[13] Amit Daniely and Shai Shalev-Shwartz. Optimal learners for multiclass problems. In *Proceedings of The 27th Conference on Learning Theory*, pages 287–316, 2014.

[14] Xiaotie Deng, Christos Papadimitriou, and Shmuel Safra. On the complexity of equilibria. In *Proceedings of the Thiry-fourth Annual ACM Symposium on Theory of Computing*, STOC '02, pages 67–71, New York, NY, USA, 2002. ACM. ISBN 1-58113-495-9.

[15] Nikhil Devanur, Jason Hartline, Anna Karlin, and Thach Nguyen. Prior-independent multi-parameter mechanism design. In *Internet and Network Economics*, pages 122–133. Springer, Singapore, 2011.

[16] Nikhil R Devanur and Thomas P Hayes. The adwords problem: online keyword matching with budgeted bidders under random permutations. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 71–78. ACM, 2009.

[17] Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. *Games and Economic Behavior*, 2014.

[18] Shaddin Dughmi, Li Han, and Noam Nisan. Sampling and representation complexity of revenue maximization. In *Web and Internet Economics*, volume 8877 of *Lecture Notes in Computer Science*, pages 277–291. Springer Intl. Publishing, Beijing, China, 2014.

[19] Andrzej Ehrenfeucht, David Haussler, Michael Kearns, and Leslie Valiant. A general lower bound on the number of examples needed for learning. *Information and Computation*, 82(3): 247–261, 1989.

[20] Edith Elkind. Designing and learning optimal finite support auctions. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 736–745. SIAM, 2007.

[21] Satoru Fujishige and Zaifu Yang. A note on Kelso and Crawford's Gross Substitutes condition. *Mathematics of Operations Research*, 28(3):463–469, 2003.

[22] Andrew V Goldberg, Jason D Hartline, Anna R Karlin, Michael Saks, and Andrew Wright. Competitive auctions. *Games and Economic Behavior*, 55(2):242–269, 2006.

[23] Faruk Gul and Ennio Stacchetti. The English auction with differentiated commodities. *Journal of Economic Theory*, 92(1):66–95, 2000.

[24] Faruk Gul and Ennio Stacchetti. Walrasian equilibrium with gross substitutes. *Journal of Economic Theory*, 87(1):95–124, 1999.

[25] Steve Hanneke. The optimal sample complexity of PAC learning. *CoRR*, abs/1507.00473, 2015.

[26] Jason D. Hartline and Tim Roughgarden. Simple versus optimal mechanisms. In *ACM Conf. on Electronic Commerce*, Stanford, CA, USA., 2009. ACM.

[27] Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. Making the most of your samples. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC '15, pages 45–60, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3410-5.

[28] Alexander Kelso and Vincent Crawford. Job matching, coalition formation, and gross substitutes. *Econometrica*, 50(6):1483–1504, 1982.

[29] Nick Littlestone and Manfred Warmuth. Relating data compression and learnability. Technical report, Technical report, University of California, Santa Cruz, 1986.

[30] Andres Munoz Medina and Mehryar Mohri. Learning theory and algorithms for revenue optimization in second price auctions with reserve. In *Proceedings of The 31st Intl. Conf. on Machine Learning*, pages 262–270, 2014.

[31] Jamie Morgenstern and Tim Roughgarden. On the pseudo-dimension of nearly-optimal auctions. In *Twenty-ninth Annual Conference on Neural Information Processing Systems (NIPS)*, page To Appear, December 2015.

[32] Kenneth Mount and Stanley Reiter. The informational size of message spaces. *Journal of Economic Theory*, 8(2):161–192, 1974.

[33] Kazuo Murota. Convexity and Steinitz's exchange property. In William H. Cunningham, S. Thomas McCormick, and Maurice Queyranne, editors, *Integer Programming and Combinatorial Optimization*, volume 1084 of *Lecture Notes in Computer Science*, pages 260–274. Springer Berlin Heidelberg, 1996. ISBN 978-3-540-61310-7.

[34] Kazuo Murota. Valuated matroid intersection I: Optimality criteria. *SIAM Journal on Discrete Mathematics*, 9(4):545–561, 1996.

[35] Kazuo Murota and Akiyoshi Shioura. M-Convex function on generalized polymatroid. *Mathematics of Operations Research*, 24(1):95–105, 1999.

[36] Noam Nisan and Ilya Segal. The communication requirements of efficient allocations and supporting prices. *Journal of Economic Theory*, 129(1):192 – 224, 2006.

[37] Michael Ostrovsky and Renato Paes Leme. Gross substitutes and endowed assignment valuations. *Theoretical Economics*, 2015.

[38] James G Oxley. *Matroid Theory*. Oxford Graduate Texts in Mathematics. Oxford University Press, 1997.

[39] David Pollard. *Convergence of stochastic processes*. David Pollard, New Haven, Connecticut, 1984.

[40] T. Roughgarden and O. Schrijvers. Ironing in the dark. Submitted, 2015.

[41] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

[42] Akiyoshi Shioura and Akihisa Tamura. Gross substitutes condition and discrete concavity for multi-unit valuations: a survey. *Journal of the Operations Research Society of Japan*, 58(1): 61–103, January 2015.

[43] Vladimir N Vapnik and A Ya Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability & Its Applications*, 16(2):264–280, 1971.

[44] Vladimir Naumovich Vapnik and Samuel Kotz. *Estimation of dependences based on empirical data*, volume 40. Springer-Verlag New York, 1982.