Equilibrium of Data Markets with Externality

Safwan Hossain [1]   **Yiling Chen** [1]

Abstract

We model real-world data markets, where sellers post fixed prices and buyers are free to purchase from any set of sellers, as a simultaneous game. A key component here is the negative externality buyers induce on one another due to data purchases. Starting with a simple setting where buyers know their valuations a priori, we characterize both the existence and welfare properties of the pure Nash equilibrium in the presence of such externality. While the outcomes are bleak without any intervention, mirroring the limitations of current data markets, we prove that for a standard class of externality functions, platforms intervening through a transaction cost can lead to a pure equilibrium with strong welfare guarantees. We next consider a more realistic setting where buyers learn their valuations over time through market interactions. Our intervention is feasible here as well, and we consider learning algorithms to achieve low regret concerning both individual and cumulative utility metrics. Lastly, we analyze the promises of this intervention under a much richer externality model.

1. Introduction

Data plays a central role in machine learning, and demand for it has grown substantially due to the increasing value it provides. It has thus become the subject of trading, and understanding markets for data (or more generally, information) has gained traction in research communities in recent years (Bergemann et al., 2018; Babaioff et al., 2012; Bergemann & Bonatti, 2019; Mehta et al., 2019; Agarwal et al., 2019; 2020). Such works have largely focused on determining the "right" price for data so that it is allocated to those who value it most (Bergemann & Bonatti, 2019; Mehta et al., 2019; Agarwal et al., 2019; 2020). The mechanisms

proposed here for two-sided markets are usually auction-based, and require buyers to report their valuations for data sources (Agarwal et al., 2019; 2020). These requirements are challenging, if not impossible, to achieve in practice due to the peculiar characteristics of data products. Buyers are generally free to purchase from any combination of data sources, and it is impractical to elicit accurate valuations in this combinatorially large space. More importantly, a data source's value to a buyer largely arises from how it can improve model performance, a metric the buyer herself may not know before making purchases and evaluating the results. Thus, she acts in the market with only partial or incomplete information. It is hence not surprising that most real-world data market platforms, such as Snowflake Data Marketplace[1] and AWS Data Exchange[2], take a very simple format: sellers use some pricing mechanism to set a price for each dataset, and buyers can freely choose which datasets to purchase, with platforms charging a simple transaction fee. While sellers can change prices over time to reflect market changes, the granularity of this is quite coarse, and prices stay stable or *fixed* for some time. The simplicity of fixed-price data markets is not just a huge operational advantage but also arguably necessary in light of the challenges data products pose. Such markets are the focus of our paper.

An important component of modeling data markets is accounting for the (usually negative) externality one buyer's purchase decision has on another. In a competitive setting, a buyer's value for data is predicated on the relative advantage it provides, with respect to their peers. Equivalently, the value of data can depend on what others have access to. While externality is not unique to data and is present for other products as well, it is especially prominent in data markets due to another salient feature: data can be replicated at a mass scale with zero marginal cost and sold to multiple buyers, a phenomenon that exacerbates any externality data products induce. Replication need not even be exact: sellers can offer different versions of the same dataset by injecting noise or by interleaving it with something innocuous (Pei, 2020). Accordingly, real-world data markets usually give no guarantee of unique ownership. Data externality is thus a persistent phenomenon, and by linking one buyer's utility to another's decisions, it turns the buyer interactions into

---

[1]Harvard University, Cambridge, USA. Correspondence to: Safwan Hossain <shossain@g.harvard.edu>.

---

[1]https://www.snowflake.com/
[2]https://docs.aws.amazon.com/data-exchange/index.html

a game. We thus model fixed-price data markets between arbitrary buyers and sellers in the presence of externality and free replicability as a game. Despite their simplicity, which sidesteps many of the above concerns, this model has not been formally analyzed in literature and its properties are unknown. By leveraging the formalisms of game theory, we systematically explore this landscape and ask:

- How well can fixed-price markets serve the purpose of allocating data to buyers, especially with respect to social welfare?

- How does buyer externality affect their strategy and market performance?

- Can simple platform interventions improve the outcome of fixed-price data markets?

- If buyers learn their value for data sources by repeatedly interacting with a data market, can they learn to act optimally, and how much social welfare can be achieved?

1.1. Our contributions

We model the interactions between buyers in fixed-price data markets as a simultaneous game, with buyer utility depending on the net gain derived from the purchased data, the externality they suffer due to others' actions, and possibly some cost administered by the platform. We formally define this in section 3. In section 4, we study the pure Nash equilibrium (PNE) of this game and show that without any intervention, this equilibrium has poor social welfare, confirming the present concerns with data markets. A market intervention in the form of a transaction cost is then proposed that significantly improves the situation by guaranteeing a dominant strategy NE whose social welfare approaches the optimal. Section 5 then considers the real-world setting wherein buyers do not know valuations a priori but instead must learn them through repeated market interactions. We model the buyer's learning problem as a multi-armed bandit instance with exponentially many arms, show the proposed platform intervention is feasible here as well, and consider online learning algorithms for buyers in this challenging combinatorial setting. Section 6 explores a richer model of buyer externality, wherein not intervening does not guarantee even a reasonable approximate equilibrium. In contrast, our proposed intervention significantly improves this, guaranteeing a $\varepsilon-$pure NE with good welfare. Nonetheless, this is a more challenging model for the learning setting, and we discuss these along with broader extensions in section 7.

2. Related Works

Bergemann & Bonatti (2019) provide a broad overview of the growing literature on markets for information and data. Works here examine how to optimally sell information according to some objectives, and take a mechanism de-

sign approach, with a focus on the seller's pricing problem. For instance, many works consider a monopoly information holder directly selling a private random signal to buyers (Eso & Szentes, 2007; Bergemann et al., 2018; Babaioff et al., 2012; Chen et al., 2020b; Cai & Velegkas, 2020; Bonatti et al., 2022). The seller decides on a menu of information products (e.g. partially revealing the signal) and an associated price for each product to maximize her profit. Mehta et al. (2019) also study selling a dataset in a monopoly seller setting. When there are multiple sellers and multiple buyers, auction-based mechanisms have also been leveraged to design two-sided data marketplaces (Agarwal et al., 2019; 2020). In such works, the marketplaces themselves are not profit-driven but intend to maximally facilitate the matching of data to buyers. Bergemann & Bonatti (2019) and Bergemann et al. (2022) also study profit-driven data intermediaries who make a bilateral deal to purchase information from data holders and then sell the information to data buyers. Here, the data holders are consumers, and the buyers are firms that use purchased information to price discriminate.

While determining how to price data is fundamentally important, the proposed mechanisms so far have not found their way into real-world data markets. A menu of partially-revealing information products can be too complex and cumbersome for buyers. As discussed in the introduction, buyers are unlikely to know their valuation for data a priori, rendering mechanisms that solicit buyer valuations (e.g. auctions) impractical. Our work thus takes an orthogonal direction and steps away from the design question of how to sell data and consider data sellers participating in a real-world fixed-price data market. We take this market mechanism as a given and adopt a game-theoretic approach to understanding buyer behavior and dynamics in such markets.

A key feature of data markets that is central to our model is buyer externality. This has been highlighted in studying monopoly data selling (Admati & Pfleiderer, 1986; 1990; Bimpikis et al., 2019; Bonatti et al., 2022) and data auctions (Agarwal et al., 2020) in competitive environments. These works explicitly model downstream competition (e.g. trading in financial markets) among data buyers and the negative buyer externality that arises therein. Our work abstracts away the specifics of competitive environments to present such externality in a general sense. Another type of externality present in information markets is the externality among data sellers. Naturally, the value of one's data decreases when others decide to sell similar data (Bergemann et al., 2022). Seller externality is an important phenomenon to consider, which we leave for future work.

Lastly, our work on the online setting relaxes the assumption that buyers know their valuations a priori. This spiritually parallels studying learning agents in other mechanisms such as auctions (Blum et al., 2003; Weed et al., 2015; Braver-

man et al., 2018) and peer prediction (Feng et al., 2022). The techniques used in the online section leverage the vast literature on bandits, particularly bandits in metric spaces (Kleinberg et al., 2019; Bubeck et al., 2008). We view studying markets with learning agents as a step toward a more realistic evaluation of market performance.

3. Model

Market Structure: Consider n buyers, $\mathcal{D} = \{d_1, \ldots, d_n\}$, and k sellers, $\mathcal{L} = \{\ell_1, \ldots \ell_k\}$, who without loss of generality have one dataset each to sell. We consider a market where buyers are free to purchase from any subset of the sellers, each of whom posts a fixed price for their dataset and cannot refuse to sell. The data offered by the sellers can be arbitrarily correlated, and multiple buyers may buy from the same seller. Let Γ denote the power set of the sellers, and let $\gamma \in \Gamma$ denote a specific subset of sellers (we often refer to it as a *seller set* or *order*). We use subscripts to distinguish between seller sets chosen by various buyers - for example, γ_i refers to the set of buyers chosen by seller i - and superscripts to distinguish or index between two seller sets irrespective of buyers - for example, γ^m and γ^n refers to two distinct seller sets. We consider buyers simultaneously submitting their orders and use $S = (\gamma_1, \ldots, \gamma_n)$ to denote the purchase orders of all buyers.

Buyer's Utility: The utility of buyer i when purchasing from a seller set γ is predicated on: (1) the increased value/performance due to purchased data, denoted by the random variable $P_i(\gamma)$ with $\mathbb{E}[P_i(\gamma)] \triangleq p_i(\gamma)$, (2) the cost charged by sellers in γ, $c(\gamma)$, (3) the negative externality caused by another buyer j's action, denoted by the random variable $E_{ij}(\cdot)$ with $\mathbb{E}[E_{ij}(\cdot)] \triangleq e_{ij}(\cdot)$, and (4) a *transaction cost* $\mathcal{T}_i(S)$ charged by the platform, which specifies the amount that buyer i needs to pay (or receive) when the set of purchase orders is S. Without loss of generality, we assume $P_i(\gamma) \in [0, 1]$ and $c(\gamma) \in [0, 1]$, and the "buy nothing" option ($\gamma = \emptyset$) has 0 increased performance and 0 cost. Note a buyer that decides to "buy nothing" and not participate in the market, will still suffer the negative externality induced by others. We primarily focus on the most externality model in literature where buyer i's externality from buyer j's action depends arbitrarily on the latter's action: $E_{ij}(\gamma_j)$ (Aseff & Chade, 2008; Li et al., 2019; Agarwal et al., 2020). To ensure that the total externality caused by a buyer is at most 1 and thus in the same range as performance and cost, we assume that $\forall i$, $\sum_{j \neq i} E_{ji}(\gamma_i) \in [0, 1]$. In section 6, we consider a richer class of externality functions and in appendix A, we include a detailed discussion on the economic interpretations of these externality models. We now formally define the buyer's utility, and for brevity, combine the performance increase $P_i(\gamma)$ and seller costs $c(\gamma)$ terms into a single *net gain* term $G_i(\gamma)$, with $\mathbb{E}[G_i(\gamma)] \triangleq g_i(\gamma)$.

Definition 1 (Utility and Welfare). *For seller cost function c, and buyer i with performance increase P_i, we define her net gain to be $G_i(\gamma) = P_i(\gamma) - c(\gamma)$, with $G_i(\gamma) \in [-1, 1]$. Using this, along with externality $E_{ij} \forall j$, and transaction cost $\mathcal{T}_i$, we define her stochastic utility for a complete order profile $S = (\gamma_1, \ldots, \gamma_n)$ to be:*

$$U_i(S) = G_i(\gamma_i) - \sum_{j \neq i} E_{ij}(\gamma_j) - \mathcal{T}_i(S).$$

The expected buyer utility is denoted by $\mathbb{E}[U_i(S)] \triangleq u_i(S)$. We define the expected social welfare of an order profile S as the sum of all expected buyer utility: $sw(S) = \sum_i u_i(S)$.

Game and Solution Concept: We model buyers in the data market as playing a simultaneous-move game to maximize their expected utility. We note that agent i's utility, and thus her best response, doesn't depend on any other agent's net gain $G_j(\cdot)$, which may be beneficial. We are primarily interested in analyzing the pure-strategy Nash equilibrium of this game under externality and transaction costs. Beyond the existence of such equilibria, we also aim to compare the social welfare at equilibrium to the optimal social welfare possible. A common notion for comparison is the *price of stability*, the ratio of the optimal social welfare to the welfare of the best equilibrium (Nisan et al., 2007; Schulz & Moses, 2003). However, this is a multiplicative metric and thus unamenable to additive notions of regret that are common in the online analysis we do in section 5. As such, we define a comparable additive notion called *welfare regret at equilibrium (WRaE)* to characterize the societal cost of self-interested behavior in games.

Definition 2 (WRaE). *Let $S^* = \arg\max_S \sum_i \overline{u}_i(S)$ be the optimal strategy with respect to social welfare and let $\mathcal{S}^q$ be the set of all equilibrium strategies. We define the* **welfare regret at equilibrium (WRaE)** *for our game as: $sw(S^*) - \max_{S \in \mathcal{S}^q} sw(S)$.*

Known vs Unknown Utilities: Buyers may not know their expected net gain for a dataset until using it with their models. Thus, there are two sources of tension in the general setting: buyers behaving strategically with respect to their utility, and buyers learning these by interacting with the market. Although intertwined, the former relates to game dynamics and the latter is a learning theoretic question. To comprehensively study both, in section 4 we start with the simpler case of buyers knowing the expected utility for their actions, which is standard in game theory (Prisner, 2014; Munoz-Garcia et al., 2019; Mailath et al., 1991), and focus on the resulting equilibrium and welfare properties. The more general online setting wherein buyers learn their valuations through repeated market participation and behave strategically based on these is studied next. Our key metrics here are online effective and online welfare regret (see section 5 for precise definition); these intuitively capture how

well buyers can learn their optimal strategy, and how this learned strategy affects welfare regret at equilibrium.

4. Data Markets Game with Known Utility

In this section, we consider the simultaneous-move data markets game wherein each buyer knows their expected net gain and externality. This allows us to settle game/market dynamics questions first, before considering the more general setting of buyers playing with learned valuations in section 5. Considering this game without any platform intervention – i.e. buyers simply pay the seller's cost and $\mathcal{T}_i(S) = 0$ – admits a unique, dominant-strategy equilibrium: each buyer simply chooses the seller set with the highest net gain, since the externality suffered depends on the others' action. However, the welfare regret of this dominant-strategy equilibrium can be maximal. Intuitively, we sketch an instance where the seller set with the highest net gain also induces a high externality (proofs for this section are in Appendix B).

Proposition 1. *For the data markets game with no intervention - $\mathcal{T}_i(S) = 0$, $\forall i$, it is a dominant strategy for any buyer i to select $\gamma_i^d = \arg\max_\gamma g_i(\gamma)$. However, there exists an instance of this game where the WRaE is maximal - (i.e. $\Theta(n)$ upper and lower bound).*

This result illustrates the very real and unsatisfactory phenomenon that occurs in modern data markets. Buyers, not incentivized to care about their impact on others, make myopic decisions based purely on their own net gain, to the detriment of both the individual and the collective. This impact is exacerbated by the easy replicability of data, allowing multiple buyers to purchase the same dataset. To address this, we consider market intervention in the form of a transaction cost $\mathcal{T}_i(S)$ levied by the platform, a standard approach in online marketplaces. We posit that platforms can use learning models to estimate externalities(Definition 3) and propose they charge each buyer a cost proportional to the net difference in estimated externality they are contributing to and suffering from (Definition 4).

Definition 3 (Predicted Externality). *For any i, j, let $\widehat{E}_{ij}(\cdot)$ denote the predicted version of the externality $E_{ij}(\cdot)$, with $\mathbb{E}[\widehat{E}_{ij}(\cdot)] = \widehat{e}_{ij}(\gamma)$ and $b_{ij}(\cdot) = e_{ij}(\cdot) - \widehat{e}_{ij}(\cdot)$ representing the bias/error of the predictor. We assume predictions satisfy the same scale – i.e. $\sum_i \widehat{E}_{ji}(\cdot) \in [0, 1]$. As shorthand, we define $\Delta_b^{max} = \sum_{i \neq j} \max_{\gamma^1, \gamma^2} |b_{ij}(\gamma^1) - b_{ij}(\gamma^2)|$*

Definition 4 (Transaction Cost). *For a platform-chosen parameter $\alpha \in [0, 1]$, and a complete set of purchase orders S, we propose the platform charge each buyer i a transaction cost $\mathcal{T}_i(S) = \alpha \left(\sum_{j \neq i} \widehat{E}_{ji}(\gamma_i) - \sum_{j \neq i} \widehat{E}_{ij}(\gamma_j) \right)$, with $\alpha = 0$ corresponding to no intervention.*

Observe that we make no assumptions about the predictions $\widehat{E}_{ij}$ and characterize it generally through their item-

ized bias or expected inaccuracy b_{ij}. Secondly, while the proposed cost $\mathcal{T}_i(S)$ can be negative, from the platform's perspective this is just a redistribution and not an actual payment, since the sum of the transaction cost is always 0 - i.e. revenue-neutral. This may be desirable since the addition of this cost does not diminish the cumulative social welfare of the buyers. However, if platforms do wish to extract revenue they can charge an additional constant cost on top of this without changing any of our results. It is also advantageous that our proposed cost only depends on the externalities and not the net gain. We denote an *instance* of our game with this intervention by the tuple $\mathcal{I} = (g_i, e_{ij}, b_{ij}, \alpha)$. Lastly, the proposed transaction cost renders each buyer i's *effective utility*, the component of their utility that depends on their action γ_i, to have the following form: $G_i(\gamma_i) - \alpha \sum_j \widehat{E}_{ji}(\gamma_i) \triangleq U_i^e(\gamma_i)$, with $\mathbb{E}[U_i^e(\gamma)] \triangleq u_i^e(\gamma)$.

With this new cost in place, Theorem 1 gives a significantly improved result — a dominant strategy still exists, with each buyer maximizing their expected effective utility, but WRaE is tightly bounded by $\Theta(n(1 - \alpha))$, as the difference in bias parameters, Δ_b^{max} tend toward 0. This is a strong result since the upper bound implies that on *all* instances, equilibrium has welfare regret less than $n(1 - \alpha) + \Delta_b^{max}$. Crucially, this decreases linearly with α, with $\alpha \to 1$ leading to the equilibrium solution attaining the optimal social welfare, up to prediction inaccuracies. The corresponding lower bound, which is independent of the bias terms, shows that this is essentially tight.

Theorem 1. *Under the proposed transaction cost, it is a dominant strategy for any buyer i to select $\gamma_i^d = \arg\max_\gamma g_i(\gamma) - \alpha \sum_{j \neq i} \widehat{e}_{ji}(\gamma) = \arg\max_\gamma u_i^e(\gamma)$. Further, WRaE is upper bounded by $n(1 - \alpha) + \Delta_b^{max}$ and lower bounded by $n(1 - \alpha)$.*

Lastly, the nature of our transaction cost, which charges buyers based on externality, does indeed have spiritual similarities to the VCG family of mechanisms. There are however some very key differences. VCG mechanisms are designed for auctions wherein the platform elicits bids for items and then determines allocations. The cost to an agent in the VCG setting is the implied externality caused by that agent's participation. This fundamentally differs from the fixed-price market interaction we model, where any buyer may purchase from any seller by paying the set price. The platform does not control this. Thus we model externality explicitly; moreover, our transaction cost is the difference between the externality caused and the externality suffered. The explicit modeling of externalities also gives rise to a rich set of analysis based on its properties (this section vs section 6). So while there are spiritual similarities with VCG, they are fundamentally different.

5. Online setting with learned valuation

We now relax the assumption that buyers know their expected valuations a priori; instead, we consider them repeatedly interacting with the market and acting upon values learned during this process. This is consistent with real-world behavior since (1) sellers usually offer their data products for a prescribed time-period with buyers needing to pay anew for continued access (typical in AWS Data Marketplace) and (2) data products often capture real-time phenomenon (weather or foot traffic data) leading consumers to routinely purchase fresh data from the same source (see Appendix C for a detailed discussion). Buyers in this online setting face an exploration vs. exploitation problem, a dichotomy well modeled by the multi-armed bandit (MAB) framework. Each time a buyer interacts with the market, they can choose between 2^k "arms", representing the different seller combinations at their disposal. By "pulling an arm" they choose one of these options and observe the stochastic gain, externality, and transaction cost associated with that choice. At each round t, all buyers make such a decision and note that buying nothing, $\gamma = \emptyset$, is a valid strategy for any buyer at any time. The additional challenges here are twofold: proposing a buyer learning algorithm that works well over time given the exponential number of arms, and showing this buyer-optimal algorithm does not degrade the social welfare regret. Our goal in this section is to explore the possible learning algorithms available to buyers and the scenarios wherein they are performant.

We consider the platform charging the proposed transaction cost $\mathcal{T}_i(S)$ at each round as described in definition 4. To implement this, we consider the platform observing a possibly biased sample of the externalities each time the buyers make a decision. This sample can originate from leveraging learning algorithms to analyze downstream decision-making or outcomes for the firms involved. The platform charges transaction costs based on these sampled observations [3]. Note that under the proposed transaction cost, it is the dominant strategy of each buyer to maximize the expected effective utility, $u_i^e(\gamma) = \mathbb{E}[U_i^e(\gamma)]$, with γ_i^d being the maximizer, and $U_i^e(\gamma)$ referring to its stochastic analogue. Together, $U_i^e(\gamma)$ and $u_i^e(\gamma)$ are analogous to the stochastic and expected "reward" for an arm in the standard MAB setting. Using this, we now define appropriate notions of individual and collective regret for the online setting.

Definition 5. *For instance $\mathcal{I}$, the expected **online effective regret** for buyer i is the difference in the effective utility between their full information dominant strategy, γ_i^d, and their strategy at time t, γ_i^t:*

$R_d^i(T; \mathcal{I}) = \mathbb{E}\left[\sum_{t=1}^T U_i^e(\gamma_i^d) - U_i^e(\gamma_i^t)\right]$, *which is expanded as:* $R_d^i(T; \mathcal{I}) =$

$$\sum_{t=1}^T \left[g_i(\gamma_i^d) - g_i(\gamma_i^t) - \alpha \sum_{j \neq i} \widehat{e}_{ji}(\gamma_i^d) - \widehat{e}_{ji}(\gamma_i^t) \right]$$

Definition 6. *For an instance $\mathcal{I}$, the expected **online cumulative welfare regret** is the difference in total expected utility (across all agents) between the social welfare maximal strategy, $S^* = (\gamma_1^*, \ldots, \gamma_n^*)$, and the strategy taken at time t, $S^t = (\gamma_1^t, \ldots, \gamma_n^t)$. Since the proposed transaction cost is revenue-neutral, this regret be expressed as:* $R_w(T; \mathcal{I}) =$

$$\sum_{i=1}^N \sum_{t=1}^T \left[g_i(\gamma_i^*) - g_i(\gamma_i^t) - \sum_{j \neq i} e_{ij}(\gamma_j^*) - e_{ij}(\gamma_j^t) \right]$$

5.1. Algorithms

The exponential number of seller sets available to each buyer make learning a challenging problem. Using the forth-coming lemma 1, one can invoke the Upper Confidence Bound (UCB) algorithm to obtain $\widetilde{\mathcal{O}}(\sqrt{2^k T})$ effective regret for each buyer. To improve upon this, additional structure is needed. At first glance, the linear/combinatorial bandit framework, wherein an agent can simultaneously pull a subset k arms at each round, may seem attractive (Combes et al., 2015; Cesa-Bianchi & Lugosi, 2012). While this captures buyer interactions in our setting and reduces regret dependency from 2^k to k, it assumes a linear reward model. That is, an unknown vector $\boldsymbol{x} \in \mathbb{R}^k$ specifies the individual reward for each of the k options, and the reward for choosing a subset γ is: $\gamma^T \boldsymbol{x}$. In our setting, each of the k coordinates represents a different data seller, this becomes a strong assumption as it implies the net gain and externality experienced by buyers when purchasing from a bundle of sellers is uncorrelated and can be linearly decomposed. The utility of data products is argued to be richly correlated and this assumption may not hold (Agarwal et al., 2019).

With the linear utility assumption being too strong, we consider a weaker structure. γ can be represented as a k bit string, with bit i denoting inclusion of seller ℓ_i. Then for any γ^1 and γ^2, define $D_h(\gamma^1, \gamma^2)$ as the normalized Hamming distance (range is $[0, \ldots, 1]$, in increments of $1/k$) between the two, counting the number of sellers in which these sets differ. If $D_h(\gamma^1, \gamma^2)$ is close in Hamming distance, then the two seller sets consist of roughly the same sellers, and it is reasonable that the gain and externalities induced by these two will not be drastically different. We formalize this with the following *metric* property: for each buyer i and any pair (γ^1, γ^2): $|g_i(\gamma^1) - g_i(\gamma^2)| \leq \lambda_g D_h(\gamma^1, \gamma^2)$ and $|\sum_j e_{ji}(\gamma^1) - \sum_i e_{ji}(\gamma^2)| \leq \lambda_e D_h(\gamma^1, \gamma^2)$ [4]. We denote

---

[3] Our only assumption is that the outcome of this learning algorithm is stationary - i.e. the bias of the predictor remains the same over the rounds. Stationarity is a reasonable assumption to maintain market stability and consistency.

[4] WLOG, we fix $\lambda_g = \lambda_e = 1$ as it only changes bounds in a multiplicative manner

Algorithm 1 Zooming algorithm for buyer i

Active set $\mathcal{A}_i$ randomly initialized with a single seller set

Confidence radius for each $\gamma \in \mathcal{A}_i$ is $c_t^i(\gamma) = \sqrt{\frac{12 \log T}{n_t^i(\gamma)+1}}$

$\gamma' \in \mathcal{A}_i$ covers γ if $\lambda D_h(\gamma, \gamma') \leq c_t^i(\gamma')$

for $t = 1$ **to** T **do**

  if there is a γ uncovered by seller-sets in $\mathcal{A}_i$ **then**

    Pick any uncovered choice and append to $\mathcal{A}_i$

  end if

  Let $\text{UCB}_i(\gamma) = \overline{u}_i^e(\gamma) + 2c_t^i(\gamma)$ for each active γ

  Select arm with highest UCB_i value.

end for

$\Delta_i(\gamma_i) = u_i^e(\gamma_i^d) - u_i^e(\gamma_i)$ as the buyer i's expected effective utility gap for γ_i, with $\Delta_i(\gamma_i) \in [0, 2]$. $n_t(\gamma_i)$ denotes how often γ_i has been selected by buyer i. The metric property implies that $\Delta_i(\gamma_i^t) = u_i^e(\gamma_i^*) - u_i^e(\gamma^1) \leq 2D_h(\gamma^1, \gamma^2)$. Lastly, no assumptions are made *across* different buyers - i.e. no relation assumed between $g_i(\gamma^1)$ and $g_j(\gamma^2)$.

The metric assumption is far weaker than linearity as it still allows rich levels of correlation between options. A common technique in metric bandits is to uniformly cover the space of arms and UCB over the arms in the cover (Slivkins et al., 2019). This is practical when arms are in a continuous metric space since covering can be arbitrarily dense; in Hamming space however, any center in the covering is at least $\frac{1}{k}$ away from an arm it covers, leading to degenerate bounds. Kleinberg et al. (2019) propose a general purpose *zooming* algorithm, which adaptively discretizes a region proportional to its reward. It maintains a set of *active* arms, each of which *covers* any arm falling within its confidence radius. It ensures all arms are covered at every round, and selects from active arms using the UCB rule. Given the near-optimality of this algorithm in metric settings, we adapt this for our purpose (Kleinberg et al., 2019). While our proposed algorithm (given below) is similar to the original, obtaining good regret bounds requires more careful analysis. The value of γ depends on both the gain and externality which requires a stronger concentration result. More importantly, distance in Hamming space is quantized. This means that any two distinct elements are at least $\frac{1}{k}$ apart and there are a large number of elements that cannot be strictly compared. While the zooming algorithm in pathologically worst instances cannot improve upon standard algorithms, we show that in many natural instances, it performs significantly better. Our insights may be of independent interest in bandit settings with exponential arms. We present the algorithm below and continue the analysis thereafter.

5.2. Online Effective Regret

We bound the online effective regret of a buyer i under the zooming algorithm. First, using McDiarmid's inequality,

Lemma 1 shows sampled effective utility $\overline{u}_i^e$ is concentrated around its expected quantity u_i^e with high probability. Conditioned upon this *clean* event, Lemma 2 bounds the number of times an active choice is selected and the distance between any two active choices, mirroring a result in Kleinberg et al. (2019). Proofs for this section are in the Appendix D.

Lemma 1. *Let $\overline{u}_i^e(\gamma, h)$ denote the sample mean of $U_i^e(\gamma)$ with h samples, and define event $\mathcal{E} = \{\forall i, \forall \gamma, \forall t, |\overline{u}_i^e(\gamma, n_t^i(\gamma)) - u_i^e(\gamma)| \leq c_t^i(\gamma)\}$. Then $Pr[\mathcal{E}] \geq 1 - 1/T^2$.*

Lemma 2. *If $\mathcal{E}$ holds, then $n_t^i(\gamma) \leq \frac{108 \log T}{\Delta_i^2(\gamma)}$, $\forall \gamma, i, t$. For two active sets $\gamma^1, \gamma^2 \in \mathcal{A}_i$, $D_h(\gamma^1, \gamma^2) \geq \max(1/k, 1/3 \min(\Delta_i(\gamma^1), \Delta_i(\gamma^2)))$.*

With the following results, theorem 2 states a precise instance dependent regret bound for each buyer. To sketch this, all seller sets are partitioned based on how close their effective utility is to the optimal γ_i^d, and add up their contribution to the regret. For seller sets where the difference is large, we appeal to lemma 2 and use a packing argument to bound the number of active elements here. For any active γ, their contribution to the total regret for buyer i can be expressed using their gap $\Delta_i(\cdot)$ and the number of times they are chosen, $n_t^i(\cdot)$. For seller sets with utility close to the optimal, the discrete property of the Hamming metric makes lemma 2 too coarse, and the exact bound is instance-dependent. We address this in subsequent results.

Theorem 2. *For instance $\mathcal{I}$ and buyer i, let $S_i^{\mathcal{I}}(r/k)$ be the set of active sellers sets where each element $\gamma \in S_i^{\mathcal{I}}(r/k)$ satisfies $r/k \leq \Delta_i(\gamma) \leq 2r/k$. Let $\mathcal{C}(\cdot, d)$ denote the maximum packing of a region with balls of Hamming diameter d. Then for some $\delta > 0$, expected dominant strategy regret $R_d^i(T, \mathcal{I})$ is upper bounded by:*

$$\frac{2T\delta}{k} + 108k \log T \left(\sum_{j=-1}^{\log \frac{1}{\delta}} 2^j \left| S_i^{\mathcal{I}}\left(\frac{1}{2^j k}\right) \right| + \sum_{j=2}^{\log k} \frac{1}{2^j} |\mathcal{C}(S_i^{\mathcal{I}}(\frac{2^j}{k}), \frac{2^j}{3k})| \right)$$

Despite being stated from an instance-dependent perspective, this result is quite general and insightful. Observe that the worst case occurs when all seller sets have gap $\Delta_i(\gamma) < 1/k$; the Hamming metric is not useful here since its smallest non-zero value is $1/k$. In this pathological case, the regret and the algorithm devolve into UCB (Corollary 1). The algorithm shines, however, when the gap to the optimal is spread out evenly as a function of the Hamming distance (Theorem 3).

Corollary 1. *For any metric instance $\mathcal{I}$ where $\Delta_i(\gamma) \leq 1/k$ for all γ, $R_d^i(T; \mathcal{I}) \leq \mathcal{O}(\sqrt{2^k T \log T})$.*

Theorem 3. *For any metric instance $\mathcal{I}$ where $\Delta_i(\gamma) \in [2D_h(\gamma_i^d, \gamma) - 2/k, 2D_h(\gamma_i^d, \gamma)]$ [5] for all γ, $R_d^i(T; \mathcal{I}) \leq \mathcal{O}\left(k(\sqrt{kT \log T} + 2^{0.58k} \log T)\right)$*

The result above improves upon UCB, disentangling the dominant $\sqrt{T}$, from the exponential $2^{0.58k}$ term. The in-

---

[5] 2 is just a scale factor since $D_h(\cdot) \leq 1$ but $\Delta_i(\gamma) \in [0, 2]$

stance condition that led to this can be naturally interpreted: as seller sets γ become more distinct from the buyer's optimal γ_i^d, a quantity captured by the Hamming metric, the gain and externality of this diminishes accordingly. In other words, options close to the optimal have similar gain and externality, with those further away getting poorer. Note that γ_i^d need not be known, and this only assumes a relationship with respect to γ_i^d and not between any arbitrary arms. This condition can be viewed as a looser form of the "target-set" assumption in bandit literature (Slivkins et al., 2019; Kleinberg et al., 2019). In general, our goal here was to sketch out possible learning algorithms for buyers to learn their strategy, with the exponential set of options available being a challenge. While linear bandits provide the best regret guarantees, they require exact independence and linearity assumptions which may be unreasonable for data products. Correspondingly, we considered a weaker metric assumption and adapted a state-of-the-art algorithm for our unique Hamming metric setting. Our analysis shows that while in the worst case it cannot outperform UCB algorithm due to the discrete property of the Hamming metric, it can in more natural settings. Specifically, as gain and externality gap between an arbitrary seller set and buyer i's optimal becomes more correlated with their Hamming distance, covering becomes more effective, and regret drops. In general, the zooming algorithm allows us to smoothly interpolate the regret as the structural assumptions of the instance become looser or tighter. UCB and Linear bandits can be seen as extreme versions of these, with no assumptions and very strong assumptions respectively. Lastly, from an operational perspective, buyers rarely consider all k sellers in a marketplace; instead, they are usually deciding between a smaller subset of sellers. In such a case, a buyer only needs to learn valuations over this smaller set, and the regret bound becomes proportional to that instead of k.

5.3. Online welfare regret

We observed in theorem 1 that agents playing their dominant strategy under the proposed transaction cost leads to $\Theta(n(1-\alpha))$ welfare regret (WRaE). Having established that buyers can adopt online learning algorithms to learn their dominant strategy, we consider the analogous question in the online setting: what is the WRaE when agents play their dominant strategy using online algorithms? We prove (proof in appendix D) this can be decomposed into online effective regret discussed above and the offline WRaE (Theorem 1):

Theorem 4. *The online welfare regret $R_w(T;\mathcal{I})$ is upper-bounded by:* $2nT(1-\alpha) + 2T\Delta_b^{max} + nR_d(T;\mathcal{I})$.

We make a few remarks before concluding the online section. First, any horizon (T) dependent online algorithm for buyers can be converted to a horizon-independent version using the simple *doubling trick* (Cesa-Bianchi & Lugosi, 2006). This

entails running the algorithm in multiple phases, and in each phase i, an instance of the algorithm is executed for $T = 2^i$ rounds. This has another beneficial property if we consider adjusting the intervention parameter α over the phases as well. Observe that re-arranging the theorem above (ignoring the bias terms) to equate $nT(1-\alpha) = nR_d(T;\mathcal{I})$, yields $\alpha = \mathcal{O}\left(1 - \frac{R_d(T)}{T}\right)$. For any sub-linear learning algorithm, the ratio $\frac{R_d(T)}{T}$ is initially high but diminishes as the phases increase. In other words, for early phases when there is significant regret due to lack of learning, α can be small. In later phases, increasing α minimizes the WRaE that the dominant strategy causes. As such, the platform does not intervene strongly in early rounds when buyers know little and are simply exploring and only does so when buyers have learned. The online welfare regret under this schedule is asymptotically equivalent to the total effective regret faced by buyers. Overall, this is a nice operational property of our intervention and holds for any learning algorithm.

6. A Richer Externality Model

With a complete picture of both the learning and game dynamics, we now consider an extension by analyzing a richer externality model, termed *joint externality*. Under this externality model, when buyer i purchases γ_i and buyer j purchases γ_j, the externality buyer i suffers depends on both buyer i's decision, γ_i, and j's decision γ_j: $E_{ij}(\gamma_i, \gamma_j)$. To clarify notation, for two specific seller sets γ^1 and γ^2, $e_{ij}(\gamma^1, \gamma^2)$ implies buyer i owning γ^1 and buyer j owning γ^2. So in general, $e_{ij}(\gamma^1, \gamma^2) \neq e_{ij}(\gamma^2, \gamma^1)$, as the ownership is reversed. However, if we write $e_{ij}(\gamma_i, \gamma_j)$ the arguments already imply which buyer owns which set, and thus notationally $e_{ij}(\gamma_i, \gamma_j) = e_{ij}(\gamma_j, \gamma_i)$. The total expected externality suffered by buyer i from all other buyers is thus: $\sum_{j \neq i} e_{ij}(\gamma_i, \gamma_j)$.

Observe that this is a richer class of externality than previously studied and can capture more diverse settings (see appendix A for a detailed discussion). To our knowledge, this work is the first to explore this more general externality model in data markets. In this section, we focus on how this affects the game dynamics with and without our proposed intervention. This richer setting naturally leads to weaker guarantees, and as such we define a weaker notion of pure equilibrium. We then move on to our first pair of results (proofs for this section are in Appendix E) that paint a bleak picture of the Nash equilibrium properties under the joint externality model without any intervention: even a good approximate equilibrium may not always exist, and even in settings where it does, the welfare maybe very poor [6].

---

[6]The worst-case instance in Proposition 3 can't be captured by the independent model. Thus with no intervention, joint externality model does not aid the welfare properties of pure equilibrium.

Definition 7. *An action profile $S^q = (\gamma_1^q, \ldots, \gamma_n^q)$ is an ε- pure Nash Equilibrium if no agent i can increase their utility by more than ε by unilaterally deviating from their strategy γ_i^q. When $\varepsilon = 0$, this coincides with a PNE.*

Proposition 2. *Without any intervention, there are instances wherein no ε-pure Nash Equilibrium exists for $\varepsilon < 1$.*

Proposition 3. *Without any intervention, if a joint externality instance does have pure Nash Equilibria, then the WRaE of any such equilibria can approach $\rightarrow n$.*

These two results show that the lack of intervention is doubly bad in this richer externality model. Not only can pure equilibrium lead to maximal welfare regret, paralleling the result in theorem 1, but even good approximate equilibrium may not exist in this setting. We now consider introducing our transaction cost based on predicted externalities, which in this setting is equivalent to $\mathcal{T}_i(S) = \alpha(\sum_{j \neq i} \widehat{E}_{ji}(\gamma_i, \gamma_j) - \widehat{E}_{ij}(\gamma_i, \gamma_j))$; as before $\widehat{E}_{ij}(\gamma_i, \gamma_j)$ is the predicted externality, and $b_{ij}(\gamma_i, \gamma_j)$ is the bias. We first show that under this cost, an ε PNE is always possible, with $\varepsilon \rightarrow 0$ linearly as $\alpha \rightarrow 0.5$ and $\mathcal{O}(b) \rightarrow 0$, where $\mathcal{O}(b)$ denotes purely bias terms.

Theorem 5. *Under the proposed transaction cost, there always exists an ε PNE with $\varepsilon = 2|\alpha - 0.5| \sum_{j \neq i} \max_{\gamma_i, \gamma_j} |\widehat{e}_{ji}(\gamma_i, \gamma_j) - \widehat{e}_{ij}(\gamma_i, \gamma_j)| + \sum_{j \neq i} \max_{\gamma_i, \gamma_j} |b_{ij}(\gamma_i, \gamma_j) - b_{ji}(\gamma_i, \gamma_j)|$.*

Corollary 2. *For an instance $\mathcal{I} = (g_i, e_{ij}, b_{ij}, \alpha)$, let instance $\widetilde{\mathcal{I}} = (\widetilde{g}_i, \widetilde{e}_{ij}, \widetilde{b}_{ij}, \widetilde{\alpha})$, with $\widetilde{\alpha} = 0$, $\widetilde{g}_i = g_i$, $\widetilde{e}_{ij}(\gamma_i, \gamma_j) = \frac{1}{2}(e_{ij}(\gamma_i, \gamma_j) + e_{ji}(\gamma_i, \gamma_j))$, and $\widetilde{b}_{ij}(\cdot, \cdot) = 0$ for all possible values. Then for $\varepsilon = 2|\alpha - 0.5| \sum_{j \neq i} \max_{\gamma_i, \gamma_j} |\widehat{e}_{ji}(\gamma_i, \gamma_j) - \widehat{e}_{ij}(\gamma_i, \gamma_j)| + \sum_{j \neq i} \max_{\gamma_i, \gamma_j} |b_{ij}(\gamma_i, \gamma_j) - b_{ji}(\gamma_i, \gamma_j)|$, $\widetilde{\mathcal{I}}$ has a potential function whose exact PNE strategy is a ε PNE of $\mathcal{I}$.*

The ε in this approximate equilibrium is parameterized by the intervention parameter α, the instance externalities, and bias b_{ij}. With the latter bias component of ε diminishing linearly to 0 as bias gets closer to 0 or becomes symmetric, we focus on the first term. This linearly approaches 0 in two ways: either α approaches 0.5, or the externalities become symmetric (i.e. $e_{ij} \approx e_{ji}$)[7]. Overall, we find this result encouraging as it implies the maximum equilibrium deviation linearly reaches 0 either through instance property or α close to 0.5. Further, $\alpha \approx 0.5$ is not extreme but quite reasonable as it charges each buyer half of the predicted net externality they induce. Corollary 2 implies that one can obtain this approximate equilibrium by solving a corresponding potential game, which is operationally simple (Roughgarden, 2010). We next study the welfare regret of the equilibrium under our transaction cost and find the results to be positive here as well.

---

[7]While it might seem that it suffices for $\widehat{e}_{ij}$ to be symmetric, if the underlying externalities are not, the bias terms will not go to 0

Theorem 6. *For ε defined in theorem 5, there always exists a ε-PNE with WRaE is at most $\frac{n}{2}$.*

To summarize these results, joint externality is a richer model where the picture without any intervention is even more grim. Theorem 2, shows the existence of instances where even ε-PNE does not exist for $\varepsilon < 1$, the maximum possible utility. Theorem 3 shows even if PNE exists, the welfare regret is maximal. These reinforce the idea that equilibrium properties without intervention are poor. If however the externalities happen to be pairwise symmetric (externality faced by i due to j, is the same as faced by j due to i), then an ε pure equilibrium exists, with ε depending only on the bias terms (this can be gleaned from the proof of theorem 5). While this may not hold generally, it becomes a key insight when we introduce our transaction cost to this richer, since the combined externality and transaction cost exhibit a similar phenomenon - i.e. if the combined effect can be made symmetric, then pure equilibrium, up to bias terms, exists. The α parameter essentially controls this symmetry, with $\alpha = 0.5$ corresponding to achieving this combined symmetry perfectly, even if the externalities themselves are not symmetric. As α moves away from 0.5, the exactness of the equilibrium becomes slowly worse. Theorem formalizes this relation: $\varepsilon \rightarrow 0$ linearly as $\mathcal{O}(b) \rightarrow 0$ and $\alpha \rightarrow 0.5$. We also show in theorem 6 that the quality of the equilibrium achieved here is significantly better, with welfare regret less than $n/2$.

To compare against the standard externality setting (section 4), $\alpha = 1$ there represents the optimal parameter value with respect to welfare regret since an exact Nash equilibrium exists for all α. In the richer externality setting, $\alpha = 0.5$ represents the optimal parameter value to achieve an exact Nash equilibrium, since such an equilibrium may not exist in general. We also note that the WRaE of the richer setting roughly matches the tight WRaE bound for the independent externality setting (theorem 1) with $\alpha = 0.5$. This suggests that α close to 0.5 is a silver bullet: it imposes a reasonable cost, charging each buyer half of the net externality they induce, while having favorable equilibrium properties regardless of the externality model.

In Appendix F we experimentally validate our findings on the positive impact of our proposed transaction cost across both externality models. On a dataset derived from AWS Data Exchange, we show a significant increase in social welfare with increasing α for the standard externality model (Figure 1); crucially even a small α leads to a meaningful increase, with nearly maximal improvements achieved already at $\alpha = 0.6$. In the joint externality setting (Figure 2), we see social welfare improvements concentrated at $\alpha = 0.5$ as expected, but meaningful increases at other values as well. These results show that a range of α may lead to improved welfare, a significant operational advantage.

7. Discussion

Despite recent research on new market structures for data, real-world data markets remain far simpler: sellers post fixed prices and buyers are unfettered in their purchases. Our work fills a gap in the literature by studying this simple market under the unique characteristics of data: imperfect valuations, free replicability, and negative externality. The presence of externality allows us to naturally model buyer interactions as a simultaneous game. While this game has poor equilibrium properties by itself, a simple transaction cost can greatly improve these characteristics. For a standard externality model, our intervention is nearly perfect, guaranteeing a dominant strategy for buyers and leading to near-optimal social welfare. This intervention also fares well in the more realistic setting where buyers learn valuations through repeated interaction. We prove that buyers can learn to play their dominant strategy, while still achieving low social welfare regret. Lastly, we analyze an extension of this model by considering a richer class of externality. Although the equilibrium guarantees we can provide in this richer setting are naturally weaker, our proposed transaction cost significantly improves upon the status quo here as well.

Our work illustrates that when coupled with simple interventions, fixed-price data markets can be an elegant solution for a challenging product. It also leaves open several intriguing questions. We analyze a single transaction cost; it is unclear whether this is optimal or even what the space of such interventions is. Optimizing for non-utilitarian notions of welfare, like egalitarianism or Nash welfare is also an interesting research direction (Moulin, 2004). Our intervention leads to platforms learning or eliciting buyer externalities since transaction cost is a simple sum of these quantities. Designing truthful mechanisms or prediction algorithms for this is an important research question (Chen et al., 2020a; Kong et al., 2020; Zohar & Rosenschein, 2008). Similarly, understanding the long-term strategic perspective of sellers and their externality within this game will be insightful. Lastly, extending the joint externality model to the online setting with buyers playing based on learned valuations remains an interesting open problem. This is spiritually similar to work on learning and regret minimization in repeated games (Cesa-Bianchi & Lugosi, 2006; Slivkins et al., 2019).

Acknowledgements

Impact Statement

Access to data is a fundamental part of building performant machine learning models. This paper studies real-world data marketplaces and reaffirms that such markets, in their current form, do not generally maximize the welfare of buyers, which can lead to inequities. We propose a simple market intervention that can improve buyer welfare, even under self-interested behavior. Further research along this line can hopefully keep improving data market outcomes and be implemented in real marketplaces. Regarding specific technical contributions, there may be potential societal consequences, but none of which we feel must be specifically highlighted here.

References

Admati, A. R. and Pfleiderer, P. A monopolistic market for information. *Journal of Economic Theory*, 39(2): 400–438, 1986.

Admati, A. R. and Pfleiderer, P. Direct and indirect sale of information. *Econometrica*, 58(4):901–928, 1990.

Agarwal, A., Dahleh, M., and Sarkar, T. A marketplace for data: An algorithmic solution. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 701–726, 2019.

Agarwal, A., Dahleh, M., Horel, T., and Rui, M. Towards data auctions with externalities. *arXiv preprint arXiv:2003.08345*, 2020.

Aseff, J. and Chade, H. An optimal auction with identity-dependent externalities. *The RAND Journal of Economics*, 39(3):731–746, 2008.

Babaioff, M., Kleinberg, R., and Paes Leme, R. Optimal mechanisms for selling information. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pp. 92–109. ACM, 2012.

Bergemann, D. and Bonatti, A. Markets for information: An introduction. *Annual Review of Economics*, 11:85–107, 2019.

Bergemann, D., Bonatti, A., and Smolin, A. The design and price of information. *American Economic Review*, 108 (1):1–48, 2018.

Bergemann, D., Bonatti, A., and Gan, T. The economics of social data. *The RAND Journal of Economics*, 53(2): 263–296, 2022.

Bimpikis, K., Crapis, D., and Tahbaz-Salehi, A. Information sale and competition. *Management Science*, 65(6):2646–2664, 2019.

Blum, A., Kumar, V., Rudra, A., and Wu, F. Online learning in online auctions. *Theoretical Computer Science*, 324, 08 2003. doi: 10.1016/j.tcs.2004.05.012.

Bonatti, A., Dahleh, M., Horel, T., and Nouripour, A. Selling information in competitive environments. *arXiv preprint arXiv:2202.08780*, 2022.

Braverman, M., Mao, J., Schneider, J., and Weinberg, M. Selling to a No-Regret Buyer. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pp. 523–538, Ithaca NY USA, June 2018. ACM. ISBN 978-1-4503-5829-3. doi: 10.1145/3219166.3219233. URL https://dl.acm.org/doi/10.1145/3219166.3219233.

Bubeck, S., Stoltz, G., Szepesvári, C., and Munos, R. Online optimization in x-armed bandits. *Advances in Neural Information Processing Systems*, 21, 2008.

Cai, Y. and Velegkas, G. How to sell information optimally: an algorithmic study, 2020.

Cesa-Bianchi, N. and Lugosi, G. *Prediction, learning, and games*. Cambridge university press, 2006.

Cesa-Bianchi, N. and Lugosi, G. Combinatorial bandits. *Journal of Computer and System Sciences*, 78(5):1404–1422, 2012.

Chen, Y., Shen, Y., and Zheng, S. Truthful data acquisition via peer prediction. *Advances in Neural Information Processing Systems*, 33:18194–18204, 2020a.

Chen, Y., Xu, H., and Zheng, S. Selling information through consulting. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 2412–2431. SIAM, 2020b.

Combes, R., Talebi Mazraeh Shahi, M. S., Proutiere, A., et al. Combinatorial bandits revisited. *Advances in neural information processing systems*, 28, 2015.

Eso, P. and Szentes, B. The price of advice. *The Rand Journal of Economics*, 38(4):863–880, 2007.

Feng, S., Yu, F.-Y., and Chen, Y. Peer prediction for learning agents. In Oh, A. H., Agarwal, A., Belgrave, D., and Cho, K. (eds.), *Advances in Neural Information Processing Systems*, 2022. URL https://openreview.net/forum?id=JoukmNwGgsn.

Kleinberg, R., Slivkins, A., and Upfal, E. Bandits and experts in metric spaces. *Journal of the ACM (JACM)*, 66 (4):1–77, 2019.

Kong, Y., Schoenebeck, G., Tao, B., and Yu, F.-Y. Information elicitation mechanisms for statistical estimation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 2095–2102, 2020.

Li, M., Mei, L., Xu, Y., Zhang, G., and Zhao, Y. Facility location games with externalities. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1443–1451, 2019.

Mailath, G. J., Samuelson, L., and Swinkels, J. Extensive form reasoning in normal form games. *Foundations in Microeconomic Theory*, pp. 451, 1991.

Mehta, S., Dawande, M., Janakiraman, G., and Mookerjee, V. How to sell a dataset? pricing policies for data monetization. *Pricing Policies for Data Monetization (February 12, 2019)*, 2019.

Moulin, H. *Fair division and collective welfare*. MIT press, 2004.

Munoz-Garcia, F., Toro-Gonzalez, D., Munoz-Garcia, F., and Toro-Gonzalez, D. Pure strategy nash equilibrium and simultaneous-move games with complete information. *Strategy and game theory: Practice exercises with answers*, pp. 39–86, 2019.

Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V. *Algorithmic game theory*. Cambridge university press, 2007.

Pei, J. A survey on data pricing: from economics to data science. *IEEE Transactions on knowledge and Data Engineering*, 34(10):4586–4608, 2020.

Prisner, E. *Game theory through examples*, volume 46. American Mathematical Soc., 2014.

Roughgarden, T. Algorithmic game theory. *Communications of the ACM*, 53(7):78–86, 2010.

Schulz, A. S. and Moses, N. E. S. On the performance of user equilibria in traffic networks. In *SODA*, volume 3, pp. 86–87, 2003.

Slivkins, A. et al. Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning*, 12(1-2):1–286, 2019.

Weed, J., Perchet, V., and Rigollet, P. Online learning in repeated auctions. *CoRR*, abs/1511.05720, 2015. URL http://arxiv.org/abs/1511.05720.

Zohar, A. and Rosenschein, J. S. Mechanisms for information elicitation. *Artificial Intelligence*, 172(16-17): 1917–1939, 2008.

A. A discussion on Externality Models

The externality model we primarily focus on, which we also denote as independent externality, is the model considered in related literature (Aseff & Chade, 2008; Li et al., 2019; Agarwal et al., 2020). While simple, it is an appropriate and important model for many real-world settings. We give two different examples. Consider two firms acquiring data to improve their in-house models. The two firms are competitors and each firm's utility depends on the relative performance of their model with respect to their competitor. In this case, $g_i(\gamma_i)$ can represent the accuracy of firm i's model due to data purchases γ_i, and e_{ij} can simply be buyer i's approximation for the improvement achieved by their competitor j due to their data purchase γ_j: $\widehat{g}_j(\gamma_j)$. Note this externality is independent of i's decision, and firm i's utility (without transaction cost) is then $u_i(\gamma_i, \gamma_j) = g_i(\gamma_i) - \widehat{g}_j(\gamma_j)$. This is called additively separable externalities in (Agarwal et al., 2020), and we note that this is a common scenario when firms purchase data.

This model can also capture indirect competition. Consider two trading firms C_a and C_b that trade different financial products. C_a focuses on equities and C_b focuses on Forex. They acquire data to improve their respective performance and are targeting different data sources since they operate on different markets and products. However, they are competitors with respect to recruiting talent and securing new investments. If C_b purchases data it deems very prescient, it still negatively affects C_a, since a more profitable C_b can better compete in the talent and investment pool it shares with C_a. Thus C_b's action induces negative externality on C_a, independent of the C_a's data purchase decision (the gain from which is encapsulated in g_i).

The joint externality model is indeed more general and can capture the above examples. In these examples, the independent model can be seen as capturing indirect competition. However, it is possible that buyers might be competitors (and exert externality) purely with respect to data. Consider two firms in the same consumer industry, with consumer preferences or market data being sold. If both firm purchase the same data, they target the same group of consumers and eat into each other's profit. If the data they purchase don't overlap, they each target their respective groups and stay out of each other's way. In this setting, externality suffered by buyer i depends on both of their actions.

B. Section 4 Proofs

B.1. Proof of Proposition 1

Proof. Note each buyer i's utility under these conditions is: $g_i(\gamma_i) - \sum_{j \neq i} e_{ij}(\gamma_j)$. The only aspect of this utility that buyer i can affect is γ_i, and thus she has a dominant strategy of choosing the source with the highest gain [8]. It is intuitive that everyone adopting such a strategy will not always lead to good welfare. To see that there is an instance that achieves maximal WRaE, suppose the number of sellers and buyers are equal ($n = k$). For $k \in [1, \ldots, n]$, define γ^k as the seller set containing only seller ℓ_k. For a buyer i, let $g_i(\gamma^i) = 1$ and $g_i(\gamma) = 1 - \epsilon$ for any $\gamma \neq \gamma^i$. Further, for any buyer pair i, j, let $e_{ij}(\gamma_j = \gamma^j) = \frac{1}{n-1}$ and $e_{ij}(\gamma_j = \gamma) = 0 \, \forall \gamma \neq \gamma^j$. In this instance, the unique dominant strategy/equilibrium is each buyer i selecting γ^i, which results in 0 utility for all buyers. However, if each buyer i selects any other seller set aside from γ^i, they achieve utility $1 - \epsilon$ each. Since this is the only equilibrium in this instance, the WRaE is $n - n\epsilon \to n$ as $\epsilon \to 0$. $\square$

B.2. Proof of Theorem 1

Proof. The expected utility under the given transaction cost is given by: $g_i(\gamma_i) - \sum_{j \neq i} e_{ij}(\gamma_j) - \alpha \left(\sum_{j \neq i} \widehat{e}_{ji}(\gamma_i) - \sum_{j \neq i} \widehat{e}_{ij}(\gamma_j) \right)$. Since a buyer can only influence γ_i, the second and last terms can essentially be ignored, and the buyer has a dominant strategy, which is to maximize the expected effective utility, regardless of how others act. Regarding WRaE, we first prove the upper bound, before showing a specific instance achieves this upto bias terms.

Upper bound: For an instance $\mathcal{I}$, let $S^* = (\gamma_1^*, \gamma_2^*, \ldots, \gamma_n^*)$ denote the social welfare optimal, and let $S^d = (\gamma_1^d, \gamma_2^d, \ldots, \gamma_n^d)$ denote the dominant strategy taken by the buyers. The following is a direct implication of S^d being the dominant strategy for all buyers: $g_i(\gamma_i^d) - \alpha \sum_{j \neq i} \widehat{e}_{ji}(\gamma_i^d) \geq g_i(\gamma_i^*) - \alpha \sum_{j \neq i} \widehat{e}_{ji}(\gamma_i^*)$, which implies:

$$\sum_{j \neq i} (\widehat{e}_{ji}(\gamma_i^d) - \widehat{e}_{ji}(\gamma_i^*)) \leq \min \left(1, \frac{g_i(\gamma_i^d) - g_i(\gamma_i^*)}{\alpha} \right) \tag{1}$$

---

[8]Note that choosing the empty set and not participating is a valid strategy. Buyer i suffers externality regardless of what she chooses

Next, the social welfare expression (sum of all utilities) is unchanged due to transaction cost being revenue neutral; thus the social welfare difference between S^* and S^d is given by: $\sum_{i=1}^{n} g_i(\gamma_i^*) - \sum_{j \neq i} e_{ij}(\gamma_j^*) - g_i(\gamma_i^d) + \sum_{j \neq i} e_{ij}(\gamma_j^d)$ which is equal to:

$$\sum_{i=1}^{n} \sum_{j \neq i} \left(\widehat{e}_{ji}(\gamma_i^d) - \widehat{e}_{ji}(\gamma_i^*) \right) - \left(g_i(\gamma_i^d) - g_i(\gamma_i^*) \right) + \sum_{j \neq i} b_{ji}(\gamma_i^d) - b_{ji}(\gamma_i^*) \tag{2}$$

which holds since we are summing all possible externality pairs for seller sets γ^d and γ^*. Next, we apply inequality 1 to upper bound the expression above:

$$(2) \leq \Delta_b^{max} + \sum_{i=1}^{n} \min \left(1, \frac{g_i(\gamma_i^d) - g_i(\gamma_i^*)}{\alpha} \right) - \left(g_i(\gamma_i^d) - g_i(\gamma_i^*) \right)$$

$$= \Delta_b^{max} + \sum_{i=1}^{n} \min \left(1 - (g_i(\gamma_i^d) - g_i(\gamma_i^*)), (g_i(\gamma_i^d) - g_i(\gamma_i^*)) \left(\frac{1}{\alpha} - 1 \right) \right)$$

Putting aside the Δ_b^{max} term, each summand above is a min of two values, which is maximum when the two values are equal. Let $v_i = (g_i(\gamma_i^d) - g_i(\gamma_i^*))$. Thus we have $\forall i, 1 - v_i = \frac{v_i}{\alpha} - v_i \implies v_i = \alpha$, and thus we can upper bound this with: $\sum_{i=1}^{n} 1 - \alpha = n(1 - \alpha)$.

Lower bound: Consider a setting with only two sellers ℓ_1 and ℓ_2. For each buyer i, we have the following: $\forall i, g_i(\emptyset) = 0; g_i(\ell_1) = 0; g_i(\ell_2) = 1 - \alpha - \epsilon; g_i(\ell_1 \cap \ell_2) = 1$ and $\forall i, j, \ i \neq j, \ e_{ij}(\emptyset) = 0; e_{ij}(\ell_1) = 0; e_{ij}(\ell_2) = 0; e_{ij}(\ell_1 \cap \ell_2) = \frac{1}{n-1}$. Regarding the social welfare optimal solution, note that there is no reason to choose $\emptyset$ or ℓ_1; suppose S is such that k buyers choose $\ell_1 \cap \ell_2$ and $n - k$ choose ℓ_2. We thus have: $sw(S) = (n - k) \left[1 - \alpha - \epsilon - \frac{k}{n-1} \right] + k \left[1 - \frac{k}{n-1} \right]$ which is equal to $n - n\alpha - n\epsilon - k \left[\frac{n}{n-1} - \alpha - \epsilon \right]$. As $\epsilon \to 0$, social welfare is maximized when $k = 0$, since $\alpha \in [0, 1]$ and $\frac{n}{n-1} > 1$. Thus the optimal is all agents choosing ℓ_2, for a social welfare of $n(1 - \alpha)$ as $\varepsilon \to 0$. As for equilibrium/dominant strategy, observe that each buyer's dominant strategy is choosing $\ell_1 \cap \ell_2$, *regardless* of platform's bias. Choosing this option guarantees them at least $1 - \alpha$ utility whereas choosing $\gamma = \ell_2$ gives them at most $1 - \alpha - \varepsilon$ utility. However, this dominant strategy, however, has social welfare 0, leading to worst-case WRaE $\to n(1 - \alpha)$ on this instance. $\qquad \square$

C. Details on Repeated Data Market Interactions

In our online setting, we consider buyers repeatedly purchasing from a set of data sellers. This captures most real-world data market interactions due to two primary reasons. First, it is standard that buyers buy access for a limited time and need to pay to renew their access. In AWS Data Exchange for example, sellers offer their dataset for a fixed period of time. That is, when buyers purchase a dataset, they have access to it for this period and thereafter must pay to renew their access. The renewal/expiration period here is extremely standardized. Out of the roughly 1600 paid datasets for offer at AWS marketplace, the vast majority have a renewal/expiration period of 12 months. A second perspective behind repeated interaction is that sellers are often offering some sort of "real time" data, whose efficacy/usefulness degrades over time, leading consumers to purchase fresh data from them. This is indeed very common in for example consumer datasets (say foot-traffic data for some large retail chain), financial data (say most common equities purchased by non-institutional investors), or climate data. In such scenarios, it is natural for the seller to offer pay-per-use access to the data, which is refreshed periodically. For example, this weather dataset is pay-per-access and refreshed daily: snowflake-climate-data. The cost remains the same although the dataset sold by the seller changes daily. Note that in both settings (renewing subscription or getting fresh real-time values), it is natural that the price, expected gain, and externality are stationary.

D. Section 5 Proofs

D.1. Proof of Lemma 1

Proof. Consider a buyer i and a seller set γ. If it is not in the active set or has never been selected, the lemma trivially holds since $n_i^t(\gamma) = 0$, implying $c_t^i(\gamma) > 2$. If γ has been selected at least once, denote the ℓth realizations by y_i^ℓ, which consists of the sampled net gain and externality-based transaction cost. For the purpose of this analysis, we care only about the following realized parameters: $y_i^\ell = (g_i^\ell(\gamma), \widehat{e}_{1i}^\ell(\gamma), \widehat{e}_{2i}^\ell(\gamma), \ldots, \widehat{e}_{ni}^\ell(\gamma))$. Next, we can write the empirical mean of the effective utility from the first h times γ has been selected by buyer i as $\overline{u}_i^e(\gamma; h) = \overline{u}_i^e(y_i^1, \ldots, y_i^h) = \frac{1}{h}\sum_{\ell=1}^h \left(g_i^\ell(\gamma) - \alpha \sum_{j\neq i} \widehat{e}_{ji}^\ell(\gamma)\right)$. Observe that $\overline{u}_i^e(y_i^1, \ldots, y_i^h)$ satisfies the bounded difference property: $\forall \ell \in [1, \ldots, h], \forall y_i^\ell, \sup_{y_i'^\ell \in [0,1]^n} |\overline{u}_i^e(\ldots, y_i^\ell, \ldots) - \overline{u}_i^e(\ldots, y_i'^\ell, \ldots)| \leq \frac{2}{h}$. This allows us to apply McDiarmid's inequality on the function $\overline{u}_i^e$ over the random samples our h selections of γ results in, denoted by $Y_i^1, \ldots, Y_i^h$. We note that $\mathbb{E}[\overline{u}_i^e(Y_1, \ldots, Y_h)] = g_i(\gamma) - \alpha \sum_{j\neq i} \widehat{e}_{ji}(\gamma) = u_i^e(\gamma)$. Thus, by McDiarmid's inequality, we have:

$$\forall i, \forall \gamma, \forall h \ \mathbb{P}\left[|\overline{u}_i^e(\gamma, h) - u_i^e(\gamma)| \leq \sqrt{\frac{12\log(T)}{h+1}}\right] \geq 1 - \frac{2}{T^6} \tag{3}$$

Thus for all h, where h is the number of times buyer i selects γ, we have the bad event (no concentration) probability is $\frac{2}{T^6}$. Since $h \leq t$ and $n_t^i(\gamma) \leq t$, we can apply a union bound over all possible h and arrive at:

$$\forall i, \forall \gamma, \forall t \ \mathbb{P}\left[|\overline{u}_i^e(\gamma, n_t^i(\gamma)) - u_i^e(\gamma)| \leq \sqrt{\frac{12\log(t)}{n_t^i(\gamma)+1}}\right] \geq 1 - \frac{2}{T^5} \tag{4}$$

By applying a union bound over all $t \in [1, \ldots, T]$, we have that an event defined by $\mathcal{E}_i(\gamma) = \left\{\forall t \ |\overline{u}_i^e(\gamma, n_t^i(\gamma)) - u_i^e(\gamma)| \leq \sqrt{\frac{12\log(t)}{n_t^i(\gamma)+1}}\right\}$ holds with probability greater than $1 - \frac{2}{T^4}$. We would like to now show that for each buyer, this holds uniformly for all possible seller sets, and not just for a fixed γ. Since the total number of seller sets is exponential, a naive application of union bound provides a poor bound. However, since any inactive seller set trivially satisfies the bound with probability 1, it suffices to consider only the active seller sets for a buyer i, $\mathcal{A}_i$. While this set's size is bounded since we add at most one arm every round, it is random in its composition. For $j \in [1, \ldots, t]$, let Z_i^j denote the jth arm activated by buyer i. Z_i^j is a random variable and $\{Z_i^1, \ldots, Z_i^t\}$ is the set of all activated arms[9]. For a seller set γ, note that the event $\{Z_i^j = \gamma\}$ depends on the outcome of *previously activated sets*, whereas $\mathcal{E}_i(Z_i^j)$ is purely based on the observations derived from the seller set Z_i^j, whatever that happens to be. In other words, the event $\{Z_i^j = \gamma\}$ is independent of $\mathcal{E}_i(Z_i^j)$. Thus we have that $\forall i, \forall Z_i^j$, the clean event holds for the Z_i^j activated arm with the following probability:

$$\mathbb{P}\left[\mathcal{E}_i(Z_i^j)\right] = \sum_\gamma \mathbb{P}[\mathcal{E}_i(Z_i^j)|Z_i^j = \gamma]\mathbb{P}[Z_i^j = \gamma] = \sum_\gamma \mathbb{P}[\mathcal{E}_i(\gamma)]\mathbb{P}[Z_i^j = \gamma] \leq 1 - \frac{2}{T^4} \tag{5}$$

where we note that $\mathbb{P}[\mathcal{E}_i(\gamma)]$ is a constant and can be taken outside the sum. Thus, we have a concentration result for each active seller set. Now we apply a union bound over the whole active set $\mathcal{A}_i$. Noting that $|\mathcal{A}_i| \leq T$, we have that:

---

[9]If the number of activated arms is less than t, then for $j > |\mathcal{A}_i|$, let Z_i^j be the last arm activated

$\forall i, \mathbb{P}\left[\forall\, \gamma \in \mathcal{A}_i\,, \mathcal{E}_i(\gamma)\right] \geq 1 - \frac{2}{T^3}$. Lastly, note that the event $\mathcal{E} = \bigwedge_{i=1}^{n} \mathcal{E}_i$, and thus we apply a union bound over all the buyers. Assuming that the number of buyers is smaller than T, we arrive at the statement of the lemma. $\qquad\square$

D.2. Proof of Lemma 2

Proof. Fix any buyer i, any seller set γ, and a time t. If γ is not in the active set or never selected, then this claim holds trivially since $c_t^i(\gamma) > 2$, allowing us to consider only active arms. Suppose seller set γ was last chosen at some time $s \leq t$. Now consider the optimal dominant strategy for buyer i, γ_i^d, and the two possibilities at time s: (1) Either γ_i^d is already part of the active set, or (2) γ_i^d is covered by some other set $\gamma_i' \in \mathcal{A}_i$ which has confidence radius $c_s^i(\gamma_i') \geq \frac{1}{k}$ (the closest element to γ_i' has to be at least $\frac{1}{k}$ away). Then the following holds at time $s \leq t$ for each case:

$$\text{if (1): UCB}_i(\gamma) \geq \text{UCB}_i(\gamma_i^d) = \overline{u}_i^e(\gamma_i^d, n_t^i(s)) + 2c_s^i(\gamma_i^d) \geq u_i^e(\gamma_i^d)$$
$$\text{if (2): UCB}_i(\gamma) \geq \text{UCB}_i(\gamma_i') = \overline{u}_i^e(\gamma_i', n_t^i(s)) + 2c_s^i(\gamma_i') \geq u_i^e(\gamma_i') + c_s^i(\gamma_i') \geq u_i^e(\gamma_i^d) \tag{6}$$

where the last inequality in (1) follows from lemma 1 and the last inequality in (2) follows from the fact that γ_i^d is covered by γ_i' (i.e. $D_h(\gamma_i', \gamma_i^d) \leq c_t^i(\gamma_i')$) and thus their difference in utility is equivalently bounded by the closeness property. The following upper-bound for $\text{UCB}_i(\gamma)$ holds regardless:

$$\text{UCB}_i(\gamma) = \overline{u}_i^e(\gamma, n_t^i(s)) + 2c_s^i(\gamma) \leq u_i^e(\gamma) + 3c_s^i(\gamma) = u_i^e(\gamma) + 3c_t^i(\gamma) \tag{7}$$

where we use that fact that $c_s^i(\gamma) = c_t^i(\gamma)$ since s is the last time γ was selected. Putting the upper and lower bounds on UCB_i together, we have:

$$u_i^e(\gamma) + 3c_t^i(\gamma) \geq \text{UCB}_i(\gamma) \geq \text{UCB}_i(\gamma') \text{ or } \text{UCB}_i(\gamma^*) \geq u_i^e(\gamma_i^*) \implies \Delta_i(\gamma) \leq 3c_t^i(\gamma) \tag{8}$$

We now move to the second part of the lemma. First note that by property of the Hamming space, any two active seller sets (in fact any two seller sets) must be at least $\frac{1}{k}$ apart. Consider two active choices γ^1 and γ^2 for buyer i, and suppose γ^1 was activated (at time step t_1) before γ^2 (activated at time-step t_2). The fact that γ^2 was activated implies that it was not covered by γ^1's confidence radius at t_2, and thus $D_h(\gamma^1, \gamma^2) > c_{t_2}^i(\gamma^1) \geq \frac{\Delta_i(\gamma^1)}{3}$. If γ^2 was activated before γ^1, we get the opposite result. Combining the two, we have the $D_h(\gamma^1, \gamma^2) \geq \frac{1}{3}\min(\Delta(\gamma^1), \Delta(\gamma^2))$. Putting this and the Hamming observation of two elements being at least $\frac{1}{k}$ apart, we have the desired result. $\qquad\square$

D.3. Proof of Theorem 2

Proof. We consider the contribution of seller sets in $S_i^{\mathcal{I}}(\frac{r}{k})$ toward the total effective regret suffered by buyer i. By adding this contribution over all values of $\frac{r}{k}$ (with $r < k$), we can express the cumulative regret. For each active seller set $\gamma \in S_i^{\mathcal{I}}(\frac{r}{k})$, their contribution toward DS regret is $\Delta_i(\gamma) \cdot n_t^i(\gamma)$. Next, assume the clean event $\mathcal{E}$ holds. By lemma 2, we have that $\Delta_i(\gamma) \leq 3c_t(\gamma) = 3\sqrt{\frac{12\log T}{n_t^i(\gamma)+1}}$ which implies $n_t(\gamma) \leq \frac{108\log T}{\Delta_i^2(\gamma)}$. Thus, $\Delta_i(\gamma) \cdot n_t^i(\gamma) \leq \frac{108\log T}{\Delta_i(\gamma)} \leq \frac{108k\log T}{r}$. Now consider three ranges of $\frac{r}{k}$: (1) $\frac{r}{k} < \frac{\delta}{k}$, for some $\delta \in (0,1)$, (2) $\frac{r}{k} \in \left[\frac{\delta}{k}, \frac{2}{k}\right]$, and (3) $\frac{r}{k} \in \left[\frac{4}{k}, 1\right]$. For (1), we will use a trivial upper bound and will leave δ to be appropriately selected later. For (2), observe that by lemma 2, any two active seller sets satisfy, $D_h(\gamma^1, \gamma^2) \geq \frac{1}{3}\min(\Delta(\gamma^1), \Delta(\gamma^2)) = \frac{c}{3k}$, where $c \leq 2$. In other words, the lower bound for the distance between any active seller is $< \frac{1}{k}$ and thus in the worst case, each seller sets only covers itself. For the elements in (3) however, applying lemma 2 implies that any two active sets are at least $\frac{r}{3k}$ apart. The maximum number of active seller sets here is thus upper bounded by the maximal packing of the space $S_i^{\mathcal{I}}(\frac{r}{k})$ with balls of diameter $\frac{r}{3k}$. Thus, the cumulative regret can be expressed as follows (r is expressed as powers of two: $r = 2^j$).

$$R_d^i(T, \mathcal{I}) \leq \frac{\delta}{k}2T + \sum_{j=-1}^{\log\frac{1}{\delta}} 108k\log T k2^j \left|S_i^{\mathcal{I}}\left(\frac{1}{2^j k}\right)\right| + \sum_{j=2}^{\log k} 108k2^{-j}\log T|\mathcal{C}(S_i^{\mathcal{I}}(\frac{2^j}{k}), \frac{2^j}{3k})| \tag{9}$$

With this, we now match the bound in the theorem statement. Denote this by $R(\cdot|\text{clean})$, which captures the conditioning event. The bad event is lemma 1 not holding (probability $\frac{2}{T^2}$). Thus, the expected regret is dominated by $R(\cdot|\text{clean})$ since: $R_d^i(T, \mathcal{I}) \leq R(\cdot|\text{clean})(1 - \frac{2}{T^2}) + 2T\frac{2}{T^2}$. $\qquad\square$

D.4. Proof of Corollary 1

Proof. Since the gap $\Delta_i(\gamma) \leq \frac{1}{k}$, the third term in the expression given in theorem 2 will be 0, and we can focus on the first two terms. Next, we note that the second term, $\sum_{j=-1}^{\log \frac{1}{\delta}} k2^j 108 \log T \left| S_i^{\mathcal{I}} \left(\frac{1}{2^j k} \right) \right|$, is increasing in j and maximized when all elements belong to $S_i^{\mathcal{I}}(\delta/k)$. This can be expressed as: $S_i^I(\delta/k) = \sum_{j=-1}^{\log \frac{1}{\delta}} k2^j 108 \log T \left| S_i^{\mathcal{I}} \left(\frac{1}{2^j k} \right) \right| \leq k2^k \frac{1}{\delta} 108 \log T$. Going back to cumulative regret, we now have two terms increasing and decreasing in δ respectively. Their sum is minimized when we set them to be equal:

$$\frac{\delta}{k} 2T = k2^k \frac{1}{\delta} 108 \log T \implies 2\delta = \sqrt{\frac{k^2 2^k 108 \log T}{T}} \implies R_d^i(T) = \mathcal{O}\left(\sqrt{2^k T \log T} \right)$$

$\square$

D.5. Proof of Theorem 3

Proof. Starting with the expression given in theorem 2, we initially focus on the first two terms. Note that the second term considers seller sets whose gap is less than $\frac{4}{k}$. By given property of $\mathcal{I}$, seller sets who satisfy this are at most $\frac{3}{k}$ Hamming distance away. Thus, we can write the second term as: $\sum_{j=-1}^{\log 1/\delta} k2^j 108 \log T \left| S_i^{\mathcal{I}} \left(\frac{1}{2^j k} \right) \right|$ which we can upper bound as: $\leq \sum_{j=-1}^{\log 1/\delta} k2^j 108 \log T \sum_{i=0}^{3} \binom{k}{i} \leq 216 \log(T) k^4 \frac{1}{\delta}$. Comparing this with the $\frac{\delta}{k} 2T$ term, we now have an increasing and decreasing term in δ, which is minimized when we set them to be equal: $\frac{\delta}{k} 2T = k^4 \frac{1}{\delta} 216 \log T \implies \delta = \sqrt{\frac{k^5 216 \log T}{2T}}$. Note that δ is not an algorithm or instance parameter but used only for analysis. Putting all this together and using a standard assumption that $T > k$, we can state the first two term to be upper bounded by $\mathcal{O}(k^{3/2} \sqrt{T \log T})$.

We now turn our attention to the third and last term which considers regions $S_i^{\mathcal{I}}(\frac{r}{k})$ with $\frac{r}{k} \geq \frac{4}{k}$. Our instance condition implies the set of choices whose gap is in region $S_i^{\mathcal{I}}(\frac{r}{k})$ correspond to γ satisfying $D_h(\gamma_i^d, \gamma) \in [\frac{r}{2k}, \frac{r}{k} + \frac{1}{k}]$. We are looking to bound the maximal packing size of this region with balls of Hamming diameter $\frac{r}{3k}$. By a simple volume argument, we have that: $|\mathcal{C}(S_i^{\mathcal{I}}(\frac{r}{k}), \frac{r}{3k})| \leq \frac{\sum_{j=r/2}^{r} \binom{k}{j}}{\sum_{j=0}^{r/6} \binom{k}{j}}$. Next, note that we wish to upper bound the sum of the packing sizes of for each $S_i^{\mathcal{I}}(\frac{r}{k})$ region with $\frac{r}{k} \in [\frac{4}{k}, \frac{8}{k}, \dots, \frac{1}{2}, 1]$. These correspond to seller sets being in the following Hamming distance ranges from γ_i^d: $[[\frac{2}{k}, \frac{5}{k}], [\frac{4}{k}, \frac{9}{k}] \dots, [\frac{1}{4}, \frac{1}{2} + \frac{1}{k}], [\frac{1}{2}, 1]]$. We observe that the packing size upper bound (given above) is increasing in r up to $r = \frac{1}{2}$, which we express as:

$$|\mathcal{C}(S_i^{\mathcal{I}}(\tfrac{1}{2}), \tfrac{1}{6})| \leq \frac{\sum_{j=k/4}^{k/2} \binom{k}{j}}{\sum_{j=0}^{k/12} \binom{k}{j}} \leq \frac{2^{k-1}}{2^{H\left(\frac{1}{12}\right)k}} \leq 2^{0.58k} \tag{10}$$

where $H(\cdot)$ is the binary entropy function. Now we can sum over all the values of $\frac{r}{k}$ under consideration and express this as follows: $108k \log T \sum_{j=2}^{\log k} 2^{-j} |\mathcal{C}(S_i^{\mathcal{I}}(\frac{2^j}{k}), \frac{2^j}{3k})| \leq 108k 2^{0.58k} \log T$. Putting this together with the bound for the first two terms gives us the desired result. $\square$

D.6. Proof of Theorem 4

Proof. Let $S^* = (\gamma_1^*, \dots, \gamma_n^*)$ be the social optimal strategy, $S^d = (\gamma_1^d, \dots, \gamma_n^d)$ the dominant strategy of each buyer, and $S^t = (\gamma_1^t, \dots, \gamma_n^t)$ the strategy taken by buyers at time t. We can add and subtract the welfare at the dominant strategy to the online welfare regret, $R_w(T)$, and express it as:

$$R_w(T) = \sum_{t=1}^{T} \sum_{i=1}^{N} \left(g_i(\gamma_i^*) - \sum e_{ij}(\gamma_j^*) - g_i(\gamma_i^d) + \sum e_{ij}(\gamma_j^d) \right) + \left(g_i(\gamma_i^d) - \sum e_{ij}(\gamma_j^d) - g_i(\gamma_i^t) + \sum e_{ij}(\gamma_j^t) \right)$$

The first part of the sum, $\sum_{i=1}^{N} g_i(\gamma_i^*) - \sum e_{ij}(\gamma_j^*) - g_i(\gamma_i^d) + \sum e_{ij}(\gamma_j^d)$, is exactly the WRaE quantity that we upper bounded to $n(1 - \alpha) + \Delta_b^{max}$ in theorem 1, which we can then sum over T. Next, observe that $e_{ji}(\gamma) = \alpha \widehat{e}_{ji}(\gamma) + (1 -$

$\alpha)\widehat{e}_{ji}(\gamma) + b_{ji}(\gamma)$ for any γ. Thus, we can re-arrange the latter sum as:

$$\sum_{i=1}^{N} \left[\sum_{t=1}^{T} g_i(\gamma_i^d) - \sum_{t=1}^{T}\sum_{j\neq i} \alpha\widehat{e}_{ji}(\gamma_i^d) - \sum_{t=1}^{T} g_i(\gamma_i^t) + \sum_{t=1}^{T}\sum_{j\neq i} \alpha\widehat{e}_{ji}(\gamma_i^t) \right] \tag{11}$$

$$+T\sum_{j\neq i} [b_{ji}(\gamma_i^t) - b_{ji}(\gamma_i^d)] + T(1-\alpha)\sum_{j\neq i} [e_{ji}(\gamma_i^t) - e_{ji}(\gamma_i^d)] \tag{12}$$

We note that the first half of the summand is equal to online effective regret for all buyers, and is equivalent to $nR_d(T; \mathcal{I})$. The bias terms are upper bounded by $T\Delta_b^{max}$ and since total externality induced by any buyer is at most 1, the last summation term is upper bounded by $nT(1-\alpha)$. Thus, the total online welfare regret is upper-bounded by $2nT(1-\alpha) + 2T\Delta_b^{max} + nR_d(T; I)$. □

E. Section 6 proofs

Joint externality can be represented in matrix form. For a pair i, j with $i < j$, the matrix $\mathcal{E}_{ij}$ tabulates the externality suffered by i due to j with $\mathcal{E}_{ij}[\ell, k] = e_{ij}(\gamma_i = \gamma^\ell, \gamma_j = \gamma^k)$; similarly, $\mathcal{E}_{ji}[\ell, k] = e_{ji}(\gamma_i = \gamma^\ell, \gamma_j = \gamma^k)$. We leverage this notation in the following proofs.

E.1. Proof of Proposition 2

Proof. Consider 2 buyers, and let γ^a and γ^b denote two arbitrary orders with $a \in \{1, \ldots, 2^k\}$, $b \in \{1, \ldots, 2^k\}$ and $a \neq b$. Then, let the externality involving γ^a and γ^b be as follows (note $\mathcal{E}_{12}^{ab}$ is a sub-matrix of $\mathcal{E}_{12}$):

$$\mathcal{E}_{12}^{ab} = \begin{bmatrix} e_{12}(\gamma_1 = \gamma^a, \gamma_2 = \gamma^a) = 1 & e_{12}(\gamma_1 = \gamma^a, \gamma_2 = \gamma^b) = 0 \\ e_{12}(\gamma_1 = \gamma^b, \gamma_2 = \gamma^a) = 0 & e_{12}(\gamma_1 = \gamma^b, \gamma_2 = \gamma^b) = 1 \end{bmatrix}$$

$$\mathcal{E}_{21}^{ab} = \begin{bmatrix} e_{21}(\gamma_1 = \gamma^a, \gamma_2 = \gamma^a) = 0 & e_{21}(\gamma_1 = \gamma^a, \gamma_2 = \gamma^b) = 1 \\ e_{21}(\gamma_1 = \gamma^b, \gamma_2 = \gamma^a) = 1 & e_{21}(\gamma_1 = \gamma^b, \gamma_2 = \gamma^b) = 0 \end{bmatrix}$$

Let the expected externality e_{12} and e_{21} be 1 for any other seller set pair not included above — that is, the full externality matrices for this example, $\mathcal{E}_{12}$ and $\mathcal{E}_{21}$, are rank 2. Next, let the expected net gain be as follows: $g_i(\gamma^a) = g_i(\gamma^b) = 1$ for $i \in \{1, 2\}$, with the expected net gain for all other seller sets $\gamma \notin \{\gamma^a, \gamma^b\}$ being 0. This clearly implies that it is strictly dominant for both buyers to choose between γ^a or γ^b over any other choices as that leads to -1 utility. Thus, it suffices to consider the utility for both buyers involving only these states which we can express in a matrix as follows (buyer 1 is the row player, and buyer 2 is the column player):

$$\begin{bmatrix} u_1(\gamma^a, \gamma^a), u_2(\gamma^a, \gamma^a) & u_1(\gamma^a, \gamma^b), u_2(\gamma^a, \gamma^b) \\ u_1(\gamma^b, \gamma^a), u_2(\gamma^b, \gamma^a) & u_1(\gamma^b, \gamma^b), u_2(\gamma^b, \gamma^b) \end{bmatrix} = \begin{bmatrix} 0, 1 & 1, 0 \\ 1, 0 & 0, 1 \end{bmatrix}$$

It is evident that for any action profile γ_1, γ_2 involving a and b, one of the buyers will benefit an amount 1 by deviating. Since any other strategy involving the buyers choosing something beyond options a, b is strictly dominated also by utility 1, the statement holds. □

E.2. Proof of Proposition 3

Proof. Consider n buyers, and let γ_a and γ_b denote two arbitrary orders with $a \in \{1, \ldots, 2^k\}$, $b \in \{1, \ldots, 2^k\}$ and $a \neq b$. Buyers 1 through $n - 1$ share the same externality whereas buyer n is different. Let the externality involving options γ^a and γ^b be as follows:

$$\forall i \in \{1, n-1\}, \forall j \; : \mathcal{E}_{ij}^{ab} = \begin{bmatrix} e_{ij}(\gamma_1 = \gamma^a, \gamma_2 = \gamma^a) = \frac{1-\varepsilon}{n-1} & e_{ij}(\gamma_1 = \gamma^a, \gamma_2 = \gamma^b) = \frac{\varepsilon}{n-1} \\ e_{ij}(\gamma_1 = \gamma^b, \gamma_2 = \gamma^a) = \frac{1-\varepsilon}{n-1} & e_{ij}(\gamma_1 = \gamma^b, \gamma_2 = \gamma^b) = \frac{2\varepsilon}{n-1} \end{bmatrix}$$

$$\forall j \; : \mathcal{E}_{nj}^{ab} = \begin{bmatrix} e_{ij}(\gamma_1 = \gamma^a, \gamma_2 = \gamma^a) = \frac{1-\varepsilon}{n-1} & e_{ij}(\gamma_1 = \gamma^a, \gamma_2 = \gamma^b) = \frac{1}{n-1} \\ e_{ij}(\gamma_1 = \gamma^b, \gamma_2 = \gamma^a) = \frac{1}{n-1} & e_{ij}(\gamma_1 = \gamma^b, \gamma_2 = \gamma^b) = 0 \end{bmatrix}$$

Let the expected externality for all (i, j) be 1 for any other seller set pair not including either γ^a or γ^b. Next, let the expected net gain be as follows: $g_i(\gamma^a) = g_i(\gamma^b) = 1$ for $i \in \{1, 2\}$, with the expected net gain for all other seller sets $\gamma \notin \{\gamma^a, \gamma^b\}$ being 0. As such, it is strictly dominant for *any* buyer to select γ^a or γ^b over any other since the former will lead to a utility of at least 0, whereas the latter has utility -1. Thus it suffices to consider strategies pertaining only to γ^a and γ^b for the purpose of equilibrium. Next, note that for agents 1 through $n - 1$, choosing γ^a is in fact strictly dominant since they suffer less externality as compared to choosing γ^b, regardless of whether buyer n choosing γ^a or γ^b. Since the first $n - 1$ agents will always choose γ^a as their dominant strategy, the only pure equilibrium consists of buyer n choosing γ^a as this has strictly lower utility. Thus, all agents choosing γ^a is the only PNE, and we note that this has social welfare $n\varepsilon$. Next, consider all buyers choosing γ^b. This leads to social welfare $1 + (1 - 2\varepsilon)(n - 1) = n - 2\varepsilon(n - 1)$. Thus as $\varepsilon \to 0$, the best case WRaE $\to n$. □

E.3. Proof of Theorem 5

Proof. For a general instance $\mathcal{I} = (g_i, e_{ij}, b_{ij}, \alpha)$, first consider a simpler version of this instance, $\widetilde{\mathcal{I}} = (\widetilde{g}, \widetilde{e}_{ij}, \widetilde{b}_{ij}, \widetilde{\alpha})$ where $\widetilde{g}_i = g_i$, $\widetilde{e}_{ij}(\gamma_i, \gamma_j) = \frac{1}{2}(e_{ij}(\gamma_i, \gamma_j) + e_{ji}(\gamma_i, \gamma_j))$, $\widetilde{\alpha} = 0$, and $\widetilde{b}_{ij}(\cdot, \cdot) = 0$ for all i, j and possible pairs of

(γ_i, γ_j). We first note that in instance $\widetilde{\mathcal{I}}$, externality are symmetric - i.e. $\widetilde{e}_{ij}(\gamma_i, \gamma_j) = \widetilde{e}_{ji}(\gamma_i, \gamma_j) \, \forall i, j$. We now show that the dynamics of instance $\widetilde{I}^s$ can be represented through a potential function $\Phi(S)$ defined as follows: $\Phi(S) = \sum_i \widetilde{g}_i(\gamma_i) - 0.5 \sum_i \sum_j \widetilde{e}_{ij}(\gamma_i, \gamma_j)$. Without loss of generality, for a joint action profile $S = (\gamma_1, \ldots, \gamma_n)$, consider buyer 1 changing her action to γ_1' and leading to profile S'. The resulting change in expected utility for her is given by $\widetilde{g}_1(\gamma_1) - \widetilde{g}_1(\gamma_1') - \left(\sum_{j \neq 1} \widetilde{e}_{1j}(\gamma_i, \gamma_j) - \widetilde{e}_{1j}(\gamma_i', \gamma_j) \right)$. Note that the potential function difference is given by $\Phi(S) - \Phi(S') = \widetilde{g}_1(\gamma_1) - \widetilde{g}_1(\gamma_1') - 0.5 \sum_{j \neq 1} \widetilde{e}_{1j}(\gamma_1, \gamma_j) - 0.5 \sum_{j \neq 1} \widetilde{e}_{j1}(\gamma_1, \gamma_j) + 0.5 \sum_{j \neq 1} \widetilde{e}_{1j}(\gamma_1', \gamma_j) - 0.5 \sum_{j \neq 1} \widetilde{e}_{j1}(\gamma_1', \gamma_j)$. Since externalities are symmetric in $\widetilde{I}$, this difference in potential functions is equivalent to the difference in agent 1's utility, making this a potential game, which implies game instance $\widetilde{I}$ has a pure Nash Equilibrium.

Let $S = (\gamma_1, \ldots, \gamma_n)$ denote an action profile. Under $\widetilde{\mathcal{I}}$, observe that utility for buyer i can be expressed as $u_i(S; \widetilde{\mathcal{I}}) = g_i(\gamma_i) - 0.5 \sum_{j \neq i} e_{ij}(\gamma_i, \gamma_j) - 0.5 \sum_{j \neq i} e_{ji}(\gamma_i, \gamma_j)$. Observe that for any strategy S, we can relate $u_i(S; \widetilde{\mathcal{I}})$ to the utility for buyer i in the general instance $\mathcal{I}$ under the same actions S as follows: $u_i(S; \widetilde{\mathcal{I}}) - u_i(S; \mathcal{I})$

$$= (\alpha - 0.5) \sum_{j \neq i} [\widehat{e}_{ji}(\gamma_i, \gamma_j) - \widehat{e}_{ij}(\gamma_i, \gamma_j)] + 0.5 \sum_{j \neq i} b_{ij}(\gamma_i, \gamma_j) - b_{ji}(\gamma_i, \gamma_j) \tag{13}$$

where $\widehat{e}_{ij}$ is the predicted externality under the true instance $\mathcal{I}$. Next, let $S^q = (\gamma_1^q, \ldots, \gamma_n^q)$ denote an equilibrium profile under simplified instance $\widetilde{\mathcal{I}}$. This implies for any buyer i unilaterally deviating from γ_i^q to another γ_i (we denote the resulting action profile S_{-i}^q) is not beneficial under $\widetilde{\mathcal{I}}$: $u_i(S^q; \widetilde{I}) \geq u_i(S_{-i}^q; \widetilde{I})$. Then by appealing to equation 13, we state that $u_i(S_{-i}^q; \mathcal{I}) - u_i(S^q; \mathcal{I})$, the deviation benefit, is bounded by: $\leq 2|\alpha - 0.5| \sum_{j \neq i} \max_{\gamma_i, \gamma_j} |\widehat{e}_{ij}(\gamma_i, \gamma_j) - \widehat{e}_{ji}(\gamma_i, \gamma_j)| + \sum_{j \neq i} \max_{\gamma_i, \gamma_j} |b_{ij}(\gamma_i, \gamma_j) - b_{ji}(\gamma_i, \gamma_j)|$. Thus, S^q is an ε equilibrium for the original instance $\mathcal{I}$ for the states ε values. $\qquad \square$

E.4. Proof of Corollary 2

This follows immediately from the above theorem: Theorem 5

E.5. Proof of Theorem 6

Proof. We consider the corresponding simpler version of instance $\mathcal{I}$, $\widetilde{\mathcal{I}} = (\widetilde{g}_i, \widetilde{e}_{ij}, \widetilde{b}_{ij}, \widetilde{\alpha})$, defined in corollary 2. We note that externality in $\widetilde{\mathcal{I}}$ is always symmetric and by corollary 5, a PNE always exists for $\widetilde{\mathcal{I}}$, which implies a ε-PNE for instance $\mathcal{I}$, where ε is equal to the one in the theorem statement. Thus, it suffices to consider the WRaE of instance $\widetilde{I}$ for the remainder of the proof.

Let the social optimal be $S^* = (\gamma_1^*, \ldots, \gamma_n^*)$, with each buyer attaining utility $\widetilde{u}_i(S^*) = \widetilde{g}_i(\gamma_i^*) - \sum_{j \neq i} \widetilde{e}_{ij}(\gamma_i^*, \gamma_j^*)$ for a total welfare of $sw(S^*)$. If the social optimal is an equilibrium, then the best-case welfare regret is 0. If not, then by theorem 5, we know that a pure equilibrium of $\widetilde{\mathcal{I}}$ can be reached by playing the sequential best response (buyers play the best response one by one). Without loss of generality, suppose buyer 1 is unhappy at the social optimal and changes their decision from γ_1^* to γ_1^1 (denotes that this is buyer 1's first best response), and define this new state as $S_{1,1}$ (denotes buyer 1 has played their first best response). Thus:

$$\widetilde{g}_1(\gamma_1^1) - \sum_{j \neq 1} \widetilde{e}_{1j}(\gamma_1^1, \gamma_j^*) > \widetilde{g}_1(\gamma_1^*) - \sum_{j \neq 1} \widetilde{e}_{1j}(\gamma_1^*, \gamma_j^*) \tag{14}$$

Define $\Delta \widetilde{g}_i(\gamma_i^t, \gamma_i) = \widetilde{g}_i(\gamma_i^t) - \widetilde{g}_i(\gamma_i)$ and $\Delta \widetilde{e}_{ij}(\gamma_i^t, \gamma_i, \gamma_j) = \widetilde{e}_{ij}(\gamma_i^t, \gamma_j) - \widetilde{e}_{ij}(\gamma_i, \gamma_j)$. With this new notation, we can express equation 14 as $\Delta \widetilde{g}_1(\gamma_1^1, \gamma_1^*) > \sum_{j \neq 1} \Delta \widetilde{e}_{1j}(\gamma_1^1, \gamma_1^*, \gamma_j^*)$ and exploiting symmetric externality of $\widetilde{I}^s$, the social optimal at $S_{1,1}$ can be succinctly expressed as $sw(S_{1,1}) = sw(S^*) + \Delta \widetilde{g}_1(\gamma_1^1, \gamma_1^*) - 2 \sum_{j \neq 1} \Delta \widetilde{e}_{1j}(\gamma_1^1, \gamma_1^*, \gamma_j^*)$ (since externalities are symmetric, buyer 1's new decision affects externality $\widetilde{e}_{1j}$ and $\widetilde{e}_{j1}$ equally). This relationship is in fact satisfied between any two consecutive best response steps: let state $S_{i,k}$ denote when buyer i plays her k^{th} best response, which occurs right after state $S_{h,\ell}$ where buyer h plays her ℓ^{th} best response (note that not every buyer needs to update at every round). Also, for any buyer j, let t_j represent the number of times they have played best response up to time t (i.e. $t_h = \ell$). The following two invariants hold:

$$\Delta \widetilde{g}_i(\gamma_i^k, \gamma_i^{k-1}) > \sum_{j \neq i} \Delta \widetilde{e}_{ij}(\gamma_i^k, \gamma_i^{k-1}, \gamma_j^{t_j}) \tag{15}$$

$$sw(S_{i,k}) = sw(S_{h,\ell}) + \Delta \widetilde{g}_i(\gamma_i^k, \gamma_i^{k-1}) - 2 \sum_{j \neq i} \Delta \widetilde{e}_{ij}(\gamma_i^k, \gamma_i^{k-1}, \gamma_j^{t_j}) \tag{16}$$

Let $S^q = (\gamma_i^q, \ldots, \gamma_n^q)$ be the pure equilibrium state reached by this best response cycle starting from social optimal. By repeatedly applying the two invariants above to $sw(S^*)$ and expanding, we can relate $sw(S^q)$ to $sw(S^*)$ as follows:

$$sw(S^q) = sw(S^*) + \sum_{i=1}^{n} \Delta \widetilde{g}_i(\gamma_i^q, \gamma_i^*) - 2 \sum_{i=1}^{n} \sum_{j \neq i} \Delta \widetilde{e}(\gamma_i^q, \gamma_i^*, \gamma_j^*) \tag{17}$$

Observe that summing the first invariant (equation 15) across n implies that:

$$\sum_{i=1}^{n} \Delta \widetilde{g}_i(\gamma_i^q, \gamma_i^*) \geq \sum_{i=1}^{n} \sum_{j \neq i} \Delta \widetilde{e}_{ij}(\gamma_i^q, \gamma_i^*) \tag{18}$$

Let $\sum_i \sum_{j \neq i} \Delta \widetilde{e}_{ij}(\gamma_i^e, \gamma_i^*) \triangleq c$, and note that $c \in [0, n]$ due to the boundedness of $\sum_{ij} \widetilde{e}_{ij}$. Thus we have: $\sum_{i=1}^{n} \Delta \widetilde{g}_i(\gamma_i^e, \gamma_i^*) \geq c \implies sw(S^*) \leq n - c$ since maximum net gain for any buyer is 1. Thus, WRaE $\leq n - c$. Next, observe by equation 17, $sw(S^q) \geq sw(S^*) + c - 2c = sw(S^*) - c \implies$ WRaE $\leq c$. Combining these, we have that WRaE $= \min(c, n - c) \leq \frac{n}{2}$. $\qquad\square$

F. Experimental Results

We experimentally validate our theoretical insights on the effectiveness of transaction costs. While no publicly available dataset exists for data markets with prices and utilities, we take inspiration from AWS marketplace (a fixed price data market platform) we design a suite of synthetic experiments. We take the 10 different data categories from AWS, and instantiate several data sellers for each category (177 total sellers). For each category, we have several buyers (57 total), where each buyer has zero gain/externality for the sellers *not* in their category, with gains and externalities for their category sampled uniformly. They may purchase from up to 10% of the sellers in their category. This is meant to model the pricing and budget constraints of buyers.
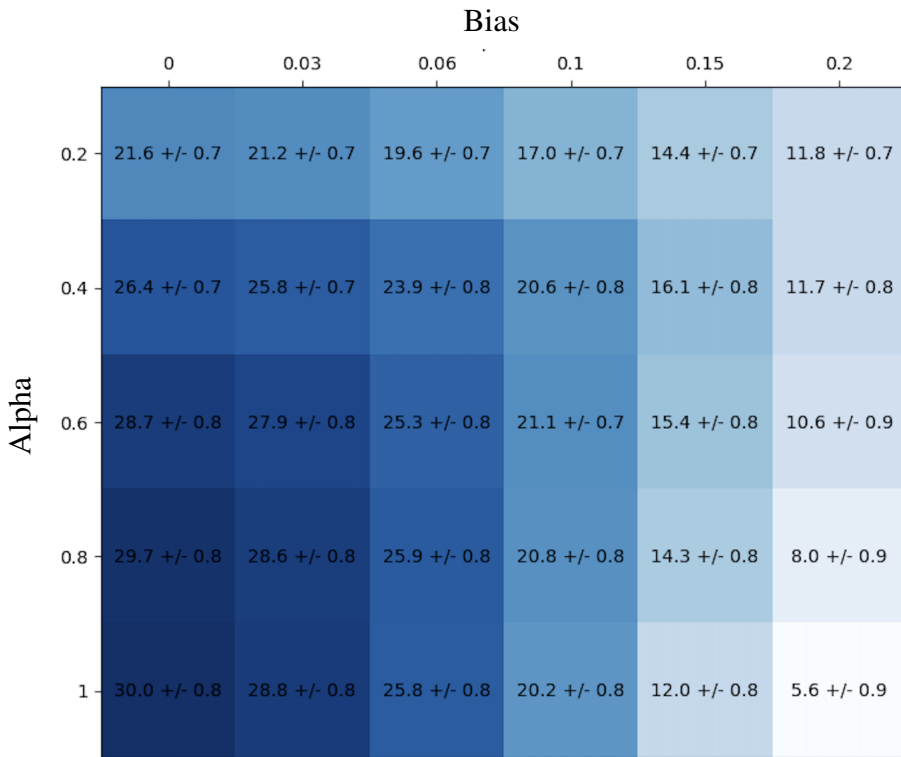

Bias

	0	0.03	0.06	0.1	0.15	0.2
0.2	21.6 +/- 0.7	21.2 +/- 0.7	19.6 +/- 0.7	17.0 +/- 0.7	14.4 +/- 0.7	11.8 +/- 0.7
0.4	26.4 +/- 0.7	25.8 +/- 0.7	23.9 +/- 0.8	20.6 +/- 0.8	16.1 +/- 0.8	11.7 +/- 0.8
0.6	28.7 +/- 0.8	27.9 +/- 0.8	25.3 +/- 0.8	21.1 +/- 0.7	15.4 +/- 0.8	10.6 +/- 0.9
0.8	29.7 +/- 0.8	28.6 +/- 0.8	25.9 +/- 0.8	20.8 +/- 0.8	14.3 +/- 0.8	8.0 +/- 0.9
1	30.0 +/- 0.8	28.8 +/- 0.8	25.8 +/- 0.8	20.2 +/- 0.8	12.0 +/- 0.8	5.6 +/- 0.9

Figure 1. Standard Externality Model: Avg % increase in social welfare (with 90% confidence interval) from gain maximizing decision.

We first consider the standard externality model, with results presented in Figure 1. We use $\alpha = 0$ (no intervention) as the baseline which corresponds to buyers picking seller sets with the highest possible gain. We measure the increase in social welfare from this baseline to the social welfare at equilibrium under varying α and bias parameters (denoted by epsilon). As expected, our plots show that increasing α leads to increased social welfare. Also as expected, as the bias of the platform's externality estimate increases, welfare decreases, with the effect more pronounced as α increases. This is also reasonable since enforcing large transaction costs with very inaccurate predictions is inadvisable. Fortunately, the results show that $\alpha = 0.6$ can capture most of the increase in welfare, without suffering too much when then the externality estimates by the platform are inaccurate.

For the joint externality setting, whose results are presented in Figure 2, there is no dominant strategy, even without any intervention. As such, we use the gain maximizing choice from before as the baseline (which isn't necessarily an equilibrium). Our theoretical results show with intervention the resulting game is an approximate potential game with an approximate equilibrium. As such, we use a best response algorithm to find the approximate equilibrium. Once again, we plot the increase in social welfare (sum of all buyer utilities) from the baseline for different values of α and bias. As expected from our theoretical results, setting α closer to 0.5 from either direction leads to a higher increase in social welfare. The results are quite symmetric. We also notice that the effect of bias is less clear here than in the standard case.

Operationally, these results corroborate our assertion that α close to 0.5 is the ideal parameter. In the standard setting, it

Bias

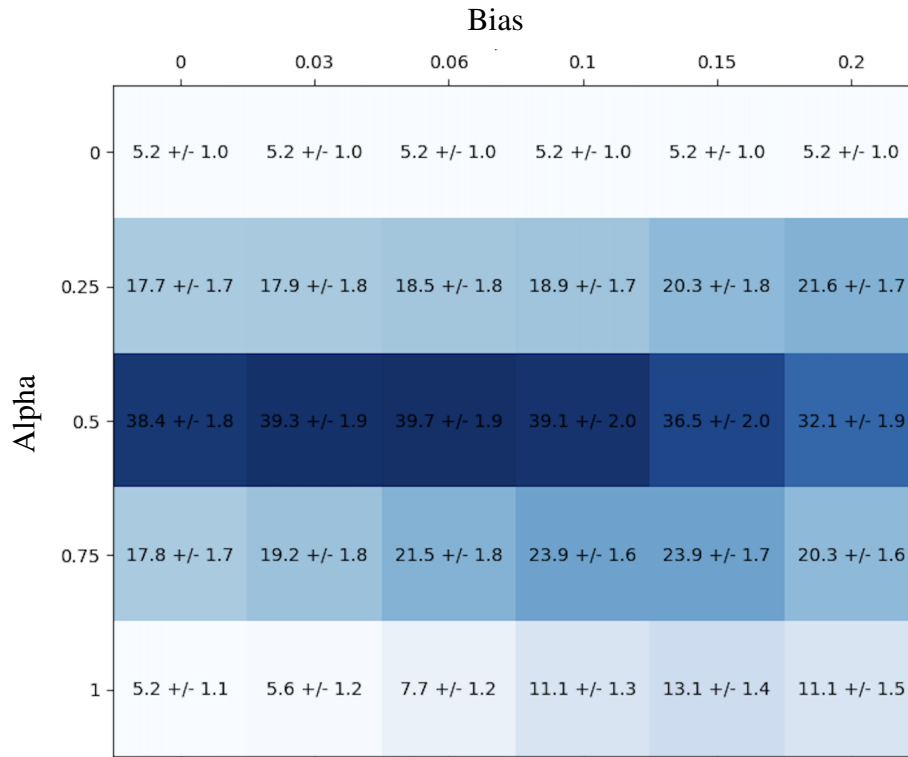| Alpha \ Bias	0	0.03	0.06	0.1	0.15	0.2
0	5.2 +/- 1.0	5.2 +/- 1.0	5.2 +/- 1.0	5.2 +/- 1.0	5.2 +/- 1.0	5.2 +/- 1.0
0.25	17.7 +/- 1.7	17.9 +/- 1.8	18.5 +/- 1.8	18.9 +/- 1.7	20.3 +/- 1.8	21.6 +/- 1.7
0.5	38.4 +/- 1.8	39.3 +/- 1.9	39.7 +/- 1.9	39.1 +/- 2.0	36.5 +/- 2.0	32.1 +/- 1.9
0.75	17.8 +/- 1.7	19.2 +/- 1.8	21.5 +/- 1.8	23.9 +/- 1.6	23.9 +/- 1.7	20.3 +/- 1.6
1	5.2 +/- 1.1	5.6 +/- 1.2	7.7 +/- 1.2	11.1 +/- 1.3	13.1 +/- 1.4	11.1 +/- 1.5

Figure 2. Joint Externality Model: Avg % increase in social welfare (with 90% confidence interval) from gain maximizing decision.

offers much of the welfare benefits of setting it higher without any of the drawbacks due to bias, and in the joint setting, it is outright the best parameter.