

Personalized Pricing via Strategic Learning of Buyers' Social Interactions

Qinqi Lin, Lingjie Duan, and Jianwei Huang

Abstract—As the sociological theory of homophily suggests, people tend to interact with those of similar preferences. This motivates product sellers to learn buyers' product preferences from the buyers' friends' purchase records. Although such learning allows sellers to enable personalized pricing to improve profits, buyers are also increasingly aware of such practices and may alter their behaviors accordingly. This paper presents the first study regarding how buyers may strategically manipulate their social interaction signals considering their preference correlations, and how an informed seller can take buyers' strategic social behaviors into consideration when designing the pricing schemes. Our analytical results show that only high-preference buyers tend to manipulate their social interactions to hurdle the seller's personalized pricing. Surprisingly, these high-preference buyers' payoff may become worse after their strategic manipulation. Furthermore, we show that the seller can greatly benefit from the learning practice, no matter whether the buyers are aware of such learning or not. In fact, buyers' learning-aware strategic manipulation only slightly reduces the seller's revenue. Considering the increasingly stricter policies on data access by authorities, it is thus advisable for sellers to make buyers aware of their access and learning based on social interaction data. This justifies well with current regulatory policies and industry practices regarding informed consent for data sharing.

Index Terms—online social networks, buyers' preference correlation, personalized pricing, dynamic Bayesian game

I. INTRODUCTION

With the ever-increasing penetration of online social media (e.g., Facebook and Twitter), people today can freely interact with one another online, exchanging product views or sharing purchase information [1]. For example, there were surprisingly 1.2 billion engagements on Facebook during the four-day Electronic Entertainment Experience 2021, where people shared *comments* or *posts* towards the newly released gaming products [2]. Such social data from people's online interactions reveal valuable information about their preferences and even the correlation behind the preferences.

“Birds of a feather flock together,” i.e., people tend to socially interact with those who have similar preferences [3]. This sociological theory of *homophily* allows product sellers

Q. Lin and J. Huang (corresponding author) are with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, and the Shenzhen Institute of Artificial Intelligence and Robotics for Society (Email: {qinqilin@link.cuhk.edu.cn, jianwei.huang@cuhk.edu.cn}). L. Duan is with Engineering Systems and Design Pillar, Singapore University of Technology and Design, Singapore (Email: lingjie_duan@sutd.edu.sg).

This work is supported by the Shenzhen Science and Technology Program (Project JCYJ20210324120011032), Guangdong Basic and Applied Basic Research Foundation (Project 2021B1515120008), and the Shenzhen Institute of Artificial Intelligence and Robotics for Society. The work is also supported in part by the Ministry of Education, Singapore, under its Academic Research Fund Tier 2 Grant (Award No. MOE-T2EP20121-0001).

to well relate buyers' mutual interactions in online social networks to their *preference correlation*. Researchers have been exploring methodologies to infer customers' similarities or differences from their interaction data to inform sellers' decision-making [4]. For example, Amazon has exploited the preferences of buyers' friends whom the buyers interact with on Facebook to provide purchase recommendations with higher accuracy [5]. Then Amazon can learn one buyer's private preference from correlated purchase records of his friends in the past. This paves the way for sellers' personalized pricing, enabling them to tailor the price to the target buyer's product preference in future sales. Apparently, such correlation information extracted from buyers' social interaction data can help the sellers improve their revenue.

However, sellers' increasing exploitation of social interaction data has raised buyers' severe concerns about publicizing their social data *without* their agreement or even awareness. For example, Amazon and Netflix were reported to have *undisclosed* data-sharing deals with Facebook, which permits the former parties privileged access to users' social interaction information [6]. Given such concerns, regulators have recently required social media platforms to inform people of their data-sharing with potential product sellers (e.g., [7], [8]). Following this, Amazon now needs to warn buyers of its access to their social interaction data when buyers log in and connect to their Facebook accounts [9].

Once *aware* of the seller's social data access and the possibility of being charged personalized prices, buyers may have the incentive to strategically manipulate their product-related social activities to thwart or mislead the product seller's follow-up learning and pricing practices. Indeed, people are known to provide untruthful data or misrepresent themselves on today's social media platforms (e.g., Facebook and Twitter) to hide their actual information [10]. For instance, buyers with high preferences for a new product may purposely stop their social discussions about it or the related products before the selling season, aiming to confuse the seller on their preference correlation. We are thus well motivated to ask the following key questions:

- **Key Question 1:** *Foreseeing the seller's learning, how should buyers manipulate their social interaction data?*
- **Key Question 2:** *How should the seller strategically learn from buyers' manipulated social data to redesign the optimal pricing scheme?*
- **Key Question 3:** *Are buyers better off through their social manipulations? Does such manipulation significantly reduce the seller's revenue?*

Note that if such manipulation greatly reduces the revenue, the seller may take a chance not to follow current regulators' requirements to keep buyers aware of the social data access. Previous works have studied how the seller should infer a single buyer's preference information from his own data (e.g., purchase records [11] and information from data broker [12]). As such, these works ignored the *correlation* among buyers' private information in today's fast-growing social networks. In contrast, we explore the opportunity that the seller learns the correlation among buyers' preferences from their social data.

Buyers use their social interactions as a means of *jointly* signaling their preference correlation with each other to the seller. In signaling parlance, buyers are signal senders, and the seller is the receiver. Our key difference from the traditional signaling paradigm [13] is that what senders signal is the correlation between their private information instead of an individual's private information alone (e.g., [11], [14]). Another difference lies in the coupling between the senders in the inter-dependent social networks, where one buyer's attempt to manipulate his social data and confuse the seller also affects his interacting friends' social decisions.

We formulate these coupled interactions among the seller and buyers under information asymmetry as a dynamic Bayesian game. In addition, we introduce two benchmarks: *no-learning benchmark* where the seller cannot access buyers' social data, and *undisclosed-learning benchmark* where the seller can access such data without buyers' awareness. By comparison, we examine the effect of the seller's learning practice and buyers' manipulation to provide useful insights.

We summarize the main contributions of this work below.

- *Novel personalized pricing via strategic learning from social data:* To the best of our knowledge, this is the first analytical study to tell how buyers proactively manipulate their social interaction signals and how an informed seller takes buyers' strategic social behaviors into consideration when designing the pricing schemes. Our work provides important insights into the ever-increasing practice of user profiling with social network data and the regulatory policy of informed consent for sharing buyers' data online.
- *Non-standard perfect Bayesian equilibrium (PBE) analysis:* It is difficult to analyze our dynamic Bayesian game, because we need to ensure sequential information consistency from coupled buyers to the seller. To resolve such challenge, we propose to first screen down the equilibrium space to facilitate the forward analysis. Then, we alternate it with backward induction to characterize the structure of buyers' social manipulations and analyze the seller's personalized pricing via strategic learning. Interestingly, we show that only high-preference buyers tend to manipulate their social interactions to hurdle the seller's personalized pricing.
- *Impacts of buyers' manipulations:* Interestingly, the seller does not always trust the buyers' manipulated social data to enable pure personalized pricing. Instead, the seller may randomize personalized pricing with uniform pricing, which in turn mitigates buyers' incentives to

manipulate. Surprisingly, we show that buyers could be even worse off by strategically manipulating their social interactions, compared to the benchmark case when they are unaware of the seller's learning practice.

- *Guidance on learning/pricing practices for the seller:* We show that the seller can greatly benefit from the learning practice, no matter whether the buyers are aware of such learning or not. In fact, buyers' learning-aware strategic manipulation only slightly reduces the seller's revenue. Considering the increasingly stricter privacy policies, it is thus advisable for sellers to make buyers aware of their access and learning over social interaction data. This justifies well with current regulatory policies and industry practices regarding informed consent for data sharing.

The rest of the paper is organized as follows. Section II reviews the related work. We present our system model in Section III and introduce two benchmarks in Section IV. In Section V, we first analyze the seller's strategic learning and then complete the PBE analysis in Section VI. Section VII examines the impact of buyers' manipulations. Section VIII concludes this paper. **Due to the page limit, we provide lengthy proofs of the key results in the online appendix [15].**

II. RELATED WORK

A. Information Sharing in Social Networks

Recent theoretic studies have shed light on information sharing in social networks (e.g., [16]–[18]). For instance, Ding *et al.* in [16] analyzed the multi-party privacy conflict (MPC) in online social networks, where the private information of one user is disclosed by others who co-own the data. Gradwohl in [18] considered network effects in social interactions as opposed to informational interdependence. These prior works mainly focused on the social interactions among users without considering any seller's engagement. Our work explicitly considers the seller's learning from users' social interaction data to infer private information and determine personalized pricing, which in turn affects the users' incentives and decision-making in social networks. On the users' side, the prior works have primarily focused on information leakage that one user incurs from the others' social activities due to information correlation or co-ownership. In contrast, we further consider the practice that users strategically compromise their social interactions to mislead the seller on their *correlation* of private information.

B. Price Discrimination with Buyer Recognition

On the seller's side, there is a growing literature on personalized pricing with buyer recognition, where the seller learns buyers' preferences from their purchase behaviors (e.g., [11], [12], [19], [20]). For example, Acquisti *et al.* in [20] and Conitzer *et al.* in [11] investigated the scenario where the seller conditions pricing on a single buyer's purchase records in repeated purchases. Particularly, authors in [11] allowed a single buyer to hide his own past purchase records to hinder the personalized pricing. Bellflamme in [19] further allowed a buyer to use technology to hide his valuation with a cost, when the seller attempts to track his private information with

the profiling technology. However, the above works did not consider the social connection among buyers nor their product preferences. By contrast, we study the seller’s personalized pricing by learning from social data and purchase records of a buyer’s friends rather than his own.

III. SYSTEM MODEL

We begin with the stylized yet fundamental model between a new product seller and any two socially connected buyers. Such a model captures buyers’ pairwise relationship for learning their preference correlation. We extend to more than two buyers’ mutual relationships in the online appendix [15] with a similar but more involved analysis. Here, the seller (e.g., Amazon) has access to the product-related discussions of two buyers, i and j , in an online social network (e.g., Facebook), aiming to learn their preference correlation and enable personalized pricing for these buyers in the coming selling season.

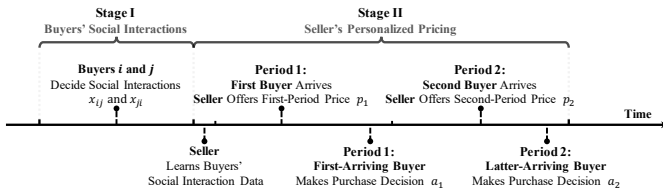


Fig. 1. Timeline of the dynamic Bayesian game

As shown in Fig. 1, we consider two stages: *Stage I (Buyers' Social Interactions)* is where the two buyers interact with each other in the online social network, and *Stage II (Seller's Personalized Pricing)* has two purchase periods where the interaction between the buyers and the seller in the product market happens. Next, we model these two stages in Section III-A and Section III-B, respectively. Finally, in Section III-C, we formally formulate our dynamic Bayesian game.

A. Model of Buyers' Social Interactions in Stage I

In Stage I, the two buyers socially interact to discuss a new product or related products on social media. As illustrated in Fig. 1, each buyer decides his *social interaction frequency* to the other buyer, jointly considering his *social interaction utility* in Stage I and his *purchase surplus* in Stage II later. Buyers' social decision-making relates to their purchase preferences, as elaborated in the following.

1) *Buyer Preference*: A buyer i (j , respectively) has a *purchase preference* v_i (v_j , respectively) for the new product, either high v_H or low v_L . This indicates his maximum willingness to pay for the product, with $v_L < v_H$. Without accessing the two buyers' social data, the seller believes that each buyer's preference is independently and identically distributed, being high (v_H) or low (v_L) with an equal prior probability $1/2$. Here, we focus on this binary preference model for ease of exposition, and we extend it similarly to a more general case of *continuous preference distributions* in the online appendix [15] to demonstrate the robustness of our key insights.

2) *Social Interaction Frequency*: The two buyers i and j simultaneously decide their *social interaction frequency* $x_{ij} \in \{0, 1\}$ and $x_{ji} \in \{0, 1\}$ with each other. Here, x_{ij} indicates how often buyer i interacts with buyer j (e.g., through messages, comments, and sharing). Specifically, $x_{ij} = 1$ implies buyer i interacts with buyer j frequently to discuss the new product on social media, whereas $x_{ij} = 0$ means buyer i seldom reaches out to buyer j .

3) *Social Interaction Utility*: As the sociological theory of homophily suggests, people tend to interact with those of similar preferences [3]. Thus, we model the *social interaction utility* $u_i(x_{ij}, x_{ji})$ in Table I to tell how much each buyer i gains (or suffers) when interacting with the other. Table I includes two sub-tables Tables I(a) and Table I(b), depending on whether their preferences of the new product are the same or different. Within each of the pairs in the same round bracket, the first number represents the buyer i 's social interaction utility and the second represents the buyer j 's, depending on their decisions (x_{ij} and x_{ji}) in social interaction frequency.

TABLE I
SOCIAL INTERACTION UTILITY OF BUYERS i AND j
(a) Same Preference $v_i = v_j$

	$x_{ji} = 1$	$x_{ji} = 0$
$x_{ij} = 1$	(1, 1)	(1 - l , l)
$x_{ij} = 0$	(l , 1 - l)	(0, 0)

(b) Different Preferences $v_i \neq v_j$

	$x_{ji} = 1$	$x_{ji} = 0$
$x_{ij} = 1$	(- c , - c)	(- c + r , - r)
$x_{ij} = 0$	(- r , - c + r)	(0, 0)

Table I(a) tells that two buyers have the same preference $v_i = v_j$. First we consider the combination case of $x_{ij} = 1$ and $x_{ji} = 1$, leading to the social interaction utilities (1, 1) for both buyers. Note that we normalize such social interaction happiness as a unit 1. In this case, when frequently interacting with the other buyer j of the same preference, a buyer i experiences social happiness through gaining empathy and reinforcing connections [21]. Then we move on to the case of $x_{ij} = 1$ and $x_{ji} = 0$, where the social interaction utilities are (1 - l , l). Here, we let $l \in (0, 1)$ denote the loss incurred from the normalized social happiness by the other buyer's low social response under the same preference. In this case, if buyer j seldom talks back to buyer i 's frequent social interaction $x_{ij} = 1$, buyer i would feel a lack of responses and gain less happiness $1 - l$ [22]. Alternatively, the other buyer j , when facing buyer i 's frequent social response $x_{ij} = 1$, also gains less happiness l as he remains silent with $x_{ji} = 0$. Finally for the case of $x_{ij} = 0$ and $x_{ji} = 0$, the social interaction utilities are (0, 0). Here, we normalize each buyer's social interaction utility from rare interactions as zero, no matter whether they share the same preference in Table I(a) or not in Table I(b).

Table I(b) tells that two buyers have different preferences $v_i \neq v_j$. First we consider the case of $x_{ij} = 1$ and $x_{ji} = 1$, the social interaction utilities are (- c , - c). When attempting to interact with the other buyer j who differs in preference, a buyer i experiences embarrassment with even a disutility c to maintain such interaction [3]. Then we move on to the

case of $x_{ij} = 1$ and $x_{ji} = 0$, where the social interaction utilities are $(-c+r, -r)$. In this case, the other buyer j 's low social responses instead relax buyer i 's awkwardness c given different preferences to a certain extent $r \in (0, c)$. Yet in turn, when receiving frequent messages with different opinions from buyer i , buyer j gets upset with a disutility r .

B. Model of Personalized Pricing in Stage II

At the beginning of Stage II, the seller can observe the *common interaction frequency* between buyer i and j , denoted as,

$$\hat{x} \triangleq \min\{x_{ij}, x_{ji}\}, \quad (1)$$

which is also binary in set $\{0, 1\}$. Here, the *minimum* operation captures the mutual essence of social interactions. This can be a reasonable approximation of the reality, as when processing interaction data in a large social network, the seller needs to locate the two buyers to identify the related data and would lose certain data if any buyer chooses not to interact with the other buyer. Consider an example involving Alice i and Bob j . Alice posts a co-owned photo with Bob ($x_{ij} = 1$), and the seller can identify their correlation after Bob comments on the post ($x_{ji} = 1$). Some privacy protection mechanism even asks Bob to grant or deny such a post [23]. That is, Bob's choice of denial $x_{ji} = 0$ disables the availability of the co-owned photo from Alice, i.e., $\min\{x_{ij}, x_{ji}\} = 0$. Overall, the seller needs to learn from such observable mutual data \hat{x} in (1) to infer the preference correlation between the two buyers.

Next, the seller announces prices p_1 and p_2 sequentially, for the two sequentially arriving buyers for the new product, respectively (with one buyer in each period of Stage II, see Fig. 1). Notice that the arrival sequence is *random* (i.e., buyer i may arrive after or before buyer j) due to information diffusion in marketing or randomness in individual behaviors. This sequential purchase pattern is commonly observed in practice and widely adopted in the literature on dynamic pricing (e.g., [11], [20]). We denote the first-arriving buyer's binary purchase decision as a_1 : $a_1 = 1$ if he decides to purchase (i.e., buyer's preference is no less than the offered price $v_1 \geq p_1$), and $a_1 = 0$ otherwise. Similarly, we define a_2 for the latter-arriving buyer's purchase decision.

Given incomplete information of buyers' preferences, the seller prices in the consecutive selling periods in Stage II to maximize the *expected sale revenue* $\tilde{\Pi}$, by taking the expectation over all possible arriving buyers in the market, i.e.,

$$\tilde{\Pi}(p_1, p_2) \triangleq \mathbb{E}_{v_1, v_2} \{p_1 a_1 + p_2 a_2\}. \quad (2)$$

On the other hand, the *final payoff* of each buyer consists of the purchase surplus in Stage II and the social interaction utility in Stage I (see Table I). To illustrate, if buyer i arrives in the selling period $t \in \{1, 2\}$ of Stage II (hence with $v_t = v_i$), then the final payoff π_i of buyer i is:

$$\pi_i(x_{ij}, x_{ji}; v_t = v_i) \triangleq \max\{v_t - p_t, 0\} + u_i(x_{ij}, x_{ji}). \quad (3)$$

Yet when deciding the social interaction frequency in Stage I, buyer i is not sure whether he arrives earlier than buyer j or not in the following Stage II. Thus, anticipating the seller's

pricing in Stage II, buyer i makes the social decision x_{ij} to maximize his *expected total payoff* $\tilde{\pi}_i$ over all possible arrival sequences, i.e.,

$$\tilde{\pi}_i(x_{ij}, x_{ji}) \triangleq \mathbb{E}_{t \in \{1, 2\}} [\pi_i(x_{ij}, x_{ji}; v_t = v_i)]. \quad (4)$$

Here, buyer i may deviate from the interaction frequency determined purely according to Table I. We broadly define such deviation as buyer i 's *manipulation* of his social interaction with buyer j , which is expected to disguise their (true) homophily information.

C. Dynamic Bayesian Game Formulation

We formally model the interactions among the seller and the buyers as a dynamic Bayesian game as follows, with the decision-making timing illustrated in Fig. 1. The solution concept we adopt is the *perfect Bayesian equilibrium (PBE)*. To differentiate from the two benchmark models to be introduced later, we will terminate the current one *strategic-learning model*.

- **Stage I:** Two buyers i and j simultaneously decide their social interaction frequencies x_{ij} and x_{ji} , with the goals of maximizing their individual expected total payoffs $\tilde{\pi}_i$ or $\tilde{\pi}_j$ in (4), respectively.
- **Stage II:** After accessing buyers' social interaction data, the seller decides and announces prices p_1 and p_2 sequentially for the arriving buyers to maximize her expected sale revenue in (2) with strategic learning. Note that here buyers in any selling period t decide to purchase or not:

$$a_t^*(v_t, p_t) = \mathbf{1}(v_t \geq p_t), \quad \forall t \in \{1, 2\}, \quad (5)$$

where $\mathbf{1}(\cdot)$ is an indicator function.

Next, we explain the *information structure* of this game. At the beginning of the game, both buyers' product preferences are private and only known to themselves;¹ the seller only knows the prior distribution. After learning from buyers' mutual interactions in the social network in Stage I, the seller infers the correlation between buyers' preferences at the beginning of Stage II. Yet, she still does not know each individual's private preference. Only after offering the first-period price p_1 in Stage II, the seller may use the price to sample the first-arriving buyer's preference from his purchase decision. Together with the correlation learned from Stage I, the seller can infer the latter-arriving buyer's preference. Overall, the seller's incomplete information about each buyer's preference gradually decreases along the timeline in Fig.1.

Notice that the seller cannot charge a personalized price to the first-arriving buyer. This is because the seller only knows the *correlation* of both buyers' preferences instead of this buyer's *individual preference*. But for the latter-arriving buyer, the seller can combine the purchase record in the prior selling period to tell this buyer's preference. In this sense, only the second-period price p_2 can be personalized.

¹Before deciding the social interaction frequencies on social media, the two buyers have already had acquaintance with each other from historical or physical interactions. Thus, they know each other's product preferences.

IV. TWO BENCHMARKS

Before we analyze the PBE of the strategic-learning model defined in Section III-C, we first introduce two benchmarks for later comparisons. Due to the page limit, we focus on discussing key insights here and leave the equilibrium details in the online appendix [15].

1) *No-learning Benchmark*: First, we consider the case where the seller is *not* allowed to access buyers' social interaction data. Here, no buyer has the incentive to manipulate, and the seller sets uniform prices in both selling periods.

2) *Undisclosed-learning Benchmark*: Next, we introduce the second benchmark case, where the seller can ideally access buyers' social interaction data in Stage I in an *undisclosed* way. In other words, buyers are *not* aware of such exploitation from the seller. Therefore, buyers' social behaviors (x_{ij}^*, x_{ji}^*) in Stage I will be the same as that under the no-learning benchmark. However, the seller's pricing decisions in Stage II can be personalized then.

As buyers do not manipulate their social data in Stage I, the seller learns the product preferences of the two buyers are *positively correlated* when observing $\hat{x}^* \triangleq \min\{x_{ij}^*, x_{ji}^*\} = 1$. Further, if the first-arriving buyer purchases the product at $p_1^* = v_H$, the seller knows that his preference is $v_1 = v_H$ to afford a high price. Thus, the seller would also charge $p_2^* = v_H$ to the latter-arriving buyer, who shall have the same high preference $v_2 = v_H$. Yet if the first-arriving buyer does not purchase, the seller learns that his preference is $v_1 = v_L$ and the latter-arriving buyer should be with the same low preference $v_2 = v_L$. Hence, the seller charges $p_2^* = v_L$. Similar arguments work when the seller observes $\hat{x}^* = 0$. Here, the seller achieves an expected revenue no smaller than that under the no-learning benchmark.

V. FORWARD ANALYSIS FOR SELLER'S LEARNING

We now move on to our focused strategic-learning model defined in Section III-C, where the seller tries to *learn* buyers' purchase preferences from their social interaction data, and the buyers are *aware* of such learning. In this section, we start with an equilibrium analysis of the seller's strategic learning from Stage I (Section V-A) and the rational pricing in Stage II (Section V-B), which allows us to *screen down* the original equilibrium space (i.e., reducing the searching space of all possible strategy combinations). By doing so, we can formulate a belief of buyers' manipulation structure in Stage I with *forward analysis*, based on which we carry on tractable analysis of PBE using *backward induction* in the next section.

Given *joint information asymmetry* in individual preference and preference correlation, there are a number of possibilities of buyers' private information behind the seller's observable social data \hat{x} in (1). This motivates us to first reduce the searching space of all possible strategy combinations to facilitate our forward analysis in this section for the seller's learning.

A. Learning from Buyers' Social Interaction Data in Stage I

We first guide the seller's strategic learning from buyers' social data in Stage I with the following Lemma 1.

Lemma 1. *At the PBE of the strategic-learning model, buyers i and j with the same low preference choose the maximum social interaction frequency in Stage I:*

$$x_{ij}^*(v_i = v_j = v_L) = x_{ji}^*(v_j = v_i = v_L) = 1. \quad (6)$$

Given different preferences, both buyers choose the minimum social interaction frequency in Stage I:

$$x_{ij}^*(v_j \neq v_i) = x_{ji}^*(v_i \neq v_j) = 0. \quad (7)$$

Intuitively, a low-preference buyer just affords a personalized price $p_2 = v_L$ in Stage II. Likewise, when facing higher prices, he also receives a zero purchase surplus by deciding not to purchase the product. When meeting another low-preference buyer, manipulation of social data from high to low interaction frequency does not improve this low-preference buyer's purchase surplus, but reduces his social interaction utility 1 in Table I(a). However, when meeting a high-preference buyer j , a low-preference buyer i loses from frequent interaction given different preferences as in Table I(b). In this case, this low-preference buyer i honestly chooses $x_{ij}^* = 0$, as any manipulation to $x_{ij}^* = 1$ does not help improve his purchase surplus in Stage II. As a result, the common interaction frequency signal in (1) to the seller is always $\hat{x} = 0$. The high-preference buyer j cannot manipulate their correlation unilaterally, and thus also chooses $x_{ji}^* = 0$ honestly given their different preferences.

The remaining case to consider is when $v_i = v_j = v_H$. In this case, we generally allow these buyers to choose a mixed strategy, i.e., choosing the low interaction frequency 0 with a manipulation probability and choosing the high frequency 1 with the complementary probability. Formally, we define such *manipulation probability* for high-preference buyers below. Note that as both high-preference buyers are symmetric, they would choose the same manipulation probability.

Definition 1. *When interacting with the other high-preference buyer, we define any high-preference buyer's manipulation probability for the event that he chooses low social interaction frequency 0 as ρ . i.e.,*

$$\begin{aligned} \rho &\triangleq \Pr(x_{ij}(v_i = v_j = v_H) = 0) \\ &\equiv \Pr(x_{ji}(v_j = v_i = v_H) = 0). \end{aligned} \quad (8)$$

To enable the seller's personalized pricing in Stage II, we now analyze the seller's posterior belief of buyers' product preferences by learning from their common interaction frequency \hat{x} . According to Bayes' theorem, we have for the seller:

$$\Pr(v_i = v_j = v_H | \hat{x} = 1) = \frac{\Pr(\hat{x} = 1 | v_i = v_j = v_H)}{\Pr(\hat{x} = 1 | v_i = v_j = v_H) + 1}, \quad (9)$$

and

$$\Pr(v_i = v_j = v_H | \hat{x} = 0) = \frac{1 - \Pr(\hat{x} = 1 | v_i = v_j = v_H)}{3 - \Pr(\hat{x} = 1 | v_i = v_j = v_H)}, \quad (10)$$

which cannot be directly determined using *backward induction* from Stage II. Indeed, to ensure belief consistency over time, we need to examine the consistency between buyers' social decisions and the beliefs above (see Section VI).

B. Learning from First-arriving Buyer's Purchase in Stage II

Next, we present another lemma that helps us screen down the seller's pricing-decision space in Stage II.

Lemma 2. *In Stage II of the strategic-learning model, the seller's equilibrium price p_t^* in any selling period $t \in \{1, 2\}$ is either v_H or v_L .*

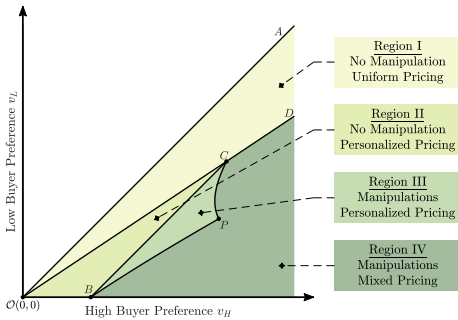
Lemma 2 helps the seller restrict to the binary pricing choices in each selling period of Stage II. Strategically, the seller may consider learning the first-arriving buyer's preference from his purchase decision by offering a high price $p_1 = v_H$. In this way, the seller can facilitate personalized pricing to the latter-arriving buyer with the preference correlation learned from Stage I. On the other hand, when offering a low price $p_1 = v_L$, the seller cannot infer the first-arriving buyer's preference, though the seller can guarantee a revenue of v_L from each arriving buyer. In this sense, the seller faces the tradeoff of whether to enable personalized pricing in Stage II. In addition, the seller's pricing also needs to account for buyers' manipulations in Stage I.

VI. ANALYSIS OF PBE FOR DYNAMIC BAYESIAN GAME

After screening down the buyers' social behaviors using Lemma 1 in Stage I and the seller's strategic pricing using Lemma 2 in Stage II, we now further analyze the PBE of the strategic-learning model in this section. The analysis is involved as we need to ensure the belief consistency from the coupled buyers to the seller over stages. For this goal, we alternate forward analysis (see Section V) with *backward induction*, while ensuring consistency over stages as follows.

- *Step 1:* We first use backward induction to analyze the seller's pricing decisions in Stage II, given the posterior belief on buyers' social decisions in (9) and (10).
- *Step 2:* Based on the seller's optimal pricing derived in Step 1, we next analyze the buyers' *coupled* social behaviors in Stage I, which should be consistent with the seller's belief to ensure *belief consistency* across stages.

We fully characterize all PBE in closed-form and classify the structural results with four preference regions in Fig.2.



OA: $v_L = v_H$	OCD: $v_L = 2v_H/3$	BC: $v_H - v_L = 2(1-l)$
CP: $(v_H - v_L)v_L = 8(1-l)^2$	BP: $(v_H - v_L)(v_H - 2v_L) = 4(1-l)^2$	

Fig. 2. PBE in four regions to tell whether high-preference buyers manipulate their interaction with the other high-preference one, and how the seller prices.

Proposition 1. *In the strategic-learning model, if low and high preferences are close in Region I of Fig.2, there exists a unique*

PBE as follows. In Stage I, buyers i and j with the same high preference choose the maximum social interaction frequency:

$$x_{ij}^*(v_i = v_j = v_H) = x_{ji}^*(v_j = v_i = v_H) = 1. \quad (11)$$

In Stage II, the seller chooses low prices $p_1^ = p_2^* = v_L$ in both selling periods.*

Note that (11) for $v_i = v_j = v_H$ together with Lemma 1 (for the other preference distributions) completes all the PBE result in Stage I. Proposition 1 shows that when low product preference v_L is close to high preference v_H , it is optimal for the seller *not* to practice the learning-enabled personalized pricing. Instead, the seller charges the low prices v_L in both selling periods without losing any buyers' demands. As the seller will not personalize the prices, buyers in Stage I behave honestly without any manipulation of their social data.

Proposition 2. *In the strategic-learning model, if low and high preferences are different with a small difference in Region II of Fig.2, there exists a unique mixed PBE as follows.*

In Stage I, buyers i and j with the same high preference choose the maximum social interaction frequency in (11). In Stage II, the seller charges a high price $p_1^ = v_H$ in the first selling period. The second-period price personalizes, depending on the observed buyers' common interaction frequency $\hat{x}^* \triangleq \min\{x_{ij}^*, x_{ji}^*\}$ in Stage I and the purchase record $a_1^* = \{0, 1\}$ of the first period in Stage II:*

$$p_2^* = \begin{cases} v_H \mathbf{1}(a_1^* = 1) + v_L \mathbf{1}(a_1^* = 0), & \text{if } \hat{x}^* = 1. \\ v_L \mathbf{1}(a_1^* = 1) + v_H \mathbf{1}(a_1^* = 0), & \text{if } \hat{x}^* = 0. \end{cases} \quad (12a) \quad (12b)$$

As the low preference v_L is no longer close to v_H , the seller has the motivation to enable personalized pricing in Stage II. In this case, she charges a high price $p_1^* = v_H$ in the first selling period to learn the buyer's preference, and further tailors the second-period price p_2 accordingly. Although facing a possible personalized price in Stage II, Proposition 2 shows that the two high-preference buyers still do not manipulate in Stage I. This is because buyers' social loss from manipulation outweighs the potential purchase gain from pretending as a low-preference type (i.e., $v_H - v_L < 2(1-l)$ in Region II of Fig.2).

Proposition 3. *In the strategic-learning model, if low and high preferences are different with a medium difference in Region III of Fig.2, there exists a unique mixed PBE as follows.*

In Stage I, buyers i and j with the same high preference do not always choose the high social interaction frequency. Instead, they randomize low/high frequencies with manipulation probability $\rho^ = 1 - 2(1-l)/(v_H - v_L)$. In Stage II, the seller charges a high price $p_1^* = v_H$ in the first selling period. The second-period price is the same as (12a) and (12b).*

With a larger potential purchase gain when viewed as a low-preference type (i.e., $v_H - v_L > 2(1-l)$ in Region III of Fig.2), high-preference buyers are motivated to manipulate with a positive probability of $\rho^* > 0$. They aim to avoid being charged high personalized prices but receive lower prices v_L instead. As v_H increases or v_L decreases, the purchase gain $v_H - v_L$ from manipulation becomes more significant. Hence,

the high-preference buyers would raise the manipulation probability ρ^* to achieve a greater purchase surplus.

Although the seller may fail to identify high-preference buyers given such manipulations, Proposition 2 suggests that the seller still takes the same personalized pricing in (12a) and (12b) as in Region II of Fig. 2. This is because the probability of these buyers' manipulations is still minor given the moderate potential purchase gain, and the seller can still capture buyers' surplus in most cases by enabling the personalized pricing.

Proposition 4. *In the strategic-learning model, if low and high preferences are different with a large difference in Region IV of Fig.2, there exists a unique mixed PBE as follows.*

In Stage I, buyers i and j with the same high preference do not always choose the high social interaction frequency. Instead, they randomize low/high frequencies with the following manipulation probability:

$$\rho^* = \begin{cases} 1 - \sqrt{v_L/2(v_H - v_L)}, & \text{if } v_L > \frac{2}{5}v_H, \\ 1 - \sqrt{(v_H - 2v_L)/(v_H - v_L)}, & \text{otherwise.} \end{cases} \quad (13)$$

In Stage II, the seller does not always charge a high price $p_1^* = v_H$ in the first selling period and enable second-period personalized pricing in (12a) and (12b). More specifically, we have the following two cases.

- **Case 1** with $v_L > 2v_H/5$: When observing buyers' high common interaction frequency $\hat{x}^* = 1$, the seller randomizes with offering low prices $p_1^* = p_2^* = v_L$ in both selling periods. i.e.,

$$\begin{cases} p_1^* = p_2^* = v_L, & \text{w.p. } \frac{v_H - v_L - 2\sqrt{2}(1-l)}{2(v_H - v_L)}, \\ p_1^* = v_H, (12a), & \text{otherwise.} \end{cases} \quad (14)$$

Whereas observing $\hat{x}^* = 0$, the seller purely charges a high price $p_1^* = v_H$ in the first period and enables personalized pricing in (12b) in the second period.

- **Case 2** with $v_L < 2v_H/5$: When observing buyers' low common interaction frequency $\hat{x}^* = 0$, the seller randomizes with offering high prices $p_1^* = p_2^* = v_H$ in both selling periods. i.e.,

$$\begin{cases} p_1^* = p_2^* = v_H, & \text{w.p. } 1 - \frac{2(1-l)}{\sqrt{(v_H - v_L)(v_H - 2v_L)}}, \\ p_1^* = v_H, (12b), & \text{otherwise.} \end{cases} \quad (15)$$

Whereas observing $\hat{x}^* = 1$, the seller purely charges a high price $p_1^* = v_H$ in the first period and enables personalized pricing in (12a) in the second period.

With a sufficiently large potential purchase gain $v_H - v_L$ in Region IV of Fig.2, high-preference buyers are now more willing to manipulate their social interactions. These high-preference buyers would raise the manipulation probability ρ^* as v_H increases or v_L decreases in Case 1, because the purchase gain $v_H - v_L$ becomes more significant under these changes. On the other hand, the seller expects such social manipulation, and introduces a mixture (mixed strategy) of uniform pricing v_L with personalized pricing in (14) when

observing the high common interaction frequency $\hat{x}^* = 1$. Given a higher v_L in Case 1, such a mixture helps the seller gain more from those low-preference buyers who never manipulate and frequently interact with each other. As a result, the seller mitigates the potential revenue loss from those manipulating high-preference buyers.

In contrast, when v_L becomes lower, as in Case 2, the seller tends to mix personalized pricing with another pricing scheme in (15), offering a uniform price v_H when observing the low common interaction frequency $\hat{x}^* = 0$. Specifically, the seller strategically infers those manipulating high-preference buyers upon $\hat{x}^* = 0$. With small v_L , the seller worries little about the demand loss from low-preference buyers, who never manipulate and seldom interact with high-preference buyers. As a result, the seller mitigates the loss by extracting revenue from those manipulating high-preference buyers. Meanwhile, such a mixture of uniform pricing $p_1^* = p_2^* = v_H$ upon $\hat{x}^* = 0$ also reduces high-preference buyers' incentives to manipulate. That is, when choosing the low interaction frequency of $\hat{x}^* = 0$, the high-preference buyers are less likely to avoid personalized pricing but still face social loss. Therefore, even though the potential purchase gain $v_H - v_L$ enlarges as v_H increases or v_L decreases, high-preference buyers' manipulation level ρ^* decreases in Case 2.

VII. IMPACTS OF BUYERS' MANIPULATIONS

After analyzing the PBE in Propositions 1-4, we are ready to understand how buyers' social data manipulations affect the payoffs of both the seller and buyers.

A. Impact on Buyers' Payoffs

First, we are interested in whether buyers receive higher payoffs from manipulating their social data to hurdle the seller's personalized pricing. Proposition 5 answers it by comparing to the *undisclosed-learning* benchmark, where buyers are *unaware* of the seller's learning and thus never manipulate.

Formally, we examine the *average buyer payoff* $\tilde{\pi}$ given by

$$\tilde{\pi} \triangleq \mathbb{E}_{v_i, v_j} \{ \tilde{\pi}_i(x_{ij}, x_{ji}) \}, \quad (16)$$

which takes expectation over various buyer preference distributions while accounting for the preference correlation between buyers. Indeed, the awareness of the seller's learning would not affect low-preference buyers' payoffs, as they never manipulate and always gain zero purchase surplus (see Lemmas 1 and 2). Hence, the investigation on (16) to compare with the undisclosed-learning benchmark actually sheds light on the high-preference buyers, as discussed in the following.

Proposition 5. *Compared with the undisclosed-learning benchmark, the average buyer payoff could be lower when buyers are aware of the seller's learning in the strategic-learning model.*

Proposition 5 reveals that *high-preference buyers strategically manipulate their social interactions but may end up with lower payoffs (than no manipulation)*. This is because each

buyer is able to confuse the seller through *unilateral manipulation*, as the seller only accesses the common mutual interactions $\hat{x} \triangleq \min\{x_{ij}, x_{ji}\}$. However, both high-preference buyers i and j may tend to manipulate unilaterally by themselves. Thus, both suffer the social loss due to the other's manipulation, imposing a *negative externality* on each other.

Due to the page limit, we present more analytical details in the online appendix [15] regarding how awareness affects the average buyer payoff across four different regions of Fig. 2.

B. Guidance on Seller's Learning from Online Social Data

Next, we discuss if it is beneficial for the seller to learn from buyers' social interaction data, and if she shall actively inform buyers of her data access. For example, Amazon today seeks explicit consent from buyers regarding Amazon's access to her buyers' Facebook data [9]. However, some other companies are found not to inform buyers regarding their data access [24]. We investigate this issue by comparing the seller's expected sale revenue obtained under the strategic-learning model with the *no-learning* and *undisclosed-learning* benchmarks.

Proposition 6. *Compared with the no-learning benchmark, the seller achieves a non-negative expected revenue gain in the strategic-learning model. The gain is positive as long as $0 < v_L < 2v_H/3$ and the maximum gain can reach 25%.*

With a big revenue gain up to 25% beyond the no-learning benchmark, the seller should always learn and exercise personalized price as long as buyers' preferences for the new product are diverse ($0 < v_L < 2v_H/3$). Notably, this makes sense no matter whether buyers are aware (or not) of such learning and the follow-up personalized pricing. Despite buyers' manipulation in social interactions to hurdle the seller's personalized pricing, the seller can strategically learn and introduce a mixture with uniform pricing (see Proposition 4) to mitigate the manipulation effect and avoid revenue loss.

Proposition 7. *Compared with the undisclosed-learning benchmark, the seller suffers a non-negative expected revenue loss in the strategic-learning model, with only 8.3% at most.*

Although buyers' awareness would reduce the seller's gain from learning, such revenue loss is insignificant, with only 8.3% in the worst case. Given the increasing trend of better personal data protection, it is advisable for the seller to inform buyers of her access to their social data. Indeed, this matches well with Amazon's current practice regarding *informed consent* for data sharing [9]. From the regulators' perspective, it is enough to impose a fine on the sellers that is equal to or larger than the potential gain of not informing the buyers (8.3% in our model) to incentivize proper behaviors.

VIII. CONCLUSION

This paper studies how buyers may strategically manipulate their social interaction signals considering their preference correlations, and how an informed seller can take buyers' strategic social behaviors into consideration when designing the pricing schemes. We find that only high-preference buyers tend to manipulate their social interactions to hurdle the

seller's personalized pricing. Yet, their payoff may become worse after such manipulation. We also show that the seller benefits from learning the buyers' social data independent of buyers' awareness. We thus advise sellers to make buyers aware of their social data access and follow-up learning.

REFERENCES

- [1] P. Jain, A. Hall-Phillips, and S. Djamasbi, "Effect of social media product reviews on buying decision when presented in augmented reality," in *International Conference on HCI in Business, Government, and Organizations*. Springer, 2018, pp. 313–326.
- [2] J. Batchelor, "What were the most talked about games of e3 week?" Jun 2021. [Online]. Available: <https://www.gamesindustry.biz/articles/2021-06-25-what-were-the-most-talked-about-games-of-e3-week>
- [3] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annual review of sociology*, 2001.
- [4] M. Braun and A. Bonfrer, "Scalable inference of customer similarities from interactions data using dirichlet processes," *Marketing Science*, vol. 30, no. 3, pp. 513–531, 2011.
- [5] L. Parfeni, "Amazon adds social recommendations based on facebook," 2010. [Online]. Available: <https://news.softpedia.com/news/Amazon-Adds-Social-Recommendations-Based-on-Facebook-149460.shtml>
- [6] G. J. X. Dance, M. LaForgia, and N. Confessore, "As facebook raised a privacy wall, it carved an opening for tech giants," 2018. [Online]. Available: <https://www.nytimes.com/2018/12/18/technology/facebook-privacy.html>
- [7] GDPR, "General data protection regulation," <https://gdpr-info.eu>, 2019.
- [8] CalOPPA, "Online privacy protection act," 2021. [Online]. Available: https://en.wikipedia.org/wiki/Online_Privacy_Protection_Act
- [9] C. Boulton, "Amazon, facebook connect for social shopping," Feb 2021. [Online]. Available: <https://www.eweek.com/it-management/amazon-facebook-connect-for-social-shopping/>
- [10] S. Sannon, N. N. Bazarova, and D. Cosley, "Privacy lies: Understanding how, when, and why people lie to protect their privacy in multiple online contexts," in *Proc. of CHI Conference on Human Factors in Computing Systems*, 2018.
- [11] V. Conitzer, C. R. Taylor, and L. Wagman, "Hide and seek: Costly consumer privacy in a market with repeat purchases," *Marketing Science*, vol. 31, no. 2, pp. 277–292, 2012.
- [12] R. Montes, W. Sand-Zantman, and T. Valletti, "The value of personal information in online markets with endogenous privacy," *Management Science*, vol. 65, no. 3, pp. 1342–1362, 2019.
- [13] M. Spence, "Job market signaling," in *Uncertainty in economics*. Elsevier, 1978, pp. 281–306.
- [14] K. Drakopoulos, S. Jain, and R. Randhawa, "Persuading customers to buy early: The value of personalized information provisioning," *Management Science*, vol. 67, no. 2, pp. 828–853, 2021.
- [15] "online appendix," 2021. [Online]. Available: <https://www.dropbox.com/sh/6xeh67cxdivr7vr/AAADsDRkefu77KTzDS1Am1neKa?dl=0>
- [16] K. Ding and J. Zhang, "Multi-party privacy conflict management in online social networks: a network game perspective," *IEEE/ACM Transactions on Networking*, vol. 28, no. 6, pp. 2685–2698, 2020.
- [17] S. Rajtmajer, A. Squicciarini, J. M. Such, J. Semonsen, and A. Belmonte, "An ultimatum game model for the evolution of privacy in jointly managed content," in *International Conference on Decision and Game Theory for Security*. Springer, 2017, pp. 112–130.
- [18] R. Gradwohl, "Information sharing and privacy in networks," in *Proc. ACM Conference on Economics and Computation*, 2017, pp. 349–350.
- [19] P. Belleflamme and W. Vergote, "Monopoly price discrimination and privacy: The hidden cost of hiding," *Economics Letters*, 2016.
- [20] A. Acquisti and H. R. Varian, "Conditioning prices on purchase history," *Marketing Science*, vol. 24, no. 3, pp. 367–381, 2005.
- [21] C. Seiter, "The secret psychology of facebook: Why we like, share, comment and keep coming back," 2015.
- [22] H. Krasnova, S. Spiekermann, K. Koroleva, and T. Hildebrand, "Online social networks: Why we disclose," *Journal of information technology*, vol. 25, no. 2, pp. 109–125, 2010.
- [23] J. M. Such and N. Criado, "Resolving multi-party privacy conflicts in social media," *IEEE Trans. on Knowl Data Eng.*, 2016.
- [24] I. A. Hamilton, "Zoom is being sued for allegedly handing over data to facebook," Mar 2020. [Online]. Available: <https://www.businessinsider.com/zoom-sued-allegedly-sharing-data-with-facebook-2020-3>