Estimating Pair-Specific Network Effects in Binary-Action Games: Two-Sided Markets via Restricted Boltzmann Machines*

Tetsuya Hoshino† Romans Pancs‡

September 25, 2026

Abstract

We extend the standard model of a two-sided market by allowing for heterogeneous, pair-specific effects from interactions. Our main contribution is to show how to estimate these effects consistently and tractably by observing agents' repeated participation decisions under stochastic best-response play. The proposed estimator exploits a formal equivalence between equilibria and restricted Boltzmann machines, a class of neural networks. We illustrate the value to a platform operator of engaging directly with heterogeneous interaction effects by showing that model misspecification can yield an arbitrarily small fraction of the revenue attainable under the correctly specified model. Finally, we show that the estimator's consistency—as well as our optimal pricing results—extends to general binary-action games with bilateral network externalities (e.g., R&D games and online recruiting marketplaces).

Keywords: two-sided markets, platforms, restricted Boltzmann machines, contrastive divergence

JEL Classification Numbers: C45, D85, L10

1 Introduction

Two-sided markets—often called platforms—connect two distinct groups of participants. On many such platforms, an agent cares about not just how many others join on the other side but also who exactly joins. For instance, a dater values logging onto a dating platform differently depending on which prospective partners he expects to be online at that time. Existing literature on two-sided markets, pioneered by Caillaud and Jullien (2003) and Rochet and Tirole (2003) and surveyed by Jullien, Pavan, and Rysman (2021), examines platform pricing, regulation, and estimation. Much of this literature makes the restrictive assumption that each agent cares only about the number

*We are grateful to the Editor, Alfonso Gambardella, the anonymous Associate Editor, and two anonymous referees for helpful comments. For feedback, we thank Nageeb Ali, Lester Chan, Ben Golub, Igal Hendel, Glen Weyl, Soomin Jung, Asen Kochov, Shaun McRae, Alessandro Pavan, Yuta Watabe, and Mauricio Romero, as well as the workshop and seminar audiences at Northwestern, Penn State, the University of Toronto, the 2023 SAET Conference, and the 2024 ESIF Economics and AI+ML Meeting. We also thank Natalia Denisenko for her excellent research assistance. The project was conceived during the second author's sabbatical in Northwestern University's Department of Economics and brought to fruition during his appointment as a visiting associate professor at the University of Rochester's Wallis Institute of Political Economy. This work is supported by JSPS KAKENHI Grant Number JP24K23942.

†Kyoto University. Email: tetsuyahoshino137@gmail.com

‡ITAM. Email: rpancs@gmail.com

of participants on the other side and not about their identities. This assumption is convenient for estimation because it ensures that a participation profile can always be summarized by two numbers: the number of participants on either side of the market. In this case, conventional maximum likelihood is numerically tractable. When the existing literature does let agents care about the identities of others, it tends to do so through agents' characteristics or in other restrictive ways, which reduce the problem's dimensionality and, once again, facilitate tractability. To this end, Hitsch, Hortaçsu, and Ariely (2010) in the context of dating and Lee (2013) in the context of video games make an agent's value of an interaction depend on a small number of estimated or observed characteristics of others rather than their identities. Gomes and Pavan (2016) restrict preferences so that all agents agree on who the desirable partners are. When estimating the consequences of platform merger, Farronato, Fong, and Fradkin (2024) model one side of a two-sided market in reduced form.

Our paper's substantive contribution is to show how to estimate the full set of pairwise interaction effects, without reducing the problem's dimensionality by making restrictive functional-form assumptions and without making pairwise effects depend on a small number of agents' characteristics. The estimator ingests a panel data set that contains agents' repeated participation decisions. The participation profiles in the data set are assumed to be i.i.d. draws from the stationary distribution generated by agents' noisy myopic best responses to others' past participation decisions. This distribution, which we call the stochastic best-response equilibrium (sBRE), happens to coincide with the restricted Boltzmann machine (RBM), a type of neural network. This coincidence lets us use contrastive divergence, a fast algorithm used to train RBMs, as the basis for our estimator, whose consistency we establish formally and whose numerical tractability we document in simulations. Thus, the paper's methodological contribution, the sBRE representation of agent behavior (Proposition 1), enables its substantive contribution, the estimator (Proposition 2). A different behavioral model would generally require a different estimator. Getting the interaction effects right is essential in practice for a revenue-maximizing platform. Pricing under the misspecified model that assumes that participants care only about the number of partners but not their identities can yield an arbitrarily small fraction of the revenue attainable under the correctly specified model that recognizes that participants care about the identities of others (Proposition 5).

We organize the paper around a story of a startup that repeatedly observes the same agents' decisions to log onto its platform. In order to price optimally, the startup must first infer each agent's standalone value from participation and his pairwise benefits (possibly negative) from interacting with agents on the other side of the market. We supply the startup with a consistent and numerically tractable estimator of these parameters. We then illustrate the utility of estimating these parameters in two ways. We exhibit the welfare- and the profit-maximizing prices. (The startup may choose to first maximize welfare, in order to grow its customer base, and only then to pursue revenue.) We also show that working with a misspecified model that assumes that agents only care about the number of counterparts may cost the startup dearly.

Formally, our model of the two-sided market extends the classic all-to-all matching model of

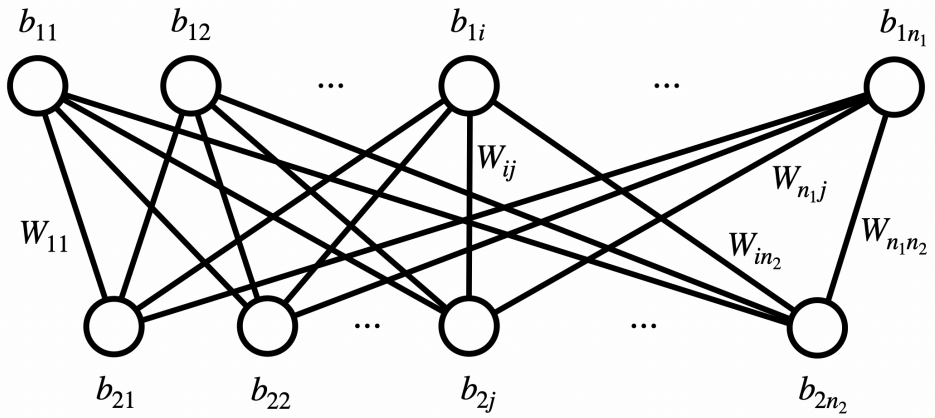

Figure 1: **Payoff parameters in a two-sided market.** The two rows of nodes are the two sides of the market. An edge connects every agent i on side 1 to every agent j on side 2, with the weight W_{ij} (which can be positive, negative, or zero) corresponding to the interaction benefit for each member of the participating pair. There are no interaction benefits for agents on the same side of the market. Agents' standalone benefits from participation, b_{1i} or b_{2j}, label the nodes.

Armstrong (2006) by introducing heterogeneous network effects. With prices fixed, each of n_1 agents on side 1 of the market and n_2 agents on side 2 makes the binary decision whether to log onto the platform. The profile of these participation decisions is denoted by the vector $\mathbf{x} \equiv (\mathbf{x}_1, \mathbf{x}_2) \in \{0,1\}^{n_1} \times \{0,1\}^{n_2}$. If agent i on side 1 and agent j on side 2 both participate (i.e., $x_{1i} = x_{2j} = 1$), then each of them enjoys the same bilateral interaction benefit W_{ij}, which can be positive, negative, or zero. (We shall later partially relax the assumption that the benefit is the same for both agents i and j.) The matrix of interaction benefits—or network effects—is denoted by $\mathbf{W} \in \mathbb{R}^{n_1 \times n_2}$. Regardless of others' participation decisions, each agent i on side 1 of the market enjoys a standalone benefit b_{1i} from participation whenever he chooses to participate (i.e., whenever $x_{1i} = 1$). This benefit, too, can be positive, negative, or zero. We define b_{2j} analogously and collect side-1 and side-2 standalone benefits into the vector $\mathbf{b} \equiv (\mathbf{b}_1, \mathbf{b}_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. The bipartite graph in Figure 1 organizes the model parameters. Matching is all-to-all: all participating agents on one side of the market benefit from interacting with all participating agents on the other side of the market. The all-to-all assumption can be interpreted either literally or as a reduced form for the stochastic many-to-many matching whereby every pair of participating agents on the opposite sides is matched with some fixed probability that is independent of the number of agents in the market.

Our equilibrium concept, sBRE, is defined as the time limit of agents' "noisy" best-response dynamics in participation decisions (Blume, 1993; Kandori, Mailath, and Rob, 1993; Young, 1993; Mele, 2017). The dynamics has a randomly called-upon agent myopically best respond to others' last-period participation decisions as he perceives his standalone value from participation with a layer of i.i.d. noise. Because the induced Markov chain is finite (i.e., there are finitely many feasible participation profiles) and irreducible (i.e., best-response dynamics can lead one from any participation profile to any other participation profile), it admits a unique stationary distribution, which we call equilibrium. That is, sBRE exists and is unique. Because the Markov chain is aperiodic (i.e.,

it does not cycle deterministically), this sBRE is eventually reached from any initial participation profile. At the sBRE, any two agents' participation decisions are correlated conditional on others' participation decisions if and only if these two agents experience a nonzero benefit from interaction. The sign of the benefit corresponds to the sign of the correlation. Under appropriate assumptions on the probability distribution of the i.i.d. noise (parametrized by a positive scale parameter σ), the equilibrium probability of a participation profile $\mathbf{x}$ is the logit distribution (Proposition 1):

$$\mathbb{P}\left(\mathbf{x} \mid \mathbf{W}, \mathbf{b}\right) = \frac{\exp\left(\frac{1}{\sigma}\left(\mathbf{b}^{\top}\mathbf{x} + \mathbf{x}_1^{\top}\mathbf{W}\mathbf{x}_2\right)\right)}{\sum_{\mathbf{y}\in\{0,1\}^{n_1+n_2}} \exp\left(\frac{1}{\sigma}\left(\mathbf{b}^{\top}\mathbf{y} + \mathbf{y}_1^{\top}\mathbf{W}\mathbf{y}_2\right)\right)}. \tag{1}$$

The startup's first task is to estimate the model parameters $(\mathbf{W}, \mathbf{b})$ from the data set that consists of a panel of agent participation profiles drawn i.i.d. from the sBRE distribution. An example of such a startup is a dating platform that matches men and women. The platform records the online–offline status of every subscriber at time intervals sufficiently separated so that the resulting profiles can be treated as independent. If the data revealed that, say, Alice and Bob each tend to log in whenever the other is online, then the startup would legitimately conclude that Alice and Bob each benefited from matching with the other.

The startup's first instinct might be to reach for maximum-likelihood estimates for $(\mathbf{W}, \mathbf{b})$: to assemble the probabilities given by (1) into a likelihood function and then maximize this function by gradient ascent. This approach runs into combinatorial explosion: How can one evaluate the expression in (1) when the sum in its denominator has $2^{n_1+n_2}$ terms, a formidable number even for moderate values of $n_1 + n_2$ (e.g., $2^{30} \approx 1,000,000,000$)? And if one cannot evaluate (1) even once, how can one hope to do so over and over again so as to numerically maximize the likelihood?

Should the startup give up on heterogeneous network effects? Models of two-sided markets with homogeneous network effects in the tradition of Armstrong (2006) do not suffer from the combinatorial explosion. Indeed, under the homogeneity of both network effects and standalone benefits, the maximum-likelihood estimation described in the paragraph above is numerically tractable, as the $2^{n_1+n_2}$ terms in the denominator collapse to just $(n_1 + 1)(n_2 + 1)$. So, what happens if the startup neglects heterogeneity and prices under the misspecified model that assumes homogeneous network effects? Betting on the misspecified model may cost dearly. We show that, under the misspecified model, not only will the startup's inference about the model parameters be wrong, but the realized revenue, induced by misguided pricing decisions, may also be an arbitrarily small fraction of the optimal revenue (Proposition 5). Therefore, the startup has no choice but to tackle the model with heterogeneous network effects head-on.

Help arrives from an unexpected place. It turns out that the sBRE distribution over participation profiles $\mathbf{x} \equiv (\mathbf{x}_1, \mathbf{x}_2)$ in (1) is known in machine learning as a restricted Boltzmann machine (RBM) (Ackley, Hinton, and Sejnowski, 1985; Smolensky, 1986). Here, "Boltzmann" is another way of saying logit. "Restricted" means that any two components of a vector $\mathbf{x}_1$ are statistically independent conditional on the values of $\mathbf{x}_2$; *mutatis mutandis* for any two components of $\mathbf{x}_2$. "Machine" is an honorific that indicates that the probability distribution in (1) has a plethora of uses

in machine learning. Indeed, an RBM can recommend movies (Salakhutdinov, Mnih, and Hinton, 2007), recognize images (Larochelle and Bengio, 2008), and discover drugs (Wang and Zeng, 2013).

In machine learning, RBMs are trained. To train an RBM means to fit the values of $(\mathbf{W}, \mathbf{b})$ that make the probability distribution in (1) resemble the empirical distribution of the data. Contrastive divergence (CD)—proposed by Hinton (2002)—is the leading algorithm for training RBMs. A seemingly unprincipled corruption of likelihood maximization by gradient ascent, CD circumvents combinatorial explosion and is responsible for the resurgence of machine learning in the early 2000s. While a good training algorithm ensures soon enough that the trained distribution fits the empirical distribution closely enough, it is not usually tasked with uncovering underlying parameter values. In fact, the underlying parameter values are of little interest in machine learning because the RBM is not a structural model of anything; its parameters admit no interpretation. Not so in economics.

The startup must know the underlying parameters $(\mathbf{W}, \mathbf{b})$ in order to execute its second task: pricing well. Therefore, to be of use, CD must estimate $(\mathbf{W}, \mathbf{b})$ consistently. CD's consistency has received attention from statisticians, on whose results we rely. We show that Jiang, Wu, Jin, and Wong's (2018) theorem for the consistency of CD applies in the context of our two-sided market model (Proposition 2). The model's special feature that makes Jiang, Wu, Jin, and Wong's results applicable is the assumption that the startup observes participation decisions by agents on both sides of the market. That is, the entire vector $\mathbf{x} \equiv (\mathbf{x}_1, \mathbf{x}_2)$ is observed. By contrast, most applications of RBMs in machine learning assume that, say, only $\mathbf{x}_1$ is observed, whereas $\mathbf{x}_2$ is hidden and must be imputed.

While CD's consistency is a theorem, CD's numerical tractability is an empirical fact supported by over two decades of successful training of neural networks and by the simulations that we report later in the paper.

Nothing we know about the theory and practice of RBMs helps us tackle the startup's pricing problem. Even the problem of evaluating welfare or revenue numerically—let alone maximizing them over prices—exhibits combinatorial explosion. (The reason, once again, is the denominator in (1).) Therefore, we focus on the pricing problem's limit case in which the noise in agents' best responses—parameterized by σ in (1)—vanishes. This case has also been the focus of existing pricing literature. We construct a class of pricing schemes each of which achieves the first-best utilitarian welfare (Proposition 4). Without estimates for heterogeneous network effects and without corrective pricing that needs these estimates, equilibrium welfare can be an arbitrarily small fraction of the first-best welfare (Proposition 3). One of the schemes that maximizes welfare also maximizes revenue by fully extracting the first-best welfare (Proposition 6). This scheme, too, requires detailed estimates of heterogeneous network effects.

While we focus on two-sided markets, our estimation and pricing results continue to apply to "stacked" markets with more than two sides. An example is a streaming platform, which has four sides. Advertisers advertise to customers, who interact with artists, who interact with recording labels that represent them and manage the rights to their output. This market is said to be stacked because each side only interacts with its adjacent sides. While conceptually the four-sided market

described above is rather different from a two-sided one, formally it reduces to a two-sided market in which advertisers and artists comprise one side (thereby ruling out interactions within and between these two groups) and customers and labels comprise the other side (again, thereby ruling out within- and between-group interactions). Across these two sides, the network effects between advertisers and labels are restricted to zero. This restriction is enabled by the model of two-sided markets that permits network effects to be heterogeneous and would have been impossible in the model with homogeneous effects, in which agents simply count the number of counterparties.

The portal that our equilibrium notion, sBRE, opens into neural networks transcends applications to two-sided (and, a fortiori, to stacked multi-sided) markets. Once one introduces same-side externalities, one is no longer dealing with a two-sided market. Any agent's payoff may be affected by any other agent's participation. A universe of economic applications opens up. Among the examples are an online recruiting marketplace with same-side congestion (Example 1) and a corporate R&D portfolio problem (Example 2). Under our equilibrium notion, binary-action games of this kind map into unrestricted Boltzmann machines (as opposed to restricted ones, for two-sided markets). As will become apparent, our estimation technique carries over and remains consistent in this more general setting. Its speed may no longer scale with the size of the market, though.

Related Literature

Below we highlight three broad strands of related literature: on discovering common structures underlying economic and machine-learning problems, on two-sided markets, and on parameter identification and estimation on networks. Further relevant papers are discussed as the story develops.

Igami (2020) and Samuelson and Steiner (2023) have pioneered the literature that recognizes economic structures in seemingly unrelated machine-learning models. Igami (2020) observes that machine-learning algorithms behind the Deep Blue expert system for chess, the Bonanza engine for shogi (Japanese chess), and the AlphaGo system for Go are related to estimation techniques for dynamic structural models. Samuelson and Steiner (2023) study an economic growth–maximizing policy and relate it to the problem of predictive coding in machine learning. Analogies between problems in asset pricing and machine learning have been established by Avramov, Cheng, and Metzker (2023), Chen, Pelger, and Zhu (2024), and Su, Tretyakov, and Newton (2025). We likewise discover a novel connection: between an equilibrium outcome in an economic model and activation patterns in a canonical neural network. This discovery opens up possibilities for repurposing machine-learning results for economic use and extends beyond two-sided markets to binary-action games.

Our paper draws inspiration from Chan's (2021) incisive observation that a two-sided market with homogeneous network effects is a potential game. (A potential game is strategically equivalent to a common-interest game in which each player's payoff is represented by a function, shared by all players, whose different arguments are controlled by different players.) We notice that Chan's potential function coincides with the negative of the "energy function" in a corresponding restricted

Boltzmann machine. From here, we conjecture and verify that a model of two-sided markets with heterogeneous network effects, too, admits a potential, and that this potential coincides with the negative of the energy function of a restricted Boltzmann machine. Our equilibrium concept, sBRE, is a probability distribution over participation profiles and has Chan's potential-maximizing Nash equilibrium as its modal outcome. The applicability of sBRE extends beyond two-sided markets and illustrates Babichenko and Tamuz's (2016, Proposition 4.6) observation that every graphical potential game corresponds to some Markov random field.

Chan (2025) studies multi-agent contracting problems in which agents' payoffs constitute a weighted potential game. His objective is different from ours: he takes the payoff environment as known and characterizes weight-ranked divide-and-conquer contracts that are optimal for all equilibrium-selection criteria more pessimistic than potential maximization. Our use of potential is complementary. We use stochastic best-response dynamics to convert the potential into the energy of a Boltzmann distribution over action profiles, thereby exploiting the potential-game representation to obtain an estimable statistical model. In the exact-potential case and in weighted-potential games after a natural scaling of noise in agents' actions, sBRE assigns each participation profile a probability that is increasing in this profile's potential. Thus, potential maximization—the equilibrium-selection criterion central to Chan's analysis—is the modal prediction of sBRE and its limit as mistakes in agents' actions vanish. In this sense, Chan's deterministic implementation problem and our stochastic estimation problem use the same mathematical object, the potential, but for different purposes.

Farronato, Fong, and Fradkin (2024) formulate a simple theoretical model of platforms in order to examine the countervailing effects that merging two platforms has on welfare. A merger enables participants to gain from interacting with a greater number of counterparties, but, because network effects are heterogeneous, participants suffer from losing the ability to choose who to interact with by choosing which of the two distinct platforms to join. Instead of estimating the model structurally, Farronato, Fong, and Fradkin estimate the effects of the merger via a reduced-form, difference-in-differences approach. By contrast, our model imposes little structure on heterogeneity and lends itself to numerically tractable structural estimation.

Estimation techniques for two-sided markets in which each agent cares about not only how many but who exactly is on the other side of the market have been developed that circumvent the curse of dimensionality by reducing agents to a bundle of a small number of characteristics. Hitsch, Hortaçsu, and Ariely (2010) study a one-to-one online dating market. They reduce agents to bundles of characteristics (e.g., age, height, BMI, income, and education) and then estimate the weights that describe how much each dater values each of a prospective partner's observed characteristics. In a model of game consoles, Lee (2013) similarly reduces video games to their prices, genres, and estimated qualities. Both papers accommodate heterogeneity by making the value of a pairwise interaction a function of measured characteristics; what then must be estimated are coefficients on these characteristics, whose number is independent of the size of the market. Our paper proceeds differently. It requires no characteristics at all. Instead, it estimates the matrix $\mathbf{W}$

of pair-specific benefits directly, inferring each pair's "chemistry" from how often they participate together across repeated observations of the market. What our approach gains in flexibility—it can recover (dis)affinities that no observed characteristics would predict—it loses in the model's ability to extrapolate to newcomers. The estimates in $\mathbf{W}$ describe only the agents in the data set and do not extend to a potential newcomer. (Of course, if extrapolation is possible, one can recover the earlier literature's summary preferences after the fact, by regressing the estimated $\mathbf{W}$ on the observable characteristics.)

Mele's (2017) seminal work is concerned with parameter identification and estimation in a setting related to, but distinct from, ours. He identifies model parameters off a single observation of a large dense network when agents engage in stochastic best-response dynamics. Mele's estimation technique of choice is Markov Chain Monte Carlo (MCMC). We, by contrast, identify model parameters off repeated observations of the same network and advocate contrastive divergence (CD). For large binary-action games, as an empirical matter, MCMC is prohibitively slow, whereas CD is fast in the special case of two-sided markets.

2 The Model

Conforming with Armstrong's (2006) workhorse model, our market has two sides, with n_1 agents on side 1, n_2 agents on side 2, and $n \equiv n_1 + n_2$ agents in total. Each agent on side 1 takes the binary decision $x_{1i} \in \{0,1\}$ whether to join the platform, with $x_{1i} = 1$ corresponding to joining and $x_{1i} = 0$ corresponding to not joining. Let $\mathbf{x}_1 \equiv (x_{1i})_{i=1}^{n_1} \in \{0,1\}^{n_1}$ denote the vector of participation decisions for side 1, let $\mathbf{x}_2 \equiv (x_{2j})_{j=1}^{n_2} \in \{0,1\}^{n_2}$ denote the corresponding vector for side 2, and let $\mathbf{x} \equiv (\mathbf{x}_1, \mathbf{x}_2) \in \{0,1\}^n$ denote the vector that is the entire participation profile.

Agent i on side 1 enjoys the standalone value $b_{1i} \in \mathbb{R}$ from participation, which can be positive, negative, or zero. Let $\mathbf{b}_1 \equiv (b_{1i})_{i=1}^{n_1}$ denote the vector of standalone values for side 1. Define $\mathbf{b}_2 \equiv (b_{2j})_{j=1}^{n_2}$ analogously. If agents i and j on sides 1 and 2, respectively, both participate, then, in addition to the standalone value, each of them enjoys the same interaction value $W_{ij} \in \mathbb{R}$, which can be positive, negative, or zero. (As we discuss shortly and then in Section 7.3, the assumption that a pair of agents split the bilateral surplus from interaction equally can be relaxed.) Let $\mathbf{W} \equiv (W_{ij})_{i,j}$ denote the $n_1 \times n_2$ matrix of interaction effects. The vector of interaction effects faced by agent i on side 1 is denoted by $\mathbf{W}_{i\bullet} \equiv (W_{ij})_{j=1}^{n_2}$. The vector of interaction effects faced by agent j on side 2 is denoted by $\mathbf{W}_{\bullet j} \equiv (W_{ij})_{i=1}^{n_1}$.

Fix a participation profile $\mathbf{x}$. The payoffs of an agent i on side 1 and an agent j on side 2 are, respectively,

$$x_{1i}\left(b_{1i} + \mathbf{W}_{i\bullet}^{\top}\mathbf{x}_2\right) \qquad \text{and} \qquad x_{2j}\left(b_{2j} + \mathbf{W}_{\bullet j}^{\top}\mathbf{x}_1\right),$$

where "$\top$" denotes transpose. From each agent's perspective, the maximization of his payoff with respect to his participation decision is equivalent to the maximization of the function $\mathbf{b}^{\top}\mathbf{x} + \mathbf{x}_1^{\top}\mathbf{W}\mathbf{x}_2$ with respect to the same decision. Games that admit such a function are called potential games, and the common maximand is called the **potential** (Monderer and Shapley, 1996).

Discussion of the Model's Assumptions

The all-to-all matching protocol implicit in the definition of payoffs is degenerate: everyone on one side of the market matches to everyone on the other side. Or such is the literal interpretation of the model. Equivalently, one can interpret the parameter W_{ij} as the expected benefit from the bilateral interaction that occurs with some probability whenever both i and j participate. This interaction probability can depend on the underlying benefit from interaction, so that, say, the pairs that value interaction more are more likely to match. What is ruled out is the dependence of this probability and of the underlying benefit on who else shows up. Formally, $\mathbf{W}$ does not depend on $\mathbf{x}$.

The assumption that agent i enjoys interacting with agent j exactly as much as j enjoys interacting with i is restrictive. The assumption describes situations in which fairness, legal norms, or equal bargaining powers dictate that the total surplus from the bilateral interaction, denoted by $2W_{ij}$, be split equally, with each agent getting W_{ij}. It is straightforward to extend the model by assuming that all side-1 agents capture, say, 30% of the bilateral surplus, and side-2 agents capture the remaining 70%. The cost of accommodating this extension is only additional notation. A further extension that would allow the benefits from interaction to differ both within and across pairs in arbitrary ways cannot be analyzed using the tools of this paper.

3 Equilibrium

Our equilibrium concept is the **stochastic best-response equilibrium** (**sBRE**). It is defined as the limit distribution of stochastic best-response dynamics, in which randomly selected agents myopically update their strategies given other agents' actions (Blume, 1993; Kandori, Mailath, and Rob, 1993; Young, 1993; Mele, 2017). The dynamics plays out in discrete steps, indexed by s. The dynamics is assembled from three components: an initial participation profile $\mathbf{x}^{[0]} \in \{0,1\}^n$, an action-revision protocol that specifies which agents are called upon to revise their actions and when, and a stochastic best-response rule according to which the called-upon agent or agents revise their actions.

The first component—the initial participation profile $\mathbf{x}^{[0]}$—will not affect the limit distribution.

Any revision protocol in a certain class will lead to the same limit distribution. This class of protocols is embedded in the equilibrium definition presented shortly (Definition 1) and includes two prominent members: random scanning and blocked sampling. **Random scanning** selects an agent uniformly at random and asks him to revise his action. Random scanning is applicable to a large class of models. **Blocked sampling** first selects all agents on one side of the market to simultaneously revise their actions, and then selects all agents on the other side to do so. Blocked sampling relies on the market's bipartite structure.

At each step $s \in \{1,2,\ldots\}$ of the stochastic **best-response dynamics**, every called-upon agent i on side 1 of the market draws a standard logistic random variable $\varepsilon_{1i}^{[s]}$ that is independent of all past draws and of other agents' contemporaneous draws (if any). He then stochastically best-responds to side 2's past actions $\mathbf{x}_2^{[s-1]}$ by revising his binary participation decision $x_{1i}^{[s]}$ according

to

$$x_{1i}^{[s]} = \begin{cases} 1 & \text{if} \quad b_{1i} + \mathbf{W}_{i\bullet}^{\top} \mathbf{x}_2^{[s-1]} + \sigma \varepsilon_{1i}^{[s]} > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{2}$$

where σ is a positive scalar, which, in contrast to Kandori, Mailath, and Rob (1993) and Young (1993), we do not insist on sending to zero. Every called-upon agent j on side 2 best-responds analogously:

$$x_{2j}^{[s]} = \begin{cases} 1 & \text{if} \quad b_{2j} + \mathbf{W}_{\bullet j}^{\top} \mathbf{x}_1^{[s-1]} + \sigma \varepsilon_{2j}^{[s]} > 0 \\ 0 & \text{otherwise} \end{cases}.$$

We interpret the epsilon-induced stochasticity in best responses as mistakes. The alternative interpretation of the epsilons ($\varepsilon_{1i}^{[s]}$ and $\varepsilon_{2j}^{[s]}$) in (2) is as shocks to utility. This alternative interpretation is compatible with our analysis of identification and estimation but is incompatible with our foray into pricing later in the paper.

We can now assemble the parts.

Definition 1. A **stochastic best-response equilibrium** (**sBRE**) is the stationary probability distribution of the Markov chain in agents' participation decisions induced by the stochastic best-response dynamics at whose every step,

- an irreducible (but possibly periodic) two-state Markov chain selects a side of the market;
- on the selected side of the market, a subset of agents is chosen according to some fixed (for that side of the market) probability distribution that chooses every agent with a positive probability;
- each chosen agent stochastically best-responds to others' past actions.

The action-revision protocol in Definition 1 is the aforementioned random scanning if, at each step, side 1 is selected with the probability n_1/n, and side 2 is selected with the complementary probability n_2/n, and then one of the agents on the selected side is chosen uniformly at random. The protocol is the aforementioned blocked sampling if side 1 is selected at odd steps, and side 2 is selected at even steps, and all agents on the selected side are chosen. Definition 1 accommodates much more general revision protocols than these two, though. It admits persistence and arbitrary imbalance in selecting the side of the market that gets to revise, and it permits different agents to be chosen with different probabilities and in a correlated manner.

Proposition 1 (Equilibrium Characterization)**.** *The stochastic best-response equilibrium exists, is unique, and, to each participation profile* $\mathbf{x}$*, assigns the probability*

$$\mathbb{P}\left(\mathbf{x} \mid \mathbf{W}, \mathbf{b}\right) = \frac{\exp\left(\frac{1}{\sigma}\left(\mathbf{b}^{\top}\mathbf{x} + \mathbf{x}_1^{\top}\mathbf{W}\mathbf{x}_2\right)\right)}{\sum_{\mathbf{y}\in\{0,1\}^n} \exp\left(\frac{1}{\sigma}\left(\mathbf{b}^{\top}\mathbf{y} + \mathbf{y}_1^{\top}\mathbf{W}\mathbf{y}_2\right)\right)}. \tag{3}$$

It is known that the logit distribution in (3) is the unique stationary distribution under random scanning and best-response dynamics in potential games, such as ours (see, e.g., Blume, 1993, and Mele, 2017). Proposition 1 asserts that the same remains true under a great variety of revision protocols. Although we believe that a version of Proposition 1 should be known, we have not found a reference to the proof, and, therefore, provide one in Appendix A.1.

Proposition 1 admits an asymptotic analogue under weaker observability assumptions, which do not require—implausibly in large markets—that each player observe what everyone else does. Remark 1 formalizes this analogue by assuming that, in a market of n agents, each agent observes a random sample of k others, drawn uniformly at random at each step. For each market size, the stationary distribution induced by sampled interactions approaches the distribution in (3) for that same market. The remark's conclusion relies on substantive asymptotic restrictions: the subsample cannot be too small ($k \propto n^{0.51}$ would work), and network effects do not swamp the stochasticity in agents' choices.

Remark 1. Suppose that, as $n \to \infty$, the components of $\mathbf{b}$ are bounded uniformly, $|W_{ij}|$ is bounded uniformly by $\bar{W}/(n-1)$ for some $\bar{W} < 4\sigma$, and $\sqrt{n}/k \to 0$. Then, as $n \to \infty$, the stationary distribution of participation profiles under sampled observation and random scanning becomes arbitrarily close to the sBRE distribution in (3) for the corresponding n-agent market.

The proof of Remark 1 is not immediate and is deferred to Hoshino and Pancs (2026).

Stochastic Best-Response Equilibrium in Context

sBRE's four critical features will enable us to quickly learn model parameters by observing i.i.d. draws from the equilibrium probability distribution of participation profiles:

- The participation decisions of interacting agents are correlated, and more so for the agents who benefit from interacting more. The strength of this correlation reveals the magnitude of network effects.
- The equilibrium distribution has full support; that is, all participation profiles are possible. As a result, the data are capable of exhibiting rich enough correlation structures to identify all network effects.
- The equilibrium is unique. In our model, uniqueness delivers point identification of model parameters.
- The probability distribution of participation profiles is logit. This functional form lends the model tractability, which is crucial for fast estimation.

Each of these features can be motivated by some alternative equilibrium concepts and is inconsistent with some others. We shall now place sBRE in the context of these alternatives. The Venn diagram in Figure 2 is a roadmap for the discussion.

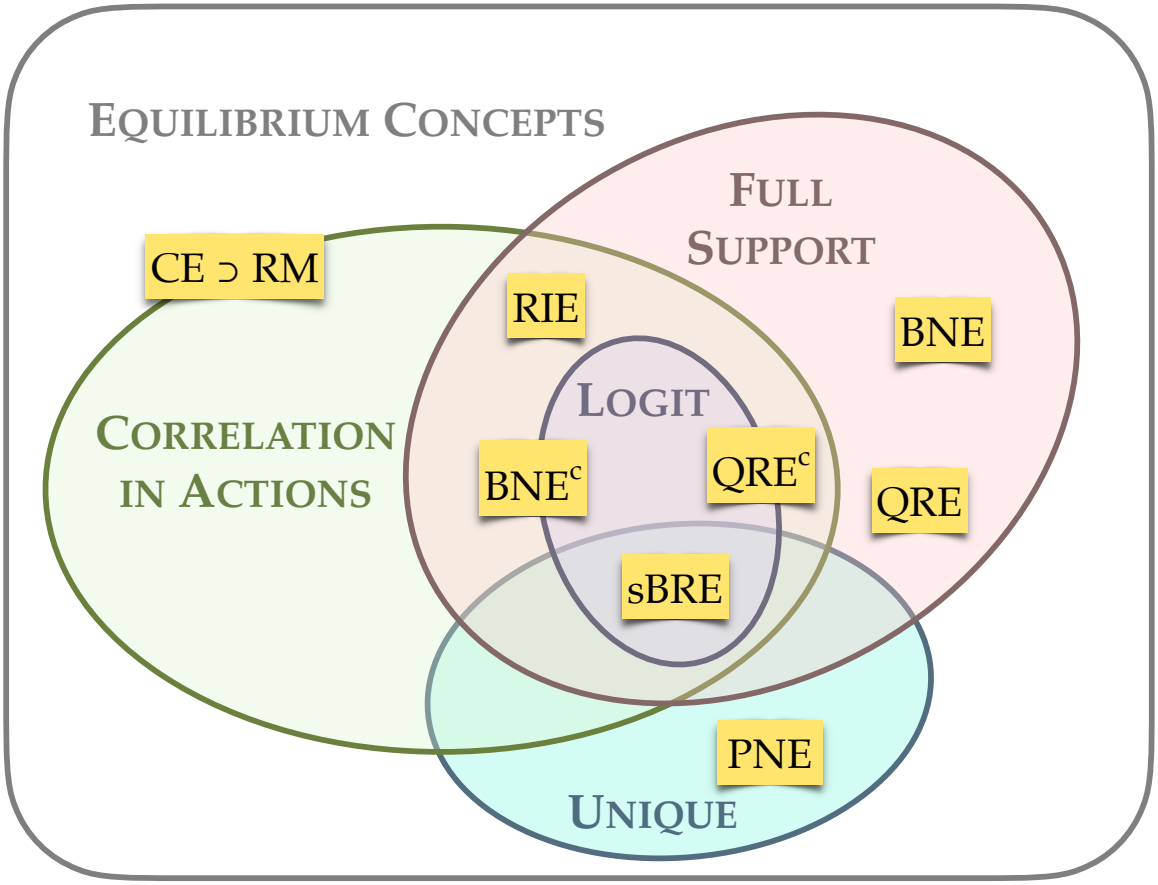

Figure 2: **Stochastic Best-Response Equilibrium in Context: the Venn Diagram.** sBRE = stochastic best-response equilibrium, RIE = rational inattention equilibrium, QRE = quantal response equilibrium with independent mistakes, QRE^c = quantal response equilibrium with correlated mistakes, BNE = Bayes–Nash equilibrium with independent payoff shocks, BNE^c = Bayes–Nash equilibrium with correlated payoff shocks, CE = correlated equilibrium, RM = regret minimization, PNE = potential-maximizing Nash equilibrium.

The best-response dynamics that undergirds sBRE effectively enables agents to spy on each other's actions before choosing their own actions. Spying begets correlation. This correlation is stronger when the benefit from the bilateral interaction is larger and, so, is likelier to swamp the random trembles—the epsilons—in best responses. An alternative way to operationalize spying is Denti's (2023) rational inattention equilibrium (RIE) (which Denti calls "equilibrium strongly robust to information acquisition"). At an RIE, agents simultaneously acquire noisy signals about others' equilibrium actions in a fixed-point manner reminiscent of the rational expectations equilibrium with private information à la Grossman and Stiglitz (1980) (Hébert and La'O, 2023). Compared to sBRE, RIE suffers from two limitations: the multiplicity of RIEs hampers point identification, and RIE's inability to deliver the especially tractable functional form in (3) hampers estimation. Even when the cost of information acquisition is entropic, the equilibrium distribution—biased logit—lacks the supreme tractability of (3).

Correlation in participation decisions can be alternatively generated by replacing spying with a randomization device and a mediator, thereby invoking the correlated equilibrium (CE). CE shares RIE's problems, though, and has more. CEs are typically multiple. Some CEs exhibit no correlation in participation decisions. Among those that do, none may conform with the logit functional form in (3).

Related to CE is Hart and Mas-Colell's (2000) regret-minimization (RM) protocol. The protocol differs from the best-response dynamics that motivates sBRE and always converges to a CE.

The existence of the aforementioned alternatives (along with some other alternatives we discuss shortly) suggests that sBRE's core feature—the correlation in participation decisions—is not fragile. It can emerge for a variety of reasons. Some common equilibrium concepts do resist correlation, though. In order to circumscribe the model's scope, we briefly discuss these concepts.

Just like sBRE, McKelvey and Palfrey's (1995) quantal-response equilibrium (QRE) assumes that agents make mistakes. As long as QRE maintains sBRE's assumption that these mistakes are statistically independent, it rules out correlation in agents' participation decisions. At QRE, agents best respond to the equilibrium probability distribution of others' actions rather than to the actions themselves. There is no mechanism to correlate actions. Such a mechanism can be introduced, however, by correlating mistakes. Let QRE^{c} denote QRE with correlated mistakes. A joint probability distribution over all agents' mistakes can be manufactured to make QRE^{c}'s probability distribution over participation profiles match the sBRE (or any other probability distribution; see Haile, Hortaçsu, and Kosenok 2008, Theorem 1). Neither QRE nor QRE^{c} need be unique. Full support, by contrast, arises naturally when mistakes are unbounded.

In our setting, the Bayes–Nash equilibrium (BNE) is formally equivalent to QRE if one interprets the stochasticity in agents' actions as arising from payoff shocks rather than mistakes. As a result, BNE, too, rules out correlation in agents' actions unless one assumes that payoff shocks are correlated.

The Nash equilibrium (NE) would seem a natural solution concept, but it predicts both too little and too much. NE predicts too little because platform-participation games may exhibit strategic

complementarities and be plagued by the associated multiplicity of NEs. The multiplicity problem admits a compelling solution, though. In two-sided market settings, Chan (2021) proposes to select the NE with the maximal potential. This refinement, denoted by PNE, is compelling because it is robust to the introduction of small amounts of incomplete information (Kajii and Morris, 1997; Ui, 2001). PNE is related to sBRE in that it is both sBRE's modal prediction and sBRE's limit as $\sigma \to 0$. NE predicts too much because it violates full support by assigning probability zero to most participation profiles, thereby forcing one to focus on the applications in which the data are not rich enough to identify model parameters. PNE shares in this shortcoming.

QRE^{c} and BNE^{c} each possess some of sBRE's critical features: correlation, full support, and the logit distribution. Even so, we favor sBRE. While QRE^{c} and BNE^{c} directly assume how the joint probability distribution over agents' mistakes or payoff shocks depends on the model parameters, sBRE treats mistakes as independent and instead derives the statistical dependence of agents' participation decisions from payoffs in an intuitive way.

4 Restricted Boltzmann Machines: An Interlude

Conceived by Smolensky (1986), the restricted Boltzmann machine (RBM) is a parameterized probability distribution of vectors $\mathbf{x} \equiv (\mathbf{x}_1, \mathbf{x}_2)$ with values in $\{0,1\}^n$. The distribution's dependence structure is represented by a complete undirected bipartite graph. With each of its nodes, we associate a binary random variable. Any two nodes on different sides of the graph—in different layers of the "neural net" the graph represents—are connected by an edge, meaning that the model permits the corresponding random variables to be statistically dependent conditional on the values of the remaining variables; they are conditionally independent when the associated weight is zero. Any two nodes in the same layer are not connected by an edge, meaning that the corresponding random variables are statistically independent conditional on the values of the remaining variables. The strength of statistical dependence is not pinned down by the graph itself and must be assumed separately. An RBM assumes that $\mathbf{x}$ is distributed according to the distribution in (3), where $\mathbf{W}$ and $\mathbf{b}$ are arbitrary parameters. That is, a probability distribution can be generated by some RBM if and only if it is the sBRE of some two-sided market.

In the context of machine learning, an RBM is classified as a neural network. It strikes a balance between tractability and expressiveness. Early on, RBMs proved successful at recognizing handwritten characters and recommending movies, among other tasks. Later on, RBMs became building blocks of more complex neural networks called deep Boltzmann machines.

To accomplish a task, first, an RBM must be trained. In most applications, only one layer of the RBM maps directly into data. This layer is said to be visible. For example, if the task at hand is to recommend movies based on past user experience, then the visible layer comprises each user's preferences over movies: thumbs up or thumbs down for every movie. If the task is to recognize handwritten characters, then the visible layer lines up the pixels in the black-and-white pixelation of a handwritten character (on or off for each pixel) and the corresponding correct answers (for

each handwritten character, whether it is 1, 2, or 3, etc.). In both applications, the other, hidden, layer comprises the latent variables whose sole purpose is to introduce a rich enough dependence structure among the visible variables. These latent variables do not directly correspond to anything observed in the data. The values of these variables are imputed in the course of training. Their interpretations are open-ended. In the movie-recommendation application, latent variables may correspond to genres and actors. In the character-recognition application, latent variables may correspond to strokes shared by multiple characters.

Training an RBM consists in choosing the underlying parameters $(\mathbf{W}, \mathbf{b})$ in such a way that the fitted distribution of the visible variables $\mathbf{x}_1$ generated by the RBM is close to the empirical distribution of $\mathbf{x}_1$ in the training data set. Training an RBM is analogous to estimating the parameters of a two-sided market when data are generated by sBRE with two exceptions:

1. In RBMs, only one layer is typically visible, whereas in two-sided markets, participation on both sides of the market is visible.

2. The goal of training RBMs is to fit the data, whereas the goal of estimating parameters of a two-sided market is to estimate the true parameter values of the underlying structural model.

The distinction between fit and estimation matters for the startup's second task: pricing. While a good fit helps predict agent participation at the prices at which the model has been trained, to price well, the startup must be able to predict agent participation in a variety of counterfactual scenarios. Positive prices discourage participation, negative prices (i.e., subsidies) encourage it, and to predict the exact magnitude of these forces, one must know the underlying model parameters.

5 Learning the Two-Sided Market Parameters

The startup has access to T i.i.d. draws from the sBRE. These draws comprise the data set, denoted by $(\mathbf{x}(t))_{t=1}^{T}$. We assume that these data are generated while the startup does not charge any prices. This assumption is inconsequential and is made to conserve notation.

5.1 Likelihood Maximization

We henceforth assume that the scale-of-noise parameter σ is known and focus on identifying the remaining model parameters $(\mathbf{W}, \mathbf{b})$. If the ultimate goal is to predict behavior, then the choice of σ is mere normalization of the units in which the components of $(\mathbf{W}, \mathbf{b})$ are measured. If the ultimate goal is to price optimally, then $(\mathbf{W}, \mathbf{b})$ must be compared to prices and, so, must be measured in dollars. In practical applications, to pin down the units, one can use the following trick: Modify the two-sided market so that whenever, say, Alice and Bob both show up, instead of letting them interact, the platform tells them to stay away from each other and instead pays each of them a dollar. As a result, by construction, $W_{Alice,Bob} = \$1$. The estimates for $(\mathbf{W}, \mathbf{b})$ assuming an arbitrary σ are then scaled to respect $W_{Alice,Bob} = \$1$; the scaling factor reveals the underlying σ.

We henceforth maintain the technical assumption that the true parameter values $(\mathbf{W}, \mathbf{b})$ are confined to the interior of a certain compact and convex set, no matter how large.

Lemma 1 exhibits an estimator that is **consistent** in the sense of converging in probability to the true parameters $(\mathbf{W}, \mathbf{b})$ as the sample size T grows. As a result, the model parameters are identified.

Lemma 1. *The maximum-likelihood estimator*

$$\left(\mathbf{W}^{ML}, \mathbf{b}^{ML}\right) \equiv \arg\max_{\tilde{\mathbf{W}}, \tilde{\mathbf{b}}} L \equiv \frac{1}{T} \sum_{t=1}^{T} \log \mathbb{P}\left(\mathbf{x}\left(t\right) \mid \tilde{\mathbf{W}}, \tilde{\mathbf{b}}\right) \tag{4}$$

with $\mathbb{P}$ defined in (3) is consistent for $(\mathbf{W}, \mathbf{b})$.

Proof. Consistency follows from the strict concavity of the log-likelihood function L. Strict concavity follows from Proposition 3.1 of Wainwright and Jordan (2008, p. 62) because sBRE is a member of a minimal exponential family. ■

The maximum-likelihood estimator in Lemma 1 achieves identification by matching moments. Roughly speaking, Alice's standalone value from participation is identified off her empirical frequency of participation, and her benefit from interacting with Bob is identified off the empirical correlation between the pair's participation decisions. Formally, this logic is captured by the first-order conditions evaluated at the maximum-likelihood (ML) estimates $\left(\mathbf{W}^{ML}, \mathbf{b}^{ML}\right)$:

$$\begin{aligned} \sum_{\mathbf{y} \in \{0,1\}^n} \mathbf{y} \mathbb{P}\left(\mathbf{y} \mid \mathbf{W}^{ML}, \mathbf{b}^{ML}\right) &= \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}\left(t\right) \\ \sum_{\mathbf{y} \in \{0,1\}^n} \mathbf{y}_1 \mathbf{y}_2^{\top} \mathbb{P}\left(\mathbf{y} \mid \mathbf{W}^{ML}, \mathbf{b}^{ML}\right) &= \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}_1\left(t\right) \mathbf{x}_2\left(t\right)^{\top}. \end{aligned} \tag{5}$$

The left-hand sides in (5) are the moments of the estimated distribution and the right-hand sides are the corresponding moments in the data. Even though, in general, moment equality need not deliver identification (Heyde, 1963), in our model it does, by Lemma 1. System (5) admits no explicit solution.

Conceptually, our model is identified because, from each agent's perspective, his counterparties' realized last-period actions are by assumption predetermined relative to his current-period private shock and act as "exogenous shifters." The exogenous variation in others' last-period actions moves the agent's best response and point-identifies the corresponding network effect. No such identification would have been possible under the Bayes–Nash equilibrium, under which moves are simultaneous, and the reflection problem prevents point identification (Manski, 1993).

5.2 Contrastive Divergence

Likelihood maximization presents computational challenges. The likelihood function suffers from **combinatorial explosion**: in equation (3), the 2^n terms in the denominator $\sum_{\mathbf{y} \in \{0,1\}^n} \exp\left(\frac{1}{\sigma}\left(\mathbf{b}^{\top}\mathbf{y} + \mathbf{y}_1^{\top}\mathbf{W}\mathbf{y}_2\right)\right)$

are too many to compute even for moderate values of n. That is, computing the likelihood function—let alone maximizing it—is a numerically intractable task.

Contrastive divergence (CD) is a bastardization of likelihood maximization by gradient ascent. Its questionable pedigree notwithstanding, it turns out to be consistent (in the sense made precise below).

In order to introduce CD, let us start with gradient ascent for likelihood maximization. Set $\sigma = 1$. The gradient with respect to $\mathbf{b}$ evaluated at step k of gradient ascent with $\left(\mathbf{W}^{[k]}, \mathbf{b}^{[k]}\right)$ being the current guess for parameter values is:

$$\nabla_{\mathbf{b}} L^{[k]} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}(t) - \sum_{\mathbf{y}\in\{0,1\}^n} \mathbf{y}\mathbb{P}\left(\mathbf{y} \mid \mathbf{W}^{[k]}, \mathbf{b}^{[k]}\right). \tag{6}$$

The gradient with respect to $\mathbf{W}$ can be analyzed analogously. In (6), the first sum involves only data and is easy to compute. The second sum contains the intractable number 2^n of terms, each of which contains the intractable function $\mathbb{P}\left(\mathbf{y} \mid \mathbf{W}^{[k]}, \mathbf{b}^{[k]}\right)$. This combinatorial explosion is due to the heterogeneity of network effects. When network effects are homogeneous (i.e., when all entries of $\mathbf{W}$ are the same), both the expression for the equilibrium probability $\mathbb{P}$ in (3) and the sum in the gradient (6) simplify, and the number of terms in the sums drops from 2^n down to $O\left(n^2\right)$, thereby making the maximum-likelihood estimation of the model feasible.

Towards making progress on gradient ascent, note that the intractable sum in (6) is an expectation. We can approximate this expectation with the tractable sum:

$$\nabla_{\mathbf{b}} L^{[k]} \approx \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}(t) - \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}^{(\infty)}(t), \tag{7}$$

where each $\mathbf{x}^{(\infty)}(t)$ is ideally an i.i.d. draw from the distribution $\mathbb{P}\left(\cdot \mid \mathbf{W}^{[k]}, \mathbf{b}^{[k]}\right)$ whose unknown expectation we seek to approximate. We approximate each draw $\mathbf{x}^{(\infty)}(t)$ by starting out at the t-th data point $\mathbf{x}(t)$ and then evolving it via a sequence of m stochastic best-response steps under the assumption that the model parameters are $\left(\mathbf{W}^{[k]}, \mathbf{b}^{[k]}\right)$; the participation profile so obtained is denoted by $\mathbf{x}^{(m)}(t)$. Each best-response step may be a random-scanning step or a blocked-sampling step, as defined in Section 3. By Proposition 1, as the number of stochastic best-response steps m goes to infinity, stochastic best-response dynamics culminates in an $\mathbf{x}^{(m)}(t)$ that converges in distribution to $\mathbf{x}^{(\infty)}$, a draw from $\mathbb{P}\left(\cdot \mid \mathbf{W}^{[k]}, \mathbf{b}^{[k]}\right)$. Then, the average $\frac{1}{T}\sum_t \mathbf{x}^{(m)}(t)$ of the T draws so obtained approximates the sought expectation $\sum_{\mathbf{y}} \mathbf{y}\mathbb{P}\left(\mathbf{y} \mid \mathbf{W}^{[k]}, \mathbf{b}^{[k]}\right)$. In statistics, the described best-response dynamics until convergence in distribution goes by the name **MCMC**, or Markov chain Monte Carlo. Because convergence is in distribution only, the approximate gradient on the right-hand side of (7) is a random variable.

The problem with running MCMC for a "large" number m of steps at every iteration k of gradient ascent is that the process is impractically slow. This slowness is an empirical fact. Of course, one can speed up MCMC by taking m to be "small" (e.g., $m = 5$ or even $m = 1$), but then

convergence to a random draw from $\mathbb{P}\left(\cdot \mid \mathbf{W}^{[k]}, \mathbf{b}^{[k]}\right)$ will no longer occur. This unprincipled act of despair—taking m to be "small"—has a name. It is called **contrastive divergence** (CD). While the moniker itself is whimsical, its existence is hopeful: someone must have found the unprincipled approach that invalidates the approximate equality in (7) to "work." And work it does, in a sense that we now make precise.

Let CD-m denote contrastive divergence whose every chain of stochastic best-response dynamics described above is terminated after just m steps. CD-m amounts to "nongradient" ascent whereby, at every iteration k, instead of climbing the likelihood function by following the computationally intractable gradient (7), we attempt to climb the likelihood function by following the computationally tractable **nongradient**:

$$\Delta_{\mathbf{b}} L^{[k]} \equiv \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}\left(t\right) - \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}^{(m)}\left(t\right)$$

—and its analogously defined counterpart $\Delta_{\mathbf{W}} L^{[k]}$—in full realization that $\Delta_{\mathbf{b}} L^{[k]} \not\approx \nabla_{\mathbf{b}} L^{[k]}$ and $\Delta_{\mathbf{W}} L^{[k]} \not\approx \nabla_{\mathbf{W}} L^{[k]}$. Let $\left(\mathbf{W}^{[k,m]}, \mathbf{b}^{[k,m]}\right)$ denote the value of the CD-m estimator as described in the paragraph above after k iterations of nongradient ascent.

Proposition 2. *Contrastive divergence is consistent for* $(\mathbf{W}, \mathbf{b})$ *in the sense that there exists a positive integer* m *such that, as the sample size* T *grows, the CD-*m *estimator*

$$\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} \left(\mathbf{W}^{[k,m]}, \mathbf{b}^{[k,m]}\right)$$

converges in probability to $(\mathbf{W}, \mathbf{b})$.

The technical nature of the proof of Proposition 2 is responsible for its relegation to Appendix A.2. Our proof relies on Jiang, Wu, Jin, and Wong's (2018) exacting study of CD for the models in the canonical exponential family. We show that our model belongs to this family and, therefore, CD is consistent if the assumptions in Jiang, Wu, Jin, and Wong's Theorem 2.1 can be verified. We verify these assumptions by following the roadmap in Jiang, Wu, Jin, and Wong's example (their Section 4.2) of a fully visible RBM with $n_1 = n_2 = 1$.

Even though motivated by the problem of training RBMs, Jiang, Wu, Jin, and Wong's (2018) consistency result does not cover the vast majority of RBMs, which have a hidden layer. Our model is covered because both sides of the market are assumed to be visible.

5.3 The Comparative Tractability of Contrastive Divergence

While CD's consistency is a formal proposition, its numerical tractability under blocked sampling is an empirical fact. This fact is suggested by two decades of neural-network training, during which practitioners would brazenly set the "positive integer m" of Proposition 2 to 1 and proceed to train large RBMs fast and effectively. For instance, Salakhutdinov, Mnih, and Hinton's (2007) Netflix recommendation system is a CD-trained RBM that contains about 10^4 nodes and 10^7 parameters. Salakhutdinov and Hinton's (2012) deep Boltzmann machine is similarly sized.

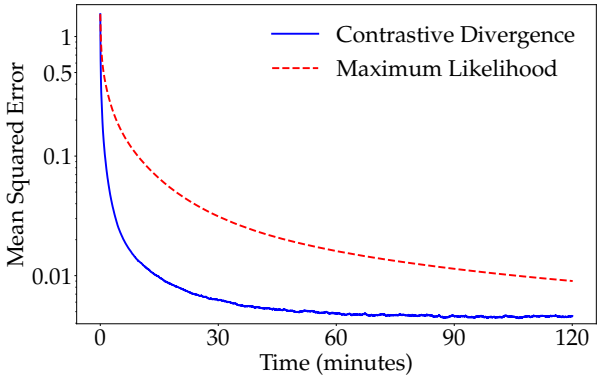

Figure 3: **With 11 agents on each side, conventional likelihood maximization is slow, whereas contrastive divergence is fast.** The underlying simulated data set has $T = 100,000$ observations.

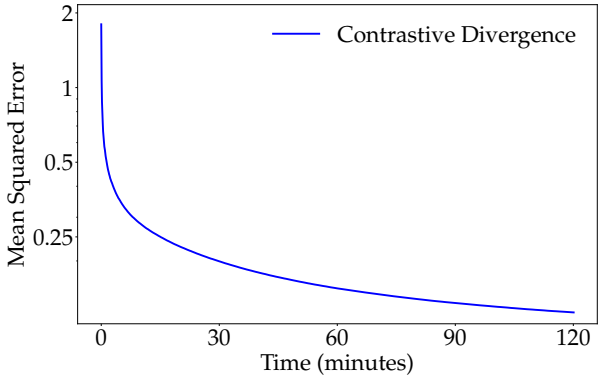

Figure 4: **With 15 agents on each side, contrastive divergence continues to converge.** The underlying simulated data set has $T = 500,000$ observations.

CD's numerical tractability is corroborated by our simulations, to which we now turn. All simulations have been conducted using Google Colab Pro, a cloud-based platform that provides access to high-performance computing resources. We used the GPU (Graphics Processing Unit) accelerator NVIDIA Tesla T4, which is optimized for machine-learning tasks. For transparency and reproducibility, our code is available at this link.

In each simulation, we first draw the underlying "true" parameters $(\mathbf{W}, \mathbf{b})$ from the multivariate standard normal distribution. We then synthesize a fictitious data set by drawing T participation profiles from the sBRE distribution induced by the parameters $(\mathbf{W}, \mathbf{b})$ and $\sigma = 1$. We then see how capable CD and likelihood maximization (ML) are of uncovering $(\mathbf{W}, \mathbf{b})$ under the assumption that $\sigma = 1$ by reporting each estimate's mean squared error in the course of each algorithm's execution. Here, CD stands for CD-m with $m = 1$ under blocked sampling, and ML stands for likelihood maximization under gradient ascent with each $2^{n_1+n_2}$-term sum computed using vectorization. Both CD and ML are initialized from the same random guess $\left(\mathbf{W}^{[0]}, \mathbf{b}^{[0]}\right)$.

Figure 3 illustrates how much faster CD converges to $(\mathbf{W}, \mathbf{b})$ than ML does in a market with $n_1 = n_2 = 11$ agents on each side. The figure shows the mean squared error of each estimator as a function of time elapsed in the simulation. Because CD is inherently stochastic, the corresponding curve is jagged. What converges in Proposition 2 is the running average of the estimates that generate this curve.

Figure 4 drops ML and focuses on the performance of CD in a larger market, with $n_1 = n_2 = 15$ agents on each side. CD continues to converge, within a "reasonable" time frame and with "few" observations. For the remainder, let us focus on the number of observations. What is remarkable about this market is that, while it has $2^{n_1+n_2} \approx 1,000,000,000$ possible participation profiles, only $T = 500,000$ (possibly non-distinct) observations suffice to estimate the model parameters. We say that $T = 500,000$ is "few" because orders of magnitude more observations would have been

required to deal with the pesky denominator in (3) by using the differencing approach commonly deployed in estimation of (smaller) discrete-choice models.

To describe the differencing approach, let $\mathbf{x}$ and $\mathbf{z}$ be two participation profiles such that $\mathbf{z} = \mathbf{0}$ and $\mathbf{x}$ is zero everywhere except at $x_{1i} = 1$. Set $\sigma = 1$. Taking the difference of the logarithms of the probabilities of these profiles at the sBRE eliminates the pesky[1] denominator in (3) to yield:

$$\log \mathbb{P}(\mathbf{x} \mid \mathbf{b}, \mathbf{W}) - \log \mathbb{P}(\mathbf{z} \mid \mathbf{b}, \mathbf{W}) = \mathbf{b}^\top (\mathbf{x} - \mathbf{z}) + \mathbf{x}_1^\top \mathbf{W} \mathbf{x}_2 - \mathbf{z}_1^\top \mathbf{W} \mathbf{z}_2,$$

whence

$$b_{1i} = \log \frac{\mathbb{P}(\mathbf{x} \mid \mathbf{b}, \mathbf{W})}{\mathbb{P}(\mathbf{z} \mid \mathbf{b}, \mathbf{W})}.$$

As a result, by inspection of the display above, b_{1i} can be estimated as the logarithm of the ratio of the observed frequencies of $\mathbf{x}$ and $\mathbf{z}$. To estimate b_{1i} like this with any degree of precision, the profiles $\mathbf{x}$ and $\mathbf{z}$—just two profiles out of the $2^{n_1+n_2}$ possible ones—must be observed over and over again, thereby requiring a data set of a size $T \gg 2^{n_1+n_2}$. This requirement contrasts with the simulation above, where $T \ll 2^{n_1+n_2}$ observations suffice because CD, just like ML, does not toss away any observations: every observation contains information that is directly or indirectly relevant for estimating b_{1i} (as well as other parameters).

In fact, CD is almost as parsimonious with the data as is ML, which converges at the rate $\sqrt{T}$. Remark 2, which makes the comparison precise, follows by applying the recent result of Glaser, Huang, and Gretton (2024, Theorem 4.2) (Appendix A.5).

Remark 2. In our setting, the CD estimator converges in probability to the true parameters $(\mathbf{W}, \mathbf{b})$ at a rate of at least $\sqrt{T/\log T}$.

6 Pricing

This section discusses optimal pricing in the special case in which σ, the scale-of-noise parameter, is "small." This case is also the focus of the existing literature on platform pricing. Our message is reassuring: estimation has not been in vain. The parameter estimates can be put to use in service of welfare or revenue.

Because the benefits derived from interactions on the platform are specific to each individual, we assume that prices, too, can be individualized. The plausibility and utility of such prices in the context of two-sided markets are argued by de Cornière, Mantovani, and Shekhar (2025).

6.1 A Class of Pricing Schemes

We consider the most general pricing scheme that preserves the strategic structure of the participation game. This pricing scheme has two components. The first component is a matrix $\mathbf{P} \equiv (P_{ij})_{i,j} \in \mathbb{R}^{n_1 \times n_2}$ of interaction prices. When agent i on side 1 and agent j on side 2 both

[1] We thank Marcin Pęski for suggesting that we make this comparison.

join the platform, each pays the amount P_{ij}. The second component is a vector $\mathbf{p} \equiv (\mathbf{p}_1, \mathbf{p}_2) \in \mathbb{R}^n$ of personalized access prices, where $\mathbf{p}_1$ and $\mathbf{p}_2$ are the price vectors faced by the agents on sides 1 and 2 of the market, respectively. As a result, at a participation profile $\mathbf{x}$, a participating agent on side 1 pays the corresponding entry of $\mathbf{p}_1 + \mathbf{P}\mathbf{x}_2$, and a participating agent on side 2 pays the corresponding entry of $\mathbf{p}_2 + \mathbf{P}^\top \mathbf{x}_1$. The equilibrium probability of a participation profile $\mathbf{x}$ when prices are $(\mathbf{P}, \mathbf{p})$ is $\mathbb{P}(\mathbf{x} \mid \mathbf{W} - \mathbf{P}, \mathbf{b} - \mathbf{p})$, which is the obvious adjustment of the corresponding probability (3).

6.2 Pricing for Welfare

We lead with welfare maximization, and we do so for two reasons: conceptual clarity and realism. Conceptually, revenue maximization in our setting requires the platform to maximize welfare before extracting it. Realistically, early in its existence, a platform wants to maximize welfare in order to attract new customers and achieve market dominance, all the while passing under the radar of antitrust authorities, who have traditionally been blind to welfare-maximizing pricing (Bork, 1978).

The Problem

The welfare maximization problem takes the form

$$\max_{\mathbf{P},\mathbf{p}} \sum_{\mathbf{x}\in\{0,1\}^n} \mathbb{P}(\mathbf{x} \mid \mathbf{W} - \mathbf{P}, \mathbf{b} - \mathbf{p}) \left(\mathbf{b}^\top \mathbf{x} + 2\mathbf{x}_1^\top \mathbf{W}\mathbf{x}_2\right), \tag{8}$$

where $\mathbf{b}^\top \mathbf{x} + 2\mathbf{x}_1^\top \mathbf{W}\mathbf{x}_2$ is utilitarian welfare, the sum of the platform's revenue and the agents' payoffs. Utilitarian welfare counts bilateral benefits from interaction twice because they accrue twice: once to each agent in the interacting pair. The welfare does not contain prices, which are a wash between the platform and the agents. Nor does the welfare contain any epsilons that appear in the definition of stochastic best response in (2); our interpretation is that these epsilons encode mistakes rather than utility shocks.

The problem in (8) suffers from combinatorial explosion because of 2^n terms in the sum and 2^n terms inside each probability $\mathbb{P}(\mathbf{x} \mid \mathbf{W} - \mathbf{P}, \mathbf{b} - \mathbf{p})$. Here, however, CD is of no avail. And yet, we plow ahead.

The Price of Anarchy

Each agent's participation decision imposes an externality on agents on the other side of the market. Because no agent internalizes this externality, equilibrium is generally inefficient in the sense of falling short of maximizing the **first-best welfare**, defined as

$$\max_{\mathbf{x}\in\{0,1\}^n} \left\{\mathbf{b}^\top \mathbf{x} + 2\mathbf{x}_1^\top \mathbf{W}\mathbf{x}_2\right\}.$$

Proposition 3 (proved in Appendix A.3) shows that equilibrium welfare can be an arbitrarily small fraction of the first-best welfare: in other words, the price of anarchy is infinite. As a result, there is a problem for pricing to solve, or at least to mitigate.

Proposition 3 (The Price of Anarchy)**.** *With no corrective pricing (i.e., under* $\mathbf{P} = \mathbf{0}$ *and* $\mathbf{p} = \mathbf{0}$*), equilibrium welfare can be an arbitrarily small fraction of the first-best welfare.*

The idea behind Proposition 3 is to conceive of economies in which network effects are strong enough for welfare to be maximized when all participate but not strong enough for anyone to actually want to participate. This discrepancy between welfare-maximizing behavior and equilibrium behavior arises because participants do not internalize the externalities.

Welfare-Maximizing Prices when $\sigma \to 0$

In the special case in which $\sigma \to 0$ (the mistakes in best responses vanish), the welfare-maximization problem in (8) admits a simple solution. This special case is interesting because the conclusion of Proposition 3 continues to hold even when $\sigma \to 0$, as is evident in the proposition's proof.

Proposition 4 (Pricing for First-Best Welfare)**.** *Fix any* $\gamma > 0$*. Suppose that* $\mathbf{P} = (1-\gamma)\,\mathbf{W}$ *and* $\mathbf{p} = \left(1 - \frac{1}{2}\gamma\right)\mathbf{b}$*. Then, as* $\sigma \to 0$*, equilibrium welfare converges to the first-best welfare.*

Proof. Under the prices in the proposition's statement, expected welfare satisfies

$$\lim_{\sigma \to 0} \frac{\sum_{\mathbf{x} \in \{0,1\}^n} \exp\left(\frac{\gamma}{2\sigma}\left[\mathbf{b}^\top \mathbf{x} + 2\mathbf{x}^\top{}_1 \mathbf{W} \mathbf{x}_2\right]\right)\left[\mathbf{b}^\top \mathbf{x} + 2\mathbf{x}_1^\top \mathbf{W} \mathbf{x}_2\right]}{\sum_{\mathbf{x} \in \{0,1\}^n} \exp\left(\frac{\gamma}{2\sigma}\left[\mathbf{b}^\top \mathbf{x} + 2\mathbf{x}_1^\top \mathbf{W} \mathbf{x}_2\right]\right)} = \max_{\mathbf{x} \in \{0,1\}^n} \left\{\mathbf{b}^\top \mathbf{x} + 2\mathbf{x}_1^\top \mathbf{W} \mathbf{x}_2\right\},$$

as desired. ■

When $\gamma = 1$, Proposition 4 prescribes $\mathbf{P} = \mathbf{0}$ and $\mathbf{p} = \frac{1}{2}\mathbf{b}$, thereby cutting each agent's standalone value from joining in half while leaving the experienced interaction benefits unchanged. As a result, the relative salience of interaction benefits doubles. This doubling induces each agent to internalize the externality from his participation. An alternative way to achieve the same goal is to set $\gamma = 2$ and, so, charge $\mathbf{P} = -\mathbf{W}$ (Pigouvian interaction prices) and $\mathbf{p} = \mathbf{0}$ (no access fees). Which value of γ is appropriate depends on whether revenue is a desideratum or a liability, and on whether nonzero interaction prices are practical.

6.3 Pricing for Revenue

Profit—here, synonymous with revenue—is the platform's ultimate reward.

The Problem

The expected-revenue maximization problem takes the form

$$\max_{\mathbf{P}, \mathbf{p}} \sum_{\mathbf{x} \in \{0,1\}^n} \mathbb{P}\left(\mathbf{x} \mid \mathbf{W} - \mathbf{P}, \mathbf{b} - \mathbf{p}\right)\left(\mathbf{p}^\top \mathbf{x} + 2\mathbf{x}_1^\top \mathbf{P} \mathbf{x}_2\right), \tag{9}$$

where $\mathbf{p}^\top \mathbf{x} + 2\mathbf{x}_1^\top \mathbf{P}\mathbf{x}_2$ is the platform's revenue. The revenue is the sum of the gathered access fees $\mathbf{p}^\top \mathbf{x}$ and the gathered interaction fees $2\mathbf{x}_1^\top \mathbf{P}\mathbf{x}_2$. Each interaction fee must be counted twice because it is collected twice, once from each member of the participating pair. From the perspective of revenue, it does not matter whether the epsilons in the best-response equation (2) are interpreted as mistakes (our preferred interpretation) or as shocks to payoffs.

The Price of Model Misspecification

Just like refraining from corrective pricing may come at a great cost to welfare (Proposition 3), the failure to account for heterogeneous network effects in learning and pricing may come at a great cost to revenue (Proposition 5 below, proved in Appendix A.4). Should the platform ignore the heterogeneity of network effects, it will make wrong inferences about the underlying model parameters and, as a result, will price incorrectly. In fact, pricing based on the misspecified model may be arbitrarily ruinous: the revenue may be arbitrarily low relative to the revenue the platform would have achieved had it used the correct model.

Proposition 5 (The Price of Model Misspecification)**.** *Even when all interaction effects are non-negative, the maximal realized revenue under the misspecified model that assumes homogeneous network effects can be an arbitrarily small fraction of the maximal realized revenue under the correctly specified model with heterogeneous network effects.*

In order to see the intuition for the arbitrary loss from model misspecification in Proposition 5, we consider an example; the proposition's proof follows the logic of this example. Suppose that Alice and Bob enjoy a positive benefit from interacting with each other, whereas Carol and Dave gain nothing. In addition, Alice and Bob each incur standalone costs from participating, whereas Carol and Dave incur none. A platform that acknowledges heterogeneous interaction effects would identify and extract some of the surplus generated by Alice and Bob's interactions. A misspecified platform, by contrast, would underestimate the strength of Alice's and Bob's bilateral benefit, which, in the correlation between men's (i.e., Bob's and Dave's) and women's (i.e., Alice's and Carol's) aggregate participation, will be diluted by Carol's and Dave's utter indifference to participants on the other side of the market. The underestimation can be so severe that the revenue lost by the misspecified platform can be arbitrarily large.

Proposition 5 is the raison d'êtr̀e for this paper.

Revenue-Maximizing Prices when $\sigma \to 0$

In the special case in which $\sigma \to 0$, the revenue-maximization problem in (9) admits a simple solution. This special case is still interesting because the conclusion of Proposition 5 continues to hold when $\sigma \to 0$.

Proposition 6 (Pricing for First-Best Revenue Extraction)**.** *Suppose that* $\mathbf{P} = (1 - \sqrt{\sigma})\,\mathbf{W}$ *and* $\mathbf{p} = \left(1 - \frac{1}{2}\sqrt{\sigma}\right)\mathbf{b}$*. Then, as* $\sigma \to 0$*, equilibrium revenue converges to the first-best welfare, which is the limiting upper bound on attainable revenue.*

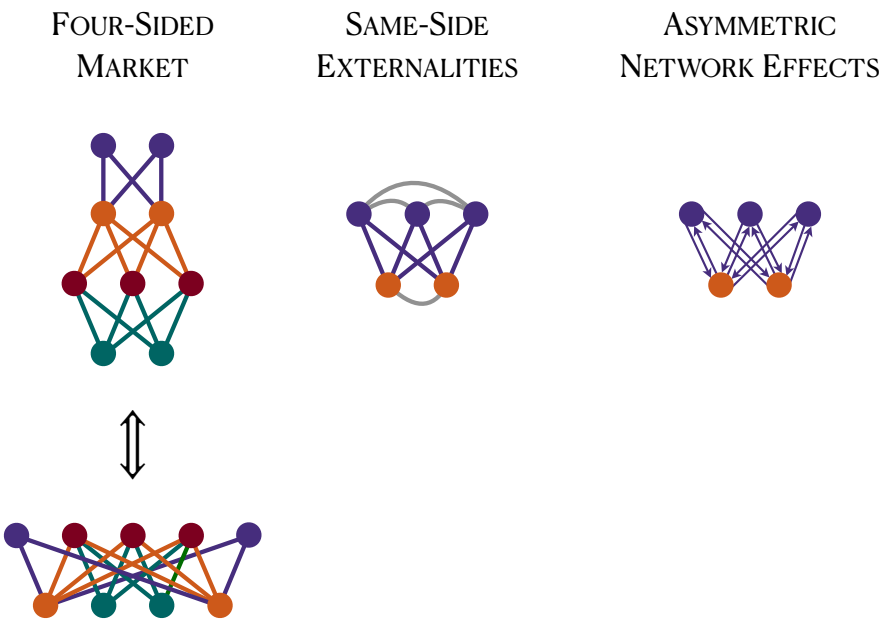

Figure 5: **Three extensions.** Multi-sided markets (left), same-side externalities (center), and asymmetric network effects.

Proof. The proposition is a special case of Proposition 7 (introduced shortly, in Section 7.3). ■

Because the prices in Proposition 6 asymptotically collect the first-best welfare in revenue, they must asymptotically generate the first-best welfare. Indeed, the prices correspond to the welfare-maximizing prices of Proposition 4 with $\gamma = \sqrt{\sigma}$.

Full extraction in Proposition 6 stands in contrast to divide-and-conquer pricing, which has been shown to emerge as optimal in related settings (Chan, 2021, 2025). Divide-and-conquer pricing "divides" by splitting agents into two groups, charging members of one group nothing, and "conquering" the other group by extracting the entire surplus from its agents. (A typical example is access to a nightclub: women get in free, while men pay an arm and a leg.) As Chan (2025, Footnote 3) notes, the optimality of divide-and-conquer relies on the commitment to forgo interaction prices, which otherwise open the door to full extraction. Full extraction in our setting does not rely on the hitherto maintained symmetry whereby the value from each interaction is assumed to be split equally between the interacting agents; we defer the formal statement and the proof of this assertion to Section 7.3 (Proposition 7).

7 Extensions

The paper's findings reach beyond two-sided markets. Figure 5 announces three extensions, each of which we consider below in isolation. The first one is the multi-sided market that formally is but a reinterpretation of the baseline model. The second one is the two-sided market that admits

same-side externalities and thereby ceases being two-sided at all: any agent may be connected to any other agent. The third extension admits asymmetric network effects: connected individuals may enjoy their interaction differently.

7.1 Multi-Sided Markets

Our two-sided markets framework accommodates without modification a class of multi-sided markets, which are markets whose participants can be partitioned into more than two economically distinct groups. To illustrate, consider the following instance of a four-sided market: a streaming platform. Its four sides are advertisers, customers, artists, and recording labels. Agents on the same side do not interact. Across sides, advertisers (side 1) interact with customers (side 2), who interact with artists (side 3), who interact with labels (side 4). In other words, the four sides are **stacked**: only the groups with consecutive indices interact. Let $\mathbf{W}_{12}$, $\mathbf{W}_{32}$, and $\mathbf{W}_{34}$ denote the matrices of interaction effects between groups 1 and 2, 3 and 2, and 3 and 4, respectively. Let $\mathbf{b}$ denote the vector of standalone values from participation, as before. The participation game can be verified to be a potential game with the potential

$$\mathbf{b}^\top \mathbf{x} + \mathbf{x}_1^\top \mathbf{W}_{12} \mathbf{x}_2 + \mathbf{x}_3^\top \mathbf{W}_{32} \mathbf{x}_2 + \mathbf{x}_3^\top \mathbf{W}_{34} \mathbf{x}_4 = \mathbf{b}^\top \mathbf{x} + \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_3 \end{pmatrix}^\top \begin{pmatrix} \mathbf{W}_{12} & \mathbf{0} \\ \mathbf{W}_{32} & \mathbf{W}_{34} \end{pmatrix} \begin{pmatrix} \mathbf{x}_2 \\ \mathbf{x}_4 \end{pmatrix}.$$

The right-hand side of the equation in the display above reveals that the four-sided market is formally equivalent to the two-sided market that treats sides 1 and 3 as one side and sides 2 and 4 as the other. The bidiagonal block matrix collects network effects and corresponds to $\mathbf{W}$ in the two-sided market model. The described reduction to two-sided markets holds for any number of sides in any stacked multi-sided market. This reduction is only possible because the underlying two-sided market admits heterogeneous network effects.

The reduction implies that all the paper's findings extend to stacked multi-sided markets. In particular, CD is fast and consistent. The proposed pricing schemes apply.

7.2 Same-Side Externalities and General Binary-Action Games

Same-side externalities arise when agents care about the participation decisions of agents on both sides of the market, including the same side as their own. Same-side externalities open up a whole new world of applications. An online recruiting marketplace with same-side congestion is one: jobseekers exert a congestion externality on other jobseekers, and each recruiter's vacancy posting exerts a congestion externality on other recruiters' postings (Example 1 below). Because there is nothing two-sided about two-sided markets with same-side externalities—anyone can exert an externality on anyone else—another application is the management of a corporate R&D portfolio: each development team's payoff from a binary decision whether to undertake its R&D project may depend on every other team's decision through competition and knowledge spillovers (Example 2

below).

The addition of same-side externalities begets a **binary-action game with pair-specific network effects**. This game is determined by a matrix $\Omega \in \mathbb{R}^{n\times n}$. Each agent i's standalone value from participation, formerly denoted by b_{1i} or b_{2j}, is now denoted by $\frac{1}{2}\Omega_{ii}$. The network effects no longer possess the bipartite structure but, for now, remain symmetric: $\Omega_{ij} = \Omega_{ji}$ for all i and all j with $j \neq i$. Letting $\mathbf{x} \in \{0,1\}^n$ denote a participation profile, agent i's payoff is $x_i\left(\frac{1}{2}\Omega_{ii} + \sum_{j\neq i}\Omega_{ij}x_j\right)$. The potential in the associated binary-action game can be verified to be

$$\Phi_\Omega(\mathbf{x}) \equiv \frac{1}{2}\mathbf{x}^\top\Omega\mathbf{x}. \tag{10}$$

Here are two examples:

Example 1 (Online Recruiting Marketplace with Same-Side Congestion)**.** The two sides of the market comprise n_1 jobseekers and n_2 recruiters. At a participation profile $\mathbf{x}$, the number of logged-on jobseekers is denoted by $U = \sum_{i=1}^{n_1} x_{1i}$, and the number of recruiters who maintain an active vacancy (at most one per recruiter) is denoted by $V = \sum_{i=1}^{n_2} x_{2i}$. Each participating jobseeker's payoff is $V - (U-1)\,u - u$ for some positive same-side congestion parameter u, which here plays a dual role by also capturing the standalone cost of job search on the platform. Each participating recruiter's payoff is $U - (V-1)\,w - w$ for some positive same-side congestion parameter w, which, here, is also the standalone cost of maintaining a vacancy. The total welfare at a participation profile $\mathbf{x}$ with (U,V) is

$$m(U,V) = 2UV - U^2u - V^2w.$$

Example 1 is a transparent two-parameter special case. In an actual recruiting marketplace, cross-side network effects and same-side congestion effects would generally be pair-specific. The unrestricted binary-action game readily accommodates both.

Example 2 (Corporate R&D Portfolio)**.** In each of a sequence of frequent project-activation rounds, each of a fixed set of n semi-autonomous R&D teams in a large technology company decides whether to undertake or continue its project. Should teams i and j both invest, each will see its payoff change by an amount Ω_{ij}. If $\Omega_{ij} > 0$, the R&D efforts of i and j are complements (as in Goyal and Moraga-González's, 2001, model of R&D); if $\Omega_{ij} < 0$, their efforts are substitutes. Team i's cost of R&D is $-\frac{1}{2}\Omega_{ii}$ (typically a positive number). The game is a special case of the linear-quadratic game of Ballester, Calvó-Armengol, and Zenou (2006) when each action x_i is binary and Ω is symmetric.

Both an operator of an online recruiting marketplace and a manager of a corporate R&D portfolio can put estimates of Ω to immediate practical use. The former would use them to set pair-specific interaction fees and subsidies depending on whether a pair creates value or imposes congestion. The latter would use them to design team-specific budgets and pair-specific incentives while accounting for how inducing one team to invest changes the returns to the other teams' projects.

The equilibrium characterization in Proposition 1 continues to hold for general binary-action games, provided in the definition of the sBRE (Definition 1) the revision protocol is assumed to be random scanning, and provided the potential in (3) is replaced by (10). Instead of corresponding to a restricted Boltzmann machine, sBRE now corresponds to a fully visible, unrestricted **Boltzmann machine** (BM) whose energy function is the negative of the potential (10).

CD remains consistent, although it is now slower because random scanning must replace blocked sampling. Alternatives to CD exist. Hyvärinen (2006) advocates pseudolikelihood maximization, which predates and sometimes computationally outperforms CD when estimating BMs. Other rivals of CD for BM include ratio matching (Hyvärinen, 2007), minimum probability flow (Sohl-Dickstein, Battaglino, and DeWeese, 2011), and noise-contrastive estimation (Gutmann and Hyvärinen, 2012).

The welfare- and revenue-maximizing pricing schemes in Propositions 4 and 6, respectively, extend in obvious ways to binary-action games with pair-specific network effects. Indeed, let $\mathbf{P} \in \mathbb{R}^{n \times n}$ be a symmetric price matrix, whose interpretation parallels that of Ω: $\frac{1}{2}P_{ii}$ is agent i's payment for participation regardless of whether others participate, and each P_{ij} with $i \neq j$ is the price agents i and j each pay if both show up. The induced game with prices has the potential $\frac{1}{2}\mathbf{x}^\top (\Omega - \mathbf{P})\,\mathbf{x}$. Consider the family of prices, parameterized by a positive γ, that for all i and all $j \neq i$, satisfy $P_{ii} = (1 - \frac{\gamma}{2})\,\Omega_{ii}$ and $P_{ij} = (1 - \gamma)\,\Omega_{ij}$. Then, $\Omega_{ii} - P_{ii} = \frac{\gamma}{2}\Omega_{ii}$ and $\Omega_{ij} - P_{ij} = \gamma\Omega_{ij}$, leading to the potential

$$\frac{1}{2}\mathbf{x}^\top (\Omega - \mathbf{P})\,\mathbf{x} = \frac{\gamma}{2}\left(\sum_{i=1}^{n}\frac{1}{2}\Omega_{ii}x_i + 2\sum_{i<j}\Omega_{ij}x_i x_j\right),$$

which is $\gamma/2$ times the total surplus. When $\sigma \to 0$, sBRE concentrates on the potential-maximizing participation profile, which under the proposed pricing scheme is also the welfare-maximizing participation profile, as in Proposition 4. Moreover, when $\gamma = \sqrt{\sigma}$, the corresponding revenue,

$$\sum_i \frac{1}{2}P_{ii}x_i + 2\sum_{i<j}P_{ij}x_i x_j,$$

converges to the total surplus as $\sigma \to 0$, just as it does in Proposition 6.

The setting of the binary-action contracting model of Chan (2025, Section 4.3) is a special case of Example 2 above as long as others' actions enter agent payoffs linearly. If others' actions aggregate in each agent's payoff nonlinearly—something that Chan allows for but we don't—sBRE behavior, while still dictated by the game's potential, no longer corresponds to a BM. One then must look beyond BM training algorithms for model estimation techniques.

7.3 Asymmetric Network Effects

So far, we have assumed symmetric network effects. In the notation of the binary-action game of Section 7.2, symmetry requires that $\Omega_{ij} = \Omega_{ji}$ for every agent pair (i, j). This assumption is restrictive and can be relaxed.

To this end, augment the binary-action game Ω by assuming that each agent i's benefit from interacting with any agent j is $\lambda_i \Omega_{ij}$ for some known positive parameter λ_i. One can interpret each λ_i in $\lambda \equiv (\lambda_1, \ldots, \lambda_n)$ as agent i's bargaining power. The agent's payoff in the **augmented binary-action game** (Ω, λ) then becomes $x_i \left(\frac{1}{2}\Omega_{ii} + \lambda_i \sum_{j \neq i} \Omega_{ij} x_j \right)$. This game nests the binary-action version of Chan's (2025) network game; conversely, with a rank restriction imposed on Ω, it is nested within Chan's game. The augmented binary-action game admits a potential:

$$\Phi_{\Omega,\lambda}(\mathbf{x}) \equiv \sum_i \frac{\frac{1}{2}\Omega_{ii}}{\lambda_i} x_i + \sum_{i<j} x_i \Omega_{ij} x_j. \tag{11}$$

If we further assume that the noise that an agent experiences when best-responding is proportional to λ_i, then the associated sBRE under random scanning obeys the conclusion of Proposition 1 with the obvious modification: the equilibrium probability of a participation profile $\mathbf{x}$ becomes $\mathbb{P}(\mathbf{x} \mid \Omega, \lambda) = e^{\Phi_{\Omega,\lambda}(\mathbf{x})/\sigma} / \sum_{\mathbf{y}} e^{\Phi_{\Omega,\lambda}(\mathbf{y})/\sigma}$. Once again, sBRE is a BM. The underlying parameter Ω can be estimated as in Section 7.2 with the obvious adjustment for λ, assumed to be known. As a result, whenever Chan's (2025) network game is a special case of ours, its parameters can be estimated using the techniques discussed in this paper.

The full-extraction pricing result of Proposition 6 extends to the asymmetric setting of augmented binary-action games.

Proposition 7 (Pricing for First-Best Revenue Extraction in the Augmented Binary-Action Game)**.** *In the augmented binary-action game (Ω, λ), suppose that each agent i pays an access price $p_i = \frac{1}{2}\left(1 - \lambda_i \frac{\sqrt{\sigma}}{2}\right)\Omega_{ii}$ and, for each partner j, pays a pair-specific interaction price $P_{ij}^{(i)} = \lambda_i \left(1 - \frac{\sqrt{\sigma}}{2}(\lambda_i + \lambda_j)\right)\Omega_{ij}$. Then, as $\sigma \to 0$, the equilibrium revenue converges to the first-best welfare, which is the limiting upper bound on attainable revenue.*

Proof. For an arbitrary participation profile $\mathbf{x}$, the total surplus is

$$W(\mathbf{x}) \equiv \sum_{i=1}^{n} \frac{1}{2}\Omega_{ii} x_i + \sum_{i<j} (\lambda_i + \lambda_j) \Omega_{ij} x_i x_j.$$

Fix the prices $\left(\left(P_{ij}^{(i)} \right)_{j \neq i}, p_i \right)_{i=1}^{n}$ at the values specified in the proposition's statement. Then, the game's potential

$$\Phi(\mathbf{x}) = \sum_{i=1}^{n} \frac{\frac{1}{2}\Omega_{ii} - p_i}{\lambda_i} x_i + \sum_{i<j} \frac{\lambda_i \Omega_{ij} - P_{ij}^{(i)}}{\lambda_i} x_i x_j$$

—which is the potential in (11) augmented by prices—can be rewritten as $\frac{\sqrt{\sigma}}{2} W(\mathbf{x})$, and the revenue

$$\sum_{i=1}^{n} p_i x_i + \sum_{i<j} \left(P_{ij}^{(i)} + P_{ji}^{(j)} \right) x_i x_j$$

can be rewritten as

$$W(\mathbf{x}) - \frac{\sqrt{\sigma}}{2}\left(\sum_{i=1}^{n}\frac{\lambda_i}{2}\Omega_{ii}x_i + \sum_{i<j}(\lambda_i+\lambda_j)^2\,\Omega_{ij}x_ix_j\right)$$

and, hence, converges uniformly to $W(\mathbf{x})$ as $\sigma \to 0$. Therefore, at the specified prices, the expected revenue has the same limit as the expected total surplus:

$$\lim_{\sigma\to 0}\frac{\sum_{\mathbf{x}\in\{0,1\}^n}\exp\left(\frac{1}{2\sqrt{\sigma}}W(\mathbf{x})\right)W(\mathbf{x})}{\sum_{\mathbf{y}\in\{0,1\}^n}\exp\left(\frac{1}{2\sqrt{\sigma}}W(\mathbf{y})\right)} = \max_{\mathbf{x}\in\{0,1\}^n}W(\mathbf{x}),$$

which is the first-best welfare. As $\sigma \to 0$, a uniform lower bound on each agent's expected payoff converges to zero. Therefore, the upper bound on the platform's revenue converges to the first-best welfare. ■

A Appendix: Omitted Proofs

A.1 Proof of Proposition 1

The sequence of participation profiles generated by the stochastic best-response dynamics in sBRE is a finite Markov chain. This chain is irreducible (every state can be reached from every state in finitely many steps if the shocks take particular values) and aperiodic (the system does not cycle deterministically over participation profiles). Under irreducibility and aperiodicity, the ergodic theorem assures existence of a unique stationary distribution that is reached from any initial participation profile. This distribution's functional form, in (3), is the concern for the rest of the proof.

The proof strategy is to verify that the distribution $\mathbb{P}$ in (3)—where we have suppressed the dependence of $\mathbb{P}$ on $(\mathbf{W}, \mathbf{b})$—is stationary by verifying a sufficient condition. In order to formulate this condition, let $\mathbb{T}(\mathbf{x}, \mathbf{z})$ denote the probability that a single step of best-response dynamics transforms a participation profile $\mathbf{x}$ into $\mathbf{z}$. The **detailed balance condition** (Brémaud, 2020, Corollary 2.4.15, p. 92) verifies that $\mathbb{P}$ is the stationary distribution if, for all $\mathbf{x}$ and all $\mathbf{z}$, we have

$$\mathbb{P}(\mathbf{x})\,\mathbb{T}(\mathbf{x},\mathbf{z}) = \mathbb{P}(\mathbf{z})\,\mathbb{T}(\mathbf{z},\mathbf{x}). \tag{A.1}$$

The condition says that $\mathbb{P}$ is a stationary distribution if the "outflow" of probability from every state $\mathbf{x}$ to every state $\mathbf{z}$ reachable from $\mathbf{x}$ in one step of best-response dynamics equals the "inflow" of probability from $\mathbf{z}$ to $\mathbf{x}$; in other words, the two flows are exactly "balanced."

The critical feature of any revision protocol in sBRE is that, at each step, actions on at most one side of the market are revised. Henceforth, we focus on the step at which actions on side 1 are (possibly) revised while actions on side 2 (definitely) remain fixed: $\mathbf{z}_2 = \mathbf{x}_2$. (The complementary case of side-2 revisions is analogous.)

Let $\mathbf{u} \in \{0,1\}^{n_1}$ represent the side-1 agents who are called upon to revise their actions at the best-response step under consideration, with $u_i = 1$ corresponding to agent i being called upon, and $u_i = 0$ corresponding to him not being called upon. Assume that $\mathbf{u}$ is distributed according to some probability distribution μ that is the same at every step of best-response dynamics. Note that

$$\mu\left(\prod_{i:x_{1i}\neq z_{1i}} u_i = 1\right) = 0 \quad \Longrightarrow \quad \mathbb{T}(\mathbf{x},\mathbf{z}) = \mathbb{T}(\mathbf{z},\mathbf{x}) = 0 \quad \Longrightarrow \quad \text{(A.1) holds.}$$

That is, whenever at least one agent whose actions in $\mathbf{x}$ and $\mathbf{z}$ differ is given no chance to revise his actions, there is no way to transition from $\mathbf{x}$ to $\mathbf{z}$ or from $\mathbf{z}$ to $\mathbf{x}$ in one step of the revision process. As a result, in the remainder of the proof, we focus on the complementary case in which $\mathbf{x}$ and $\mathbf{z}$ are such that

$$\mu\left(\prod_{i:x_{1i}\neq z_{1i}} u_i = 1\right) > 0. \tag{A.2}$$

Let $\pi_i(z_{1i} \mid \mathbf{x}_2)$ denote the probability that the agent i who is called upon to revise his action by best-responding to the previous-step action profile $\mathbf{x}$ opts for an action $z_{1i} \in \{0,1\}$. This probability depends on $\mathbf{x}$ only through $\mathbf{x}_2$ thanks to the bipartite structure of the two-sided market. Because ε_{1i} is distributed logistically, $\pi_i(z_{1i} \mid \mathbf{x}_2)$ satisfies

$$\pi_i(z_{1i} \mid \mathbf{x}_2) = \frac{e^{z_{1i}\left(b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2\right)}}{1 + e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}}, \tag{A.3}$$

where we have innocuously set $\sigma = 1$ to carry less notation; $\sigma = 1$ stays for the rest of the proof. Because all $(\varepsilon_{1i})_{i=1}^{n_1}$ are independent, the transition probability $\mathbb{T}$ takes the form

$$\mathbb{T}(\mathbf{x},\mathbf{z}) = \sum_{\mathbf{u}\in\{0,1\}^{n_1}} \mu(\mathbf{u}) \prod_{i=1}^{n_1} \left(u_i \pi_i(z_{1i} \mid \mathbf{x}_2) + (1-u_i)\,\mathbf{1}_{\{z_{1i}=x_{1i}\}}\right). \tag{A.4}$$

Using the definitions of $\mathbb{T}$ in (A.4) and $\mathbb{P}$ in (3) and $\mathbf{z}_2 = \mathbf{x}_2$, the detailed balance condition in (A.1) can be written as

$$\begin{aligned}
&\frac{e^{\mathbf{b}_1^\top \mathbf{x}_1 + \mathbf{b}_2^\top \mathbf{x}_2 + \mathbf{x}_1^\top \mathbf{W} \mathbf{x}_2}}{\sum_{\mathbf{y}} e^{\mathbf{b}^\top \mathbf{y} + \mathbf{y}_1^\top \mathbf{W} \mathbf{y}_2}} \sum_{\mathbf{u}} \mu(\mathbf{u}) \prod_{i=1}^{n_1} \left(u_i \pi_i(z_{1i} \mid \mathbf{x}_2) + (1-u_i)\,\mathbf{1}_{\{z_{1i}=x_{1i}\}}\right) \\
&\qquad = \frac{e^{\mathbf{b}_1^\top \mathbf{z}_1 + \mathbf{b}_2^\top \mathbf{x}_2 + \mathbf{z}_1^\top \mathbf{W} \mathbf{x}_2}}{\sum_{\mathbf{y}} e^{\mathbf{b}^\top \mathbf{y} + \mathbf{y}_1^\top \mathbf{W} \mathbf{y}_2}} \sum_{\mathbf{u}} \mu(\mathbf{u}) \prod_{i=1}^{n_1} \left(u_i \pi_i(x_{1i} \mid \mathbf{x}_2) + (1-u_i)\,\mathbf{1}_{\{z_{1i}=x_{1i}\}}\right).
\end{aligned}$$

Cancelling the normalizing constant in the denominator and simplifying gives

$$\frac{\sum_{\mathbf{u}} \mu(\mathbf{u}) \prod_{i=1}^{n_1} \left(u_i \pi_i(x_{1i} \mid \mathbf{x}_2) + (1-u_i)\,\mathbf{1}_{\{z_{1i}=x_{1i}\}}\right)}{\sum_{\mathbf{u}} \mu(\mathbf{u}) \prod_{i=1}^{n_1} \left(u_i \pi_i(z_{1i} \mid \mathbf{x}_2) + (1-u_i)\,\mathbf{1}_{\{z_{1i}=x_{1i}\}}\right)} = e^{\mathbf{b}_1^\top (\mathbf{x}_1 - \mathbf{z}_1) + (\mathbf{x}_1 - \mathbf{z}_1)^\top \mathbf{W} \mathbf{x}_2}. \tag{A.5}$$

For some partition $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ of the index set $\{1, 2, \ldots, n_1\}$, we have $x_{1i} = 1$ and $z_{1i} = 0$ for all

$i \in \mathcal{A}$, $x_{1i} = 0$ and $z_{1i} = 1$ for all $i \in \mathcal{B}$, and $x_{1i} = z_{1i}$ for all $i \in \mathcal{C}$. With this notation, the left-hand side of (A.5) becomes

$$\frac{\sum_{\mathbf{u}} \mu(\mathbf{u}) U(\mathbf{u}) \prod_{i \in \mathcal{A}} u_i \pi_i(1 \mid \mathbf{x}_2) \prod_{i \in \mathcal{B}} u_i \pi_i(0 \mid \mathbf{x}_2)}{\sum_{\mathbf{u}} \mu(\mathbf{u}) U(\mathbf{u}) \prod_{i \in \mathcal{A}} u_i \pi_i(0 \mid \mathbf{x}_2) \prod_{i \in \mathcal{B}} u_i \pi_i(1 \mid \mathbf{x}_2)},$$

where $U(\mathbf{u}) \equiv \prod_{i \in \mathcal{C}} (u_i \pi_i(x_{1i} \mid \mathbf{x}_2) + 1 - u_i)$. Substituting the definition of π_i from (A.3) into the display above gives

$$\begin{aligned}
&\frac{\sum_{\mathbf{u}} \mu(\mathbf{u}) U(\mathbf{u}) \prod_{i \in \mathcal{A}} \frac{u_i e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}}{1 + e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}} \prod_{i \in \mathcal{B}} \frac{u_i}{1 + e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}}}{\sum_{\mathbf{u}} \mu(\mathbf{u}) U(\mathbf{u}) \prod_{i \in \mathcal{A}} \frac{u_i}{1 + e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}} \prod_{i \in \mathcal{B}} \frac{u_i e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}}{1 + e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}}} \\
&\quad = \frac{\sum_{\mathbf{u}} \mu(\mathbf{u}) U(\mathbf{u}) \prod_{i \in \mathcal{A} \cup \mathcal{B}} \frac{u_i}{1 + e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}} \prod_{i \in \mathcal{A}} e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}}{\sum_{\mathbf{u}} \mu(\mathbf{u}) U(\mathbf{u}) \prod_{i \in \mathcal{A} \cup \mathcal{B}} \frac{u_i}{1 + e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}} \prod_{i \in \mathcal{B}} e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}} \\
&\quad = \frac{\prod_{i \in \mathcal{A}} e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}}{\prod_{i \in \mathcal{B}} e^{b_{1i} + \mathbf{W}_{i\bullet}^\top \mathbf{x}_2}},
\end{aligned}$$

where the last equality relies on the cancellation that can be carried out thanks to the assumption in (A.2). The right-hand side in the display above equals the right-hand side in the detailed balance condition (A.5), which is thereby verified.

A.2 Proof of Proposition 2

We begin by mapping our notation into that of Jiang, Wu, Jin, and Wong (2018) (JWJW, henceforth). Set $\sigma = 1$ without loss of generality. Define the parameter vector θ by stacking the model parameters in $(\mathbf{W}, \mathbf{b})$, and define the function $\phi(\mathbf{x})$ by stacking $\mathbf{x}$ and the interactions of the variables in $\mathbf{x}$:

$$\theta \equiv \begin{pmatrix} \mathbf{b} \\ \mathbf{W}_{1\bullet} \\ \mathbf{W}_{2\bullet} \\ \vdots \\ \mathbf{W}_{n_1\bullet} \end{pmatrix} \in \mathbb{R}^{n + n_1 n_2} \qquad \text{and} \qquad \phi(\mathbf{x}) \equiv \begin{pmatrix} \mathbf{x} \\ x_{11}\mathbf{x}_2 \\ x_{12}\mathbf{x}_2 \\ \vdots \\ x_{1n_1}\mathbf{x}_2 \end{pmatrix} \in \mathbb{R}^{n + n_1 n_2}. \tag{A.6}$$

Let $\Theta \subset \mathbb{R}^d$ with $d \equiv n + n_1 n_2$ denote the compact parameter space where θ is a priori known to lie. We distinguish the true model parameters, $\theta_\star$, by the asterisk.

Define the cumulant generating function $\Lambda(\theta) \equiv \log \sum_{\mathbf{x} \in \{0,1\}^n} e^{\theta^\top \phi(\mathbf{x})}$. The sBRE probability (3) of a participation profile $\mathbf{x}$ belongs to the **canonical exponential family** because it can be rewritten in the form $\mathbb{P}(\mathbf{x} \mid \theta) = e^{\theta^\top \phi(\mathbf{x}) - \Lambda(\theta)}$.

We now verify each of the assumptions A1–A6 for JWJW's Theorem 2.1 one by one.

Assumption A1 requires that the parameter space Θ be a convex and compact subset of the

natural parameter domain $\mathcal{D} \equiv \left\{\theta \in \mathbb{R}^d : \Lambda(\theta) < \infty\right\}$, and that the true parameter, denoted here by $\theta_\star$, lie in the interior of Θ. The compactness of Θ and the interiority of $\theta_\star$ are the maintained assumptions in our paper. The inclusion $\Theta \subset \mathcal{D}$ is immediate because $\Lambda(\theta) < \infty$ for all $\theta \in \mathbb{R}^d$.

Assumption A2 requires that there exist a positive constant L such that, for all $\theta \in \Theta$, we have $\chi\left(\mathbb{P}(\cdot \mid \theta_\star), \mathbb{P}(\cdot \mid \theta)\right) \leq L\|\theta_\star - \theta\|$, where χ is a notion of divergence between two distributions described in JWJW's Definition 2.1. JWJW note that A2 holds whenever the model belongs to an exponential family with $\Lambda(\theta) < \infty$, which is the case here.

Informally, assumption A3 requires that the best-response dynamics converge to the stationary distribution uniformly fast in θ. Formally, JWJW state assumption A3 as the requirement that a certain $\mathcal{L}^2$-spectral gap be bounded away from zero uniformly in θ. In order to restate their condition for our finite setting, we introduce additional notation. Let $\mathbb{T}_\theta = \left(\mathbb{T}_\theta(\mathbf{x}, \mathbf{z})\right)_{\mathbf{x},\mathbf{z}}$ denote the (ergodic) transition matrix of the Markov chain induced by the best-response dynamics in the two-sided market parameterized by θ. Let $\mathbf{h} = \left(h(\mathbf{x})\right)_{\mathbf{x}} \in \mathbb{R}^{2^n}$ denote a vector whose entries are indexed by a participation profile $\mathbf{x}$. Let $\mathbb{T}_\theta \mathbf{h} \in \mathbb{R}^{2^n}$ denote the vector defined by $\left(\mathbb{T}_\theta \mathbf{h}\right)(\mathbf{x}) \equiv \sum_{\mathbf{z}} \mathbb{T}_\theta(\mathbf{x}, \mathbf{z}) h(\mathbf{z})$. Let $\langle \cdot, \cdot \rangle_\theta$ denote the inner product defined by $\langle \mathbf{h}, \mathbf{h}' \rangle_\theta \equiv \sum_{\mathbf{x} \in \{0,1\}^n} h(\mathbf{x}) h'(\mathbf{x}) \mathbb{P}(\mathbf{x} \mid \theta)$ for all vectors $\mathbf{h}$ and $\mathbf{h}'$. Let $\|\cdot\|_\theta$ denote the norm induced by $\langle \cdot, \cdot \rangle_\theta$. Define $\mathbf{1} = (1, \ldots, 1)$. With this notation, JWJW's assumption A3 becomes

$$\sup_{\theta \in \Theta} \sup_{\|\mathbf{h}\|_\theta = 1} \left\{\|\mathbb{T}_\theta \mathbf{h}\|_\theta : \langle \mathbf{h}, \mathbf{1} \rangle_\theta = 0\right\} < 1, \tag{A.7}$$

which—as we explain below—is a fancy way of specifying that the second largest eigenvalue modulus of the transition matrix $\mathbb{T}_\theta$ must be bounded away from 1 uniformly in θ. That is, the spectral gap must be positive. Because, for all the action-revision protocols that we consider, the Markov chain induced by the associated best-response dynamics is irreducible and aperiodic for all θ, the Perron–Frobenius theorem implies uniqueness of the largest eigenvalue, 1. This uniqueness immediately implies a positive spectral gap. This gap is bounded away from zero uniformly in θ because θ is assumed to come from a compact set Θ, and because all eigenvalues are continuous in θ.

It remains to explain why the inequality in (A.7) is concerned with a spectral gap. Because the stationary distribution $\mathbb{P}(\cdot \mid \theta)$ satisfies the detailed balance condition (A.1), introduced and verified in the proof of Proposition 1, the operator $\mathbb{T}_\theta$ is self-adjoint (see, e.g., Brémaud, 2020, Theorem 9.1.2, p. 290); that is, $\langle \mathbf{h}, \mathbb{T}_\theta \mathbf{h}' \rangle_\theta = \langle \mathbb{T}_\theta \mathbf{h}, \mathbf{h}' \rangle_\theta$ for all $\mathbf{h}$ and $\mathbf{h}'$. Then, by the spectral theorem, there exists an orthonormal basis of the space on which $\mathbb{T}_\theta$ acts: $\mathbb{T}_\theta = \sum_{i=1}^{2^n} \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$, where each λ_i is an eigenvalue, and $\mathbf{v}_i$ is the corresponding eigenvector with $\|\mathbf{v}_i\|_\theta = 1$. (Even though the eigenvalues and the eigenvectors depend on θ, we suppress this dependence for brevity.) Let $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{2^n}|$ without loss of generality. By the Perron–Frobenius theorem, the largest eigenvalue is $|\lambda_1| = 1$, with the corresponding eigenvector being $\mathbf{v}_1 = \mathbf{1}$, and the second-largest eigenvalue modulus is $|\lambda_2| < 1$. Therefore, for any $\mathbf{h}$ with $\langle \mathbf{h}, \mathbf{1} \rangle_\theta = 0$, there exist coefficients

$\{c_i\}_{i=2}^{2^n}$ such that $\mathbf{h} = \sum_{i=2}^{2^n} c_i v_i$. If, in addition, $\mathbf{h}$ satisfies $\|\mathbf{h}\|_\theta = 1$, then

$$\|\mathbb{T}_\theta \mathbf{h}\|_\theta = \left(\sum_{i=2}^{2^n} c_i^2 \lambda_i^2\right)^{\frac{1}{2}} \leq |\lambda_2| \left(\sum_{i=2}^{2^n} c_i^2\right)^{\frac{1}{2}} = |\lambda_2| \, \|\mathbf{h}\|_\theta = |\lambda_2| \,,$$

with equality if and only if $\mathbf{h} = \mathbf{v}_2$. Therefore, the inner sup in (A.7) equals the second largest eigenvalue modulus.

Assumption A4 requires the random variable $\phi(\mathbf{x})$ when $\mathbf{x}$ is distributed according to a certain probability distribution on $\{0,1\}^n$ to be subexponential (i.e., to abhor "fat tails"). Because each $\mathbf{x}$ in $\{0,1\}^n$ is bounded, $\phi(\mathbf{x})$ is also bounded and, therefore, subexponential for all distributions of $\mathbf{x}$.

Assumption A5 requires the expected potential after m steps of the best-response dynamics to depend continuously on θ. Formally, letting $\mathbf{x}^m$ denote the random variable that is the participation profile after m steps of the best-response dynamics, A5 requires the expectation $\mathbb{E}\left[\theta^\top \phi(\mathbf{x}^m) \mid \mathbf{x}, \theta\right]$ to be Lipschitz continuous in θ for all $\mathbf{x}$. Note that, in our case, the m-step transition matrix is continuously differentiable in θ, and, therefore, the expectation is also continuously differentiable. A continuously differentiable function is Lipschitz continuous on the compact set Θ.

Assumption A6 bounds the variance-covariance matrix of $\phi(\mathbf{x}^m)$ conditional on $\mathbf{x}$ and θ. Because all $\mathbf{x}$ in $\{0,1\}^n$ are bounded, $\phi(\mathbf{x})$ is also bounded. The variance-covariance matrix is therefore also appropriately bounded.

A.3 Proof of Proposition 3

Consider an instance of the two-sided market in which each entry of the matrix $\mathbf{W}$ is 1, and each entry of the vector $\mathbf{b}$ is $-\beta$ for some positive scalar β that satisfies

$$\frac{1}{\beta} < \frac{1}{n_1} + \frac{1}{n_2} < \frac{2}{\beta}. \tag{A.8}$$

For any participation profile $\mathbf{x}$, the potential satisfies:

$$\begin{aligned}
\mathbf{b}^\top \mathbf{x} + \mathbf{x}^\top{}_1 \mathbf{W} \mathbf{x}_2 &= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} x_{1i}x_{2j} - \beta\left(\sum_{i=1}^{n_1} x_{1i} + \sum_{j=1}^{n_2} x_{2j}\right) \\
&= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\left(x_{1i}x_{2j} - \frac{\beta x_{1i}}{n_2} - \frac{\beta x_{2j}}{n_1}\right) \\
&\leq \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \max_{(x,y)\in\{0,1\}^2}\left(xy - \frac{\beta x}{n_2} - \frac{\beta y}{n_1}\right) \\
&= n_1 n_2 \max\left\{0, 1 - \frac{\beta}{n_2} - \frac{\beta}{n_1}\right\} \\
&= 0,
\end{aligned}$$

where the last equality follows from the first inequality in (A.8). As a result, the potential is bounded above by zero. This upper bound is uniquely attained at $\mathbf{x} = \mathbf{0}$. Therefore, as $\sigma \to 0$, sBRE converges in probability to $\mathbf{x} = \mathbf{0}$, the maximizer of the potential. The associated welfare converges to zero.

For any participation profile $\mathbf{x}$, the utilitarian welfare satisfies:

$$\begin{aligned}
\mathbf{b}^\top \mathbf{x} + 2\mathbf{x}^\top{}_1 \mathbf{W} \mathbf{x}_2 &= 2\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} x_{1i}x_{2j} - \beta\left(\sum_{i=1}^{n_1} x_{1i} + \sum_{j=1}^{n_2} x_{2j}\right) \\
&= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\left(2x_{1i}x_{2j} - \frac{\beta x_{1i}}{n_2} - \frac{\beta x_{2j}}{n_1}\right) \\
&\leq n_1 n_2 \max_{x,y}\left(2xy - \frac{\beta x}{n_2} - \frac{\beta y}{n_1}\right) \\
&= n_1 n_2 \max\left\{0, 2 - \frac{\beta}{n_2} - \frac{\beta}{n_1}\right\} \\
&= n_1 n_2 \left(2 - \frac{\beta}{n_2} - \frac{\beta}{n_1}\right),
\end{aligned}$$

where the last equality follows from the second inequality in (A.8). As a result, the welfare is bounded above by the positive scalar $n_1 n_2 \left(2 - \frac{\beta}{n_2} - \frac{\beta}{n_1}\right)$. This upper bound is achieved at the participation profile $\mathbf{x} = (1, \ldots, 1)$.

Conclude that, as $\sigma \to 0$, for the specified instance $(\mathbf{W}, \mathbf{b})$, the welfare at sBRE converges to zero, whereas the first-best welfare is a positive constant, independent of σ. The ratio of equilibrium welfare to the first-best welfare vanishes, as claimed.

A.4 Proof of Proposition 5

Set $\sigma = 1$. Let $n_1 = n_2 = 2$. Side-1 agents are denoted by 11 and 12 and side-2 agents by 21 and 22. For some positive scalar parameter w, the strength of the heterogeneous network effects, denoted by a matrix $\mathbf{W}$, and the standalone values from participation, denoted by a vector $\mathbf{b} = (\mathbf{b}_1, \mathbf{b}_2)$, are given by

$$\mathbf{W} = \begin{pmatrix} w & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{b}_1 = \mathbf{b}_2 = \begin{pmatrix} -w/2 \\ 0 \end{pmatrix}.$$

With the parameter w implicit, let $\mathbb{E}^{true}$ denote the expectation under the sBRE distribution $\mathbb{P}$ defined in (3), which we now denote by $\mathbb{P}^{true}$ and, in the context of the present example, calculate to be

$$\mathbb{P}^{true}(\mathbf{x} \mid \mathbf{W}, \mathbf{b}) \propto \exp\left(-\left(\frac{w}{2}x_{11} + \frac{w}{2}x_{21}\right) + w x_{11} x_{21}\right). \tag{A.9}$$

The realizations of the participation profiles drawn from $\mathbb{P}^{true}$ are the (unlimited) **data** that a platform uses to infer model parameters. The **correctly specified platform** anticipates heterogeneous network effects and, as a result, identifies w and $\mathbf{b}$ correctly from the data (e.g., using the

maximum-likelihood technique, as per Lemma 1). Then, the platform chooses the pricing scheme $(\mathbf{P}, \mathbf{p})$ to maximize the expected realized profit, denoted by $\Pi_w^{true}(\mathbf{P}, \mathbf{p})$.

We can bound the optimal revenue $\sup_{\mathbf{P},\mathbf{p}} \Pi_w^{true}(\mathbf{P}, \mathbf{p})$ from below by the revenue under the feasible pricing scheme that has $P_{11} = w/2$, $p_{11} = p_{21} = -w/4$, and all the remaining entries in $\mathbf{P}$ and $\mathbf{p}$ set to zero. The associated revenue bound can be verified to be

$$B(w) = \frac{w}{4} \times \frac{1 - e^{-\frac{w}{4}}}{1 + e^{-\frac{w}{4}}},$$

where the second term is increasing in w. For $w \geq 4 \log 3$, we have $\left(1 - e^{-\frac{w}{4}}\right) / \left(1 + e^{-\frac{w}{4}}\right) \geq \frac{1}{2}$, in which case the bound itself can be bounded from below to conclude:

$$\sup_{\mathbf{P},\mathbf{p}} \Pi_w^{true}(\mathbf{P}, \mathbf{p}) \geq B(w) \geq \frac{w}{8} \qquad \text{for} \quad w \geq 4 \log 3. \tag{A.10}$$

The **misspecified platform** assumes that the network effects—which it denotes by a matrix $\mathbf{U}$—are uniform:

$$\mathbf{U} = \begin{pmatrix} u & u \\ u & u \end{pmatrix} \qquad \text{for some} \quad u > 0.$$

The platform denotes the standalone values by a vector $\mathbf{a}$. With this notation, the misspecified platform believes the sBRE probability of a participation profile $\mathbf{x}$ to be

$$\mathbb{P}^{miss}(\mathbf{x} \mid \mathbf{U}, \mathbf{a}) \equiv \frac{\exp\left(\mathbf{a}^\top \mathbf{x} + u\,(x_{11} + x_{12})\,(x_{21} + x_{22})\right)}{\sum_{\mathbf{y} \in \{0,1\}^n} \exp\left(\mathbf{a}^\top \mathbf{y} + u\,(y_{11} + y_{12})\,(y_{21} + y_{22})\right)}.$$

With the parameters u and $\mathbf{a}$ implicit, let $\mathbb{E}^{miss}$ denote the expectation under the misspecified distribution $\mathbb{P}^{miss}$. In order to identify u and $\mathbf{a}$, the misspecified platform maximizes the likelihood derived from the misspecified model. The associated first-order conditions, analogous to those in (5), equate the moments in the data (the left-hand sides below) to the theoretical moments based on the misspecified model (the right-hand sides below):

$$\mathbb{E}^{true}\left[(x_{11} + x_{12})\,(x_{21} + x_{22})\right] = \mathbb{E}^{miss}\left[(x_{11} + x_{12})\,(x_{21} + x_{22})\right] \tag{A.11}$$

$$\mathbb{E}^{true}\left[x_{1i}\right] = \mathbb{E}^{miss}\left[x_{1i}\right], \qquad i = 1, 2 \tag{A.12}$$

$$\mathbb{E}^{true}\left[x_{2j}\right] = \mathbb{E}^{miss}\left[x_{2j}\right], \qquad j = 1, 2. \tag{A.13}$$

The distribution $\mathbb{P}^{true}$ in (A.9) implies that x_{12} and x_{22}—the variables that are not linked by the network effect w—are i.i.d. Bernoulli$\left(\frac{1}{2}\right)$ and independent of the variables x_{11} and x_{21}, which are linked by the network externality. For the linked variables x_{11} and x_{21}, the probabilities of the four outcomes $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$ are proportional to

$$1, e^{-\frac{w}{2}}, e^{-\frac{w}{2}}, 1.$$

As a result, the moments to be matched in (A.12)–(A.13) are

$$\mathbb{E}^{true}[x_{11}] = \mathbb{E}^{true}[x_{12}] = \mathbb{E}^{true}[x_{21}] = \mathbb{E}^{true}[x_{22}] = \frac{1}{2}. \tag{A.14}$$

The moment to be matched in (A.11) is computed from $\mathbb{E}^{true}[x_{11}x_{21}] = 1/\left(2\left(1+e^{-w/2}\right)\right)$ and $\mathbb{E}^{true}[x_{1i}x_{2j}] = \mathbb{E}^{true}[x_{1i}]\,\mathbb{E}^{true}[x_{2j}] = \frac{1}{4}$ whenever $(i,j) \neq (1,1)$. This moment equals

$$\mathbb{E}^{true}[(x_{11}+x_{12})(x_{21}+x_{22})] = \frac{3}{4} + \frac{1}{2\left(1+e^{-w/2}\right)}, \tag{A.15}$$

which exceeds $\frac{3}{4}$, does not exceed $\frac{5}{4}$, and, as $w \to \infty$, converges to $\frac{5}{4}$.

Because all expectations in (A.14) are the same, the misspecified platform uses (A.12) and (A.13) to infer (mistakenly) that the standalone values are the same for all agents: $\mathbf{a} = (a, \ldots, a)$ for some a. With this simplification, the probability $\mathbb{P}^{miss}(\mathbf{x} \mid \mathbf{U}, \mathbf{a})$ of a participation profile $\mathbf{x}$ can be written as

$$\mathbb{P}^{miss}(\mathbf{x} \mid \mathbf{U}, \mathbf{a}) \propto \exp\left(a\left(x_{11}+x_{12}+x_{21}+x_{22}\right) + u\left(x_{11}+x_{12}\right)\left(x_{21}+x_{22}\right)\right).$$

Then, the moment conditions (A.11)–(A.13) jointly require $a = -u$, where u solves the moment condition (A.11), which, using (A.15) and the display above, can be rewritten as

$$\frac{3}{4} + \frac{1}{2\left(1+e^{-w/2}\right)} = g(u), \qquad \text{where} \quad g(u) \equiv \frac{2+6e^{-u}}{1+6e^{-u}+e^{-2u}}. \tag{A.16}$$

Because the left-hand side is contained in $\left(\frac{3}{4}, \frac{5}{4}\right)$, and because g satisfies $g'(u) > 0$, $g(0) = 1$, and $\lim_{u\to\infty} g(u) = 2$, the equation in (A.16) has a unique solution for u, which we denote by u_w. This solution is positive, continuous in w, and bounded above uniformly in w.

The misspecified platform maximizes the expected revenue under misspecified belief:

$$\max_{\mathbf{P},\mathbf{p}} \sum_{\mathbf{x}\in\{0,1\}^n} \mathbb{P}^{miss}(\mathbf{x} \mid \mathbf{U}_w - \mathbf{P}, \mathbf{a}_w - \mathbf{p})\left(\mathbf{p}^\top\mathbf{x} + 2\mathbf{x}_1^\top\mathbf{P}\mathbf{x}_2\right),$$

where $(\mathbf{U}_w, \mathbf{a}_w)$ depend on w through u_w. For each u_w, the optimal prices in the display above are bounded because prices enter the revenue linearly but are collected with probabilities that vanish exponentially when prices rise. Because u_w is bounded in w, the optimal prices are also bounded in w. As a result, the maximization problem in the display above consists in maximizing a continuous function over a compact set. By the Weierstrass theorem, a solution, denoted by $(\mathbf{P}_w^{miss}, \mathbf{p}_w^{miss})$, exists. Because each component of the solution is bounded uniformly in w by a constant that we denote by M, the revenue that $(\mathbf{P}_w^{miss}, \mathbf{p}_w^{miss})$ delivers under the correctly specified model is bounded, too:

$$\mathbb{E}^{true}\left[\left(\mathbf{p}_w^{miss}\right)^\top\mathbf{x} + 2\mathbf{x}_1^\top\mathbf{P}_w^{miss}\mathbf{x}_2\right] \leq 4M + 2\times 4M = 12M < \infty. \tag{A.17}$$

We conclude that

$$\lim_{w\to\infty}\frac{\mathbb{E}^{true}\left[\left(\mathbf{p}_w^{miss}\right)^\top\mathbf{x}+2\mathbf{x}_1^\top\mathbf{P}_w^{miss}\mathbf{x}_2\right]}{\sup_{\mathbf{P},\mathbf{p}}\Pi_w^{true}\left(\mathbf{P},\mathbf{p}\right)}\leq\lim_{w\to\infty}\frac{12M}{B\left(w\right)}=\lim_{w\to\infty}\frac{12M}{w/8}=0,$$

where use has been made of profit bounds in (A.10) and (A.17). The proposition's conclusion follows.

A.5 Proof of Remark 2

Below we relate the assumptions of Theorem 4.2 of Glaser, Huang, and Gretton (2024) (GHG, for short) to the previously verified (in the proof of Proposition 2) assumptions of Jiang, Wu, Jin, and Wong (2018) (JWJW, for short).

GHG's assumptions A1, A2, and A5 are the same as JWJW's respective assumptions A1, A2, and A5.

GHG's A3 is weaker than JWJW's A3, as discussed towards the end of GHG's Section 3.1.

GHG's A6 is weaker than JWJW's corresponding condition A4 (not A6); GHG's A6 imposes subexponential tails only for the true parameter, whereas JWJW's A4 imposes subexponential tails for all parameters.

GHG's A4 is stronger than JWJW's corresponding condition A6 (not A4) but still holds in our setting, as we now show. In order to state GHG's A4, just as in the proof of Proposition 2 in Appendix A.2, we let $\mathbb{T}_\theta$ denote the transition matrix of the Markov chain induced by the best-response dynamics in the two-sided market with parameters $(\mathbf{W},\mathbf{b})$ stacked into the vector θ, as in (A.6). We let $\mathbb{T}_\theta^{[m]}$ denote the associated m-step CD transition matrix. For each positive integer m and all $\mathbf{x}\in\{0,1\}^n$, let $\mathcal{T}_\theta^{[m]}\left(\mathbf{x}\right)$ denote a random variable whose law is given by the transition distribution $\mathbb{T}_\theta^{[m]}\left(\mathbf{x},\cdot\right)$ for m CD steps. GHG's A4 requires that there exist a $\nu>2$ such that for all $m\in\mathbb{N}$, there exist a $\kappa_{\nu,m}<\infty$ such that

$$\sup_{\mathbf{x}}\sup_{\theta}\left[\mathbb{E}\left\|\phi\left(\mathcal{T}_\theta^{[m]}\left(\mathbf{x}\right)\right)-\mathbb{E}\left[\phi\left(\mathcal{T}_\theta^{[m]}\left(\mathbf{x}\right)\right)\right]\right\|^\nu\right]^{1/\nu}\leq\kappa_{\nu,m},$$

where $\|\cdot\|$ denotes the Euclidean norm, and ϕ maps a participation profile into a binary vector as described in (A.6). (JWJW's A6—not A4—is weaker because it replaces $\nu>2$ in the condition above with $\nu\geq 2$.) Because $\phi\left(\mathbf{x}\right)\in\{0,1\}^{n+n_1n_2}$ for all $\mathbf{x}\in\{0,1\}^n$, the norm in the display above almost surely satisfies

$$\left\|\phi\left(\mathcal{T}_\theta^{[m]}\left(\mathbf{x}\right)\right)-\mathbb{E}\left[\phi\left(\mathcal{T}_\theta^{[m]}\left(\mathbf{x}\right)\right)\right]\right\|\leq\sqrt{n+n_1n_2}\qquad\text{for all}\quad\mathbf{x}\in\{0,1\}^n.$$

As a result, we have

$$\sup_{\mathbf{x}}\sup_{\theta}\left[\mathbb{E}\left\|\phi\left(\mathcal{T}_\theta^{[m]}\left(\mathbf{x}\right)\right)-\mathbb{E}\left[\phi\left(\mathcal{T}_\theta^{[m]}\left(\mathbf{x}\right)\right)\right]\right\|^\nu\right]^{1/\nu}\leq\kappa_{\nu,m}\equiv\sqrt{n+n_1n_2}$$

for all $\nu > 0$ and, in particular, for $\nu > 2$, as GHG's A4 requires.

Conclude that our model satisfies all the hypotheses in GHG's Theorem 4.2.

References

D. H. Ackley, G. E. Hinton, and T. J. Sejnowski. A learning algorithm for boltzmann machines. *Cognitive Science*, 9(1):147–169, 1985.

M. Armstrong. Competition in two-sided markets. *RAND Journal of Economics*, 37(3):668–691, 2006.

D. Avramov, S. Cheng, and L. Metzker. Machine learning vs. economic restrictions: Evidence from stock return predictability. *Management Science*, 69(5):2587–2619, May 2023.

Y. Babichenko and O. Tamuz. Graphical potential games. *Journal of Economic Theory*, 163: 889–899, 2016.

C. Ballester, A. Calvó-Armengol, and Y. Zenou. Who's who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417, 2006.

L. E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5: 387–424, 1993.

R. H. Bork. *The Antitrust Paradox: A Policy at War with Itself*. Basic Books, New York, 1978.

P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation and Queues*, volume 31. Springer, 2nd edition, 2020.

B. Caillaud and B. Jullien. Chicken & egg: competition among intermediation service providers. *RAND Journal of Economics*, 34(2):309–328, 2003.

L. T. Chan. Divide and conquer in two-sided markets: A potential-game approach. *RAND Journal of Economics*, 52(4):839–858, 2021.

L. T. Chan. Weight-ranked divide-and-conquer contracts. *Theoretical Economics*, 20:857–882, 2025.

L. Chen, M. Pelger, and J. Zhu. Deep learning in asset pricing. *Management Science*, 70(2): 714–750, 2024.

A. de Cornière, A. Mantovani, and S. Shekhar. Third-degree price discrimination in two-sided markets. *Management Science*, 71(4):3340–3356, April 2025.

T. Denti. Unrestricted information acquisition. *Theoretical Economics*, 18:1101–1140, 2023.

C. Farronato, J. Fong, and A. Fradkin. Dog eat dog: Balancing network effects and differentiation in a digital platform merger. *Management Science*, 70(1):464–483, 2024.

P. Glaser, K. H. Huang, and A. Gretton. Near-optimality of contrastive divergence algorithms. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 91036–91090. Curran Associates, Inc., 2024.

R. Gomes and A. Pavan. Many-to-many matching and price discrimination. *Theoretical Economics*, 11:1005–1052, 2016.

S. Goyal and J. L. Moraga-González. R&d networks. *RAND Journal of Economics*, 32(4):686–707, 2001.

S. J. Grossman and J. E. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408, 1980.

M. Gutmann and A. Hyvärinen. Noise-contrastive estimation of unnormalized statistical models, with applications to natural image statistics. *Journal of Machine Learning Research*, 13:307–361, 2012.

P. A. Haile, A. Hortaçsu, and G. Kosenok. On the empirical content of quantal response equilibrium. *American Economic Review*, 98(1):180–200, 2008.

S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

B. Hébert and J. La'O. Information acquisition, efficiency, and nonfundamental volatility. *Journal of Political Economy*, 131(10):2666–2723, 2023.

C. C. Heyde. On a property of the lognormal distribution. *Journal of the Royal Statistical Society: Series B (Methodological)*, 25(2):392–393, 1963.

G. E. Hinton. Training products of experts by minimizing contrastive divergence. *Neural Computation*, 14(8):1771–1800, 2002.

G. J. Hitsch, A. Hortaçsu, and D. Ariely. Matching and sorting in online dating. *American Economic Review*, 100(1):130–163, 2010.

T. Hoshino and R. Pancs. Logit-response dynamics under sampled observation. *Working Paper*, 2026.

A. Hyvärinen. Consistency of pseudolikelihood estimation of fully visible boltzmann machines. *Neural Computation*, 18(10):2283–2292, 2006.

A. Hyvärinen. Some extensions of score matching. *Computational Statistics & Data Analysis*, 51: 2499–2512, 2007.

M. Igami. Artificial intelligence as structural estimation: Deep blue, bonanza, and alphago. *Econometrics Journal*, 23:S1–S24, 2020.

B. Jiang, T.-Y. Wu, Y. Jin, and W. H. Wong. Convergence of contrastive divergence algorithm in exponential family. *Annals of Statistics*, 46(6A):3067–3098, 2018.

B. Jullien, A. Pavan, and M. Rysman. Two-sided markets, pricing, and network effects. In K. Ho, A. Hortaçsu, and A. Lizzeri, editors, *Handbook of Industrial Organization*, volume 4, chapter 7, pages 485–592. Elsevier, Amsterdam, 2021.

A. Kajii and S. Morris. The robustness of equilibria to incomplete information. *Econometrica*, 65 (6):1283–1309, 1997.

M. Kandori, G. J. Mailath, and R. Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56, 1993.

H. Larochelle and Y. Bengio. Classification using discriminative restricted boltzmann machines. In *Proceedings of the 25th International Conference on Machine Learning*, pages 536–543. Association for Computing Machinery, 2008.

R. S. Lee. Vertical integration and exclusivity in platform and two-sided markets. *American Economic Review*, 103(7):2960–3000, 2013.

C. F. Manski. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531–542, 1993.

R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10:6–38, 1995.

A. Mele. A structural model of dense network formation. *Econometrica*, 85(3):825–850, 2017.

D. Monderer and L. Shapley. Potential games. *Games and Economic Behavior*, 14(1):124–143, 1996.

J.-C. Rochet and J. Tirole. Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029, 2003.

R. Salakhutdinov and G. Hinton. An efficient learning procedure for deep boltzmann machines. *Neural Computation*, 24:1967–2006, 2012.

R. Salakhutdinov, A. Mnih, and G. Hinton. Restricted boltzmann machines for collaborative filtering. In *Proceedings of the 24th International Conference on Machine Learning*, pages 791–798. Association for Computing Machinery, 2007.

L. Samuelson and J. Steiner. Growth and likelihood. *CEPR Discussion Paper*, (18339), 2023.

P. Smolensky. Information processing in dynamical systems: Foundations of harmony theory. In D. Rumelhart and J. McClelland, editors, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 1: Foundations, pages 194–281. MIT Press, 1986.

J. Sohl-Dickstein, P. B. Battaglino, and M. R. DeWeese. New method for parameter estimation in probabilistic models: Minimum probability flow. *Physical Review Letters*, 107(22):220601, 2011.

H. Su, M. V. Tretyakov, and D. P. Newton. Deep learning of transition probability densities for stochastic asset models with applications in option pricing. *Management Science*, 71(4):2922–2952, 2025.

T. Ui. Robust equilibria of potential games. *Econometrica*, 69(5):1373–1380, 2001.

M. J. Wainwright and M. I. Jordan. *Graphical Models, Exponential Families, and Variational Inference*, volume 1 of *Foundations and Trends in Machine Learning*. Now Publishers, 2008.

Y. Wang and J. Zeng. Predicting drug-target interactions using restricted boltzmann machines. *Bioinformatics*, 29(13):i126–i134, 2013.

H. Young. The evolution of conventions. *Econometrica*, 61(1):57–84, 1993.