Peer Effects in Consideration and Preferences[*]

Nail Kashaev   Natalia Lazzati   Ruli Xiao[†]

This version: September 1, 2025

First version: April 13, 2019, arXiv:1904.06742

Abstract   We develop a general model of discrete choice that incorporates peer effects in preferences and consideration sets. We characterize the equilibrium behavior and establish conditions under which all parts of the model can be recovered from a sequence of choices. We allow peers to affect preferences, consideration, or both. We show that these peer-effect mechanisms have different behavioral implications in the data. This allows us to recover the set and the type of connections between the agents in the network. We then use this information to recover each agent's preferences and consideration mechanisms. These nonparametric identification results allow for general forms of heterogeneity across agents and do not rely on the variation of either exogenous covariates or the set of available options (menus). We apply our results to model expansion decisions by tea chains and find evidence of limited consideration. We simulate counterfactual predictions and show how limited consideration slows market penetration and competition.

JEL codes: C31, C33, D83, O33

Keywords: Peer Effects, Random Preferences, Random Consideration Sets

1. Introduction

Agents are known to be influenced by others when making decisions (Durlauf and Young, 2001). This social influence has been shown to be important for people in areas such as health and education and for the decisions of firms such as opening a new store. It has also been argued that agents can influence each other in different ways (Manski, 2000). A comprehensive social influence approach is needed to understand the mechanisms by which the interactions operate in practice and inform private and public policies. We offer a model of social influence where the choices of connected agents or peers shape the subset of options that the agent ends up considering[1] and her preferences over the alternatives. We show that these two mechanisms have different behavioral implications in the data and that variation in the choices made by connected agents over time allows us to recover all parts of the model. We illustrate our ideas with an empirical application of expansion decisions of the two largest tea chains in the high-end tea industry in China.

In our model, a fixed group of agents are linked through a given network. A linked agent or peer can affect the options an agent considers, how she ranks the alternatives, or both.[2] More to the point, at a randomly given time, an agent gets the opportunity to select a new option from a finite set of alternatives. Like in standard consideration set models, the agent does not pay attention to all the available options. Instead, she first forms a consideration set and then picks an option from it. The distinctive feature of our model is that the probability that a given alternative enters the consideration set depends on the number of peers currently adopting that option.[3] We also allow the choices made by her

---

[1]This possibility has been (explicitly or implicitly) discussed by other researchers in specific contexts —e.g., the choices made by friends may help us discover a new television show (Godes and Mayzlin, 2004), a new welfare program (Caeyers, 2014), a new retirement plan (Duflo and Saez, 2003), a new restaurant (Qiu et al., 2018), or an opportunity to protest (Enikolopov et al., 2020).

[2]Throughout the paper, we use a behavioral definition of peers: for a given agent, her peers are defined as all other agents that directly impact the choices that the agent makes.

[3]As in most of the literature, we assume that consideration sets are (ex-ante) random. While there are alternative interpretations, one could motivate random preferences by saying that when subjects make a decision, they maximize a well-defined utility function (or preference), but this changes stochastically over time. We interpret random consideration in our model in a similar way. Our set-up allows current consideration to depend on current choices, but does not allow the dependence of current consideration on the previously considered (but not picked) alternatives —Remark 3 discusses this important limitation of the model.

peers to affect the preferences of the agent over alternatives.

Our model might fit in a large number of applications. In the domain of consumer behavior, we can think of an online platform that offers video games to a set of players (agents).[4] The number of games offered by the platforms is often quite large, so agents might not be able to pay attention to all of them when making a decision. Platforms often allow agents to form social networks and make the last purchased or played game by peers visible to the agent. This might shrink the subset of games she ends up considering. Some of these games are played in groups, so the choices made by her peers can directly affect the utility the agent gets from playing a particular game. If we observe choices made by the players over time, our model can help the platform personalize each agent's reference groups to maximize profits. While well-suited for many applications, our model requires repeated choices of agents in the network. Thus, this framework does not directly fit, for instance, a situation where alternatives involve durable goods such as different brands of vehicles.

After showing equilibrium existence, we consider a long sequence of choices made by the network members from an invariant, latent network structure and stable-over-time preferences and consideration set formation mechanisms. We show that all primitives of the model can be uniquely recovered: the network structure, the consideration probabilities, and the distribution of choices given a consideration set (which we associate with preferences). Three aspects of our nonparametric identification results deserve special attention. First, we allow for very general forms of heterogeneity across agents. Second, we identify not only who influences whom, but also whether a peer affects consideration, preferences, or both. Third, we do not rely on changes in exogenous covariates or the set of available options (menus) to identify the model. Instead, we use variation in the choices made by peers. One can think of them as excluded covariates that affect different parts of the decision process.[5] These excluded covariates are special as they vary *endogenously* in the model and we need to *identify* them in the data.

In our model, the observed choices are generated by a system of conditional choice probabilities (CCPs): each CCP specifies the frequency of choices made by a given agent

---

[4]Lee (2015) finds that the likelihood of a player adopting a particular game increases as more of her online friends have previously adopted it.

[5]We thank Francesca Molinari for pointing this out.

conditional on the choice configuration (i.e., choices of everyone in the network) at the moment of revising her selection. The identification strategy has two parts. First, we identify the primitives of the model from the CCPs. Second, we recover the CCPs from the data.

To identify the model from the CCPs, we proceed in steps. First, we observe that changes in an agent's peers' choices lead to changes in her own CCPs. This helps us identify who the peers are, but not whether their influence comes from affecting consideration or preferences. We separate the two mechanisms using the following feature: for a given agent, the probability of selecting an alternative can be written as the probability of considering the alternative times the probability of selecting it conditional on it being considered. These two terms capture the peer effect channels by the consideration and the preference, respectively. While the first probability changes when a consideration peer switches to that alternative, the second probability remains constant since the agent is already considering the alternative. Also, while the second probability varies when a preference peer chooses something different from the alternative, the first probability is not affected by this change. In other words, an interdependence between alternatives is present in preferences, but not in consideration. We use these observations to separate the two mechanisms via a cross order effect of peers in alternatives in the CCPs.

When the network structure is recovered, choices made by different types of peers can be used as double exclusion restrictions to identify consideration mechanisms and preferences. For example, variation in choices made by peers that only affect consideration (consideration-only peers) can be used to identify changes in consideration probabilities. To recover preferences, we first show that variation in choices made by consideration-only peers can be used to mimic variation in menus. Thus, one can identify the CCPs for the cases in which a subset of alternatives has been completely removed from the menu. This artificial variation in menus generated by consideration-only peers can then be used to identify preferences.

To identify the CCPs, we assume the researcher observes a sequence of choices for the network members in a long time-series with a time-stable structure. We consider two datasets: In continuous-time datasets, the researcher observes agents' choices in real time, so one can identify the CCPs directly. Our empirical application is an example of this type of "ideal" dataset. In discrete-time datasets, the researcher observes the joint choices at fixed time

intervals (e.g., every Monday). In this case, we use Blevins (2017, 2018) to show that the CCPs are also uniquely identified under a mild condition.

We offer empirically relevant extensions: we study finite history dependence; and we explain what to do when one of the choices (e.g., "do nothing") is not observed in the data.

To showcase our methodology, we study the possibility of limited attention in the expansion decisions of China's top two high-end tea chains, Nayuki and Heytea. The set of markets in which these firms can open a new store is very large, and we observe their choices near the time when they started their business. As newcomers, they could lack knowledge about current and future market conditions (e.g., the level of consumer demand and local government regulations) or the potential actions of competitors, and thus, not be able to form expectations about market profitability. We think that managers may have employed simple rules (heuristics) to first narrow down the markets they considered, and then used their limited resources to evaluate those options more carefully. This idea builds on previous theories of bounded rationality in firms offered by Simon (1955).[6] We add neighborhood effects to these heuristics by assuming that, for a given market, the number of firm stores in the nearby markets affects the probability of considering that market, but not its profitability. This exclusion restriction allows us to identify the network structure —neighboring markets that affect consideration only (i.e. consideration-peers)— and to recover profits and consideration. Our results show that firms in our application limit their consideration, that ignoring this behavior can mislead the analysis of market profitability, and that limited consideration may play a key role in shaping market structure.

Finally, we relate our results with the existing literature. From a modeling perspective, our setup combines the dynamic model of social interactions of Blume (1993, 1995) with the (single-agent) model of random consideration sets of Manski (1977) and Manzini and Mariotti (2014). By adding peer effect in consideration sets, we use variation in the choices made by the peers of a given agent as the main tool to recover her random preferences. The literature on the identification of single-agent consideration set models has mainly relied on the variation of the set of available options (menus), e.g., Aguiar (2017), Aguiar et al.

---

[6]See also the discussion of boundedly rational firm behavior in Simon (1945), Armstrong and Huck (2010), and Heidhues and Kőszegi (2018).

(2016), Brady and Rehbeck (2016), Caplin et al. (2019), Cattaneo et al. (2020), Horan (2019), Kashaev ⓡ Aguiar (2022), Lleras et al. (2017), Manzini and Mariotti (2014), and Masatlioglu et al. (2012); or offered partial identification results, e.g., Barseghyan et al. (2021a). (See Aguiar et al., 2023, for a comparison of several consideration set models in an experiment.) Other papers have relied on covariates that shift preferences or consideration sets, e.g., Barseghyan et al. (2021b), Crawford et al. (2021), Conlon and Mortimer (2013), Draganska and Klapper (2011), Gaynor et al. (2016), Goeree (2008), Mehta et al. (2003), Lu (2022), and Roberts and Lattin (1991). Abaluck and Adams-Prassl (2021) show that choice probabilities in full consideration models satisfy a symmetry property analogous to Slutsky symmetry in continuous-choice models that breaks down in consideration set models when changes in characteristics perturb consideration. They use unbounded alternative-specific covariates to generate exogenous menu variation to identify the consideration probabilities. Aguiar and Kashaev (2024), Allen and Rehbeck (2023), Crawford et al. (2021), and Dardanoni et al. (2020) use repeated choices (i.e., panel data) but do not allow for peer effects.

There is a vast econometric literature on the identification of models of social interactions in which friends' choices affect the preferences but not the consideration set of a given person (see, for example, Blume et al., 2011, Bramoullé et al., 2020, De Paula, 2017 and Graham, 2015, for comprehensive reviews of this literature). We add to this literature a second mechanism of peer effects that might be important in specific applications.

As we mentioned earlier, we can recover from the data the set of connections between the agents in the network. In the context of linear models, a few recent papers have made progress in the same direction. Among them, Blume et al. (2015), Bonaldi et al. (2015), De Paula et al. (2023), Lewbel et al. (2023), and Manresa (2013). In the context of discrete choice, Chambers et al. (2023) also identify the network structure. However, in this paper, peers do not affect consideration sets but directly change preferences (among other differences).

Two theoretical papers incorporate peer effects in the consideration sets: Borah and Kops (2018) do so in a static framework and use menu variation for identification. Lazzati (2020) considers a dynamic model but focuses on two alternatives that can be acquired together.

In relation to the application, a large empirical literature addresses how firms make entry and expansion decisions in oligopolies. Part of this literature studies strategic interactions in

6

payoffs in static entry games (e.g., Bresnahan and Reiss, 1990, Berry, 1992, and Ciliberto and Tamer, 2009). These studies typically use cross-sectional data and address the challenge of multiple equilibria across markets. Another branch, more closely related to our work, models expansion decisions as the result of dynamic games (e.g., Aguirregabiria and Mira, 2007, Bajari et al., 2007, Arcidiacono et al., 2016, and Blevins et al., 2018). These papers use panel data, often assume a single equilibrium in the data-generating process, and add forward-looking behavior. By capturing the trade-off between the upfront cost and the potential benefits in the future, they show the incentives of firms to preempt their opponents by entering the market early. We study a third possible channel of interdependencies in firms' entry decisions via bounded rationality in managerial decision-making and introduce interaction effects in both payoffs and consideration sets. By modeling sequential revisions in a continuous-time setting, the underlying mechanisms and factors that produce the observed data in our set-up lead to only one equilibrium. Although we abstract away from firms' forward-looking behavior, we allow the history of rivals' actions to affect the firm's payoffs.[7] One could view marginal profit in our framework as the reduced-form value from expansion, which embeds forward-looking behaviors without explicitly modeling them. We leave for future research the modeling of limited consideration, multiplicity of equilibria, and forward-looking behavior altogether —this new model will require a different interpretation.

The rest of the paper is organized as follows. Section 2 presents the model, the main assumptions, and establishes equilibrium existence. Section 3 studies identification. Section 4 offers some extensions. Section 5 applies our model to expansion decisions by firms in the tea market, and Section 6 concludes. The Online Appendix contains the regularity conditions for identification, all the formal proofs, and extra results for identification and estimation.

---

[7]As we mentioned earlier, our application studies the expansion decisions of Nayuki and Heytea when they were relatively new in the market and of moderate size. Thus, a less-sophisticated model might explain their behavior well. We understand that a fully rational, forward-looking approach would be more appropriate to study well-established large-scale companies.

2. The Model

This section describes the model and the main assumptions in the paper. It also establishes the existence and uniqueness of equilibrium.

2.1. Network, Consideration Sets, and Preferences

Network and Choice Configuration There is a finite set of agents $\mathcal{A} = \{1, 2, \ldots, A\}$, $A \geq 2$, and a finite set (menu) of alternatives $\mathcal{Y} = \{0, 1, 2, \ldots, Y\}$, $Y \geq 1$, from which the agents might choose. Alternative 0 is called the default alternative. We refer to $\mathbf{y} = (y_a)_{a \in \mathcal{A}} \in \mathcal{Y}^A$ as a choice configuration.[8]

Agents are connected through a fixed given network and interact with others in different ways. Specifically, the choices made by the peers of a given agent can affect the set of alternatives the agent ends up considering, her preferences over the alternatives, or both. Thus, the network comprises two sets of edges in $\mathcal{A}$, $\Gamma = (\Gamma_C, \Gamma_R)$, where Γ_C and Γ_R are the sets of consideration and preference edges, respectively. Each edge identifies two connected agents and the direction of the connection. Hence, the reference group of Agent a consists of reference groups for consideration, $\mathcal{NC}_a$, and for preferences, $\mathcal{NR}_a$. Formally, for each $a \in \mathcal{A}$

$$\mathcal{NC}_a = \{a' \in \mathcal{A} : \exists \text{ edge from } a \text{ to } a' \text{ in } \Gamma_C\}, \mathcal{NR}_a = \{a' \in \mathcal{A} : \exists \text{ edge from } a \text{ to } a' \text{ in } \Gamma_R\}.$$

The full reference group is $\mathcal{N}_a = \mathcal{NC}_a \bigcup \mathcal{NR}_a$. We follow the convention and assume that $a \notin \mathcal{N}_a$. Since we allow for the possibility that some peers affect both consideration and preferences, $\mathcal{NCR}_a = \mathcal{NC}_a \bigcap \mathcal{NR}_a$ can be nonempty.

We will use a simple example throughout the paper to illustrate the main concepts, assumptions, and identification results.

Example 1. There are four agents and three alternatives. That is, $\mathcal{A} = \{1, 2, 3, 4\}$ and

---

[8]The model easily extends to settings where menus are agent-specific if, for every pair of agents, there is a one-to-one mapping between their choice sets.

$\mathcal{Y} = \{0, 1, 2\}$. The reference groups for consideration are:

$$\mathcal{NC}_1 = \{2, 3\}, \quad \mathcal{NC}_2 = \{1\}, \quad \mathcal{NC}_3 = \{2\}, \quad \mathcal{NC}_4 = \emptyset.$$

This means that, for instance, Agents 2 and 3 affect the set of alternatives that Agent 1 ends up considering. We assume that only Agents 1 and 3 affect each other preferences:

$$\mathcal{NR}_1 = \{3\}, \quad \mathcal{NR}_2 = \emptyset, \quad \mathcal{NR}_3 = \{1\}, \quad \mathcal{NR}_4 = \emptyset.$$

Agent 4 has no peers and Agent 3 affects both preferences and consideration of Agent 1. $\square$

Choice Revision Process We model the revision process as a standard continuous-time Markov process on the space of choice configurations. We assume agents are endowed with independent Poisson "alarm clocks" with rates $(\lambda_a)_{a \in \mathcal{A}}$. At randomly given moments (exponentially distributed with mean $1/\lambda_a$) the alarm of Agent a goes off.[9] When this happens, the agent forms a consideration set and then selects an alternative from it. Since the probability of any two alarm clocks going off simultaneously is zero, the probability that two agents make choices simultaneously is also zero. This fact simplifies identification and rules out multiple equilibria in the data-generating process.

Peer Effect in Consideration Sets The probability that Agent a pays attention to and, thereby, includes a particular alternative in her consideration set depends on the choice configuration at the moment of revising decisions. We indicate by $Q_a(v \mid \mathbf{y}, \mathcal{NC}_a)$ the ex-ante probability that Agent a considers alternative v given a choice configuration $\mathbf{y}$ and her consideration reference group $\mathcal{NC}_a$. We assume that each alternative is added to the consideration set independently from other alternatives.

Assumption 1 (Independent Consideration)**.** For each $a \in \mathcal{A}$ and $\mathbf{y} \in \mathcal{Y}^A$, the probability of facing consideration set $\mathcal{C}$, which is a subset of menu $\mathcal{Y}$, is

$$C_a(\mathcal{C} \mid \mathbf{y}, \mathcal{NC}_a, \mathcal{Y}) = \prod_{v \in \mathcal{C}} Q_a(v \mid \mathbf{y}, \mathcal{NC}_a) \prod_{v \in \mathcal{Y} \setminus \mathcal{C}} (1 - Q_a(v \mid \mathbf{y}, \mathcal{NC}_a)).$$

---

[9] That is, each Agent a is endowed with a collection of random variables $\{\tau_n^a\}_{n=1}^{\infty}$ such that $\tau_n^a - \tau_{n-1}^a$ is exponentially distributed with mean $1/\lambda_a$. These differences are independent across agents and time. All the identification results in Section 3.1 are still valid if agents have perfectly synchronized clocks.

Since the consideration set cannot be empty, we assume that the default alternative is always considered. That is, $Q_a(0 \mid \mathbf{y}, \mathcal{NC}_a) = 1$ for all $a \in \mathcal{A}$ and $\mathbf{y} \in \mathcal{Y}^A$. This restriction is the only one imposed on the default alternative, and it is satisfied in many applications (including the one we present in Section 5). We include $\mathcal{Y}$ as a determinant of C_a to simplify the notation when we later identify counterfactual choice distributions for different menus.

Example 1 (continued)**.** Recall that the consideration peers of Agent 1 are $\mathcal{NC}_1 = \{2, 3\}$. Thus, the probability that Agent 1 considers at the moment of choosing, say, set $\{0, 1\}$ is

$$C_1(\{0, 1\} \mid \mathbf{y}, \{2, 3\}, \{0, 1, 2\}) = Q_1(1 \mid \mathbf{y}, \{2, 3\})[1 - Q_1(2 \mid \mathbf{y}, \{2, 3\})].$$

Assume that the attention to option v given by Agent 1 is modeled as follows: v is considered if and only if $c_{1,v}(\mathbf{y}, \{2, 3\}) \geq \varepsilon_{1,v}$, where $c_{1,v}(\mathbf{y}, \mathcal{NC}_1)$ measures the mean attention of Agent 1 to v as a function of her consideration peers and $\varepsilon_{1,v}$ is a random attention shock independent of $\mathbf{y}$. Then, the probability of considering v is $Q_1(v \mid \mathbf{y}, \{2, 3\}) = F_{\varepsilon_{1,v}}(c_{v,1}(\mathbf{y}, \{2, 3\}))$, where $F_{\varepsilon_{1,v}}$ is the cumulative distribution function (c.d.f.) of $\varepsilon_{1,v}$. $\qquad\square$

Peer Effect in Preferences After the consideration set is formed, the agent selects an alternative according to some choice rule. The choices made by the agent from any consideration set may be random from the researcher's perspective if there are latent preference shocks across choice instances, as in the example below. Choice rules summarize the decision process after the consideration set is formed. Since, in practice, the underlying preferences (utility function parameters) can be identified and estimated from the choice rule, we focus on the choice rule and leave its association with preferences to be flexible.

We incorporate the peer effect in preferences by allowing the choice rule of agents to depend on the configuration of choices and her preference peers. Formally, given consideration set $\mathcal{C}$, the choice rule $R_a(\cdot \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C})$ is a discrete probability distribution supported on $\mathcal{C}$. That is, $R_a(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C}) \geq 0$ for all v, and $\sum_{v \in \mathcal{C}} R_a(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C}) = 1$.

Example 1 (continued)**.** Recall the preference peer of Agent 1 is $\mathcal{NR}_1 = \{3\}$. Hence, her probability of selecting alternative 1 in consideration set $\{0, 1\}$ is $R_1(1 \mid \mathbf{y}, \{3\}, \{0, 1\})$. Assume the utility Agent 1 gets from v in set $\mathcal{C}$ is given by $u_{1,v,\mathcal{C}}(\mathbf{y}, \mathcal{NR}_1) + \xi_{1,v,\mathcal{C}}$, where

10

$u_{1,v,\mathcal{C}}$ captures the mean utility from the alternative in the given consideration set. The vector of agent- and consideration-set-specific taste shocks $\xi_{a,\mathcal{C}} = (\xi_{a,v,\mathcal{C}})_{v \in \mathcal{C}}$ is continuously distributed[10] with the conditional c.d.f. $F_{a,\xi}(\cdot \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C})$. Then, for $v \in \mathcal{C}$,

$$\mathrm{R}_1\left(v \mid \mathbf{y}, \{3\}, \mathcal{C}\right) = \int \mathbb{1}\left(v = \arg\max_{v' \in \mathcal{C}}\{u_{1,v',\mathcal{C}}\left(\mathbf{y}, \{3\}\right) + \xi_{1,v',\mathcal{C}}\}\right) dF_{1,\xi}(\xi_{1,\mathcal{C}} \mid \mathbf{y}, \{3\}, \mathcal{C}),$$

where $\mathbb{1}\left(\cdot\right)$ is the indicator function. If the $\xi_{v,a,\mathcal{C}}$s are independent and identically distributed (i.i.d.) shocks, distributed according to the standard Type I extreme value distribution, then the above expressions simplify to

$$\mathrm{R}_1\left(v \mid \mathbf{y}, \{3\}, \mathcal{C}\right) = \frac{\exp\left(u_{1,v,\mathcal{C}}\left(\mathbf{y}, \{3\}\right)\right)}{\sum_{v' \in \mathcal{C}} \exp\left(u_{1,v',\mathcal{C}}\left(\mathbf{y}, \{3\}\right)\right)}. \qquad \square$$

Note that in the above example, consideration sets and choices of preference peers can directly affect mean utilities and the distribution of the latent taste shocks. Since the choices of consideration-only peers cannot affect the distribution of utilities conditional on consideration sets, in contrast to Barseghyan et al. (2021a) and Lu (2022), our model imposes some restrictions on the dependence between random preferences and consideration.[11]

We extend the model in Section 4.1 to allow the dependence of Q_a and R_a on the current or past choices made by Agent a (e.g., a Markov process with memory). This allows the possibility that an agent considers for sure (i.e., with probability 1) her current option in the next revision. We write the model in a stricter way here only to simplify the exposition.

By combining preferences and consideration sets, the conditional choice (ex-ante) probability (CCP) that Agent a selects (at the time of choosing) alternative v given $\mathbf{y}$ is

$$\mathrm{P}_a\left(v \mid \mathbf{y}\right) = \sum_{\mathcal{C} \subseteq \mathcal{Y}} \mathrm{R}_a\left(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C}\right) \prod_{v' \in \mathcal{C}} \mathrm{Q}_a\left(v' \mid \mathbf{y}, \mathcal{NC}_a\right) \prod_{v' \in \mathcal{Y} \setminus \mathcal{C}} \left(1 - \mathrm{Q}_a\left(v' \mid \mathbf{y}, \mathcal{NC}_a\right)\right).$$

We aim to identify $\mathcal{NR}_a$, $\mathcal{NC}_a$, R_a, and Q_a from a sequence of choices over time.

Remark 1. Our identification arguments *only* use variation in the choices made by connected agents. That is, we do not use exogenous variation in observable characteristics or menus.

---

[10]Continuity is only needed to handle cases where two alternatives give the same utility.

[11]Aguiar and Kashaev (2024) allow for a similar form of dependence of mean utilities on consideration.

Thus, to simplify the exposition, we do not include observable covariates in the model. We can interpret our setting as if we were conditioning on them. We show in our application that covariates can be easily incorporated into the model for estimation purposes and could also offer extra sources of identification. In particular, alternative-specific covariates can serve as further exclusion restrictions for the consideration or preferences —e.g., product-specific advertisement might only affect attention to a specific product (Goeree, 2008).

Remark 2. The dynamic interaction process we model assumes that each agent best responds to the observed choices made by others and does not anticipate their actions in the future or how her choice could affect them. Allowing for these possibilities would require a different interpretation of our model. For instance, an agent could select an option so that others incorporate the alternative in their consideration sets. We believe this set up could lead to a new model of endogenous social norms or rules within a group of people.

Remark 3. Our framework does not allow the dependence of current consideration on the previously considered (but not picked) alternatives. Allowing this invalidates our exclusion restrictions. Specifically, the introduction of the dependence of the consideration on past consideration leads to a hidden state Markov process, where the hidden state is the past consideration set. Since the past consideration sets and the past choices are correlated, and the past choices depend on preference peers, we obtain that the past consideration sets are correlated with choices made by preference peers. As a result, the choices made by *all* peers affect consideration. We leave this important and hard problem for future research.

2.2. Main Assumptions

Our results for equilibrium existence and identification build on four main assumptions. We have already discussed Assumption 1. We introduce next the other three main restrictions. Let $\mathrm{NC}_a^v(\mathbf{y})$ and $\mathrm{NR}_a^v(\mathbf{y})$ be the number of agents in the consideration and preference groups of Agent a who select option v in choice configuration $\mathbf{y}$, respectively. Formally,

$$\mathrm{NC}_a^v(\mathbf{y}) = |\{a' \in \mathcal{NC}_a \ : \ y_{a'} = v\}| \text{ and } \mathrm{NR}_a^v(\mathbf{y}) = |\{a' \in \mathcal{NR}_a \ : \ y_{a'} = v\}|,$$

where $|A|$ is the cardinality of A. Let $\text{NR}_a^{\mathcal{S}}(\mathbf{y}) = (\text{NR}_a^v(\mathbf{y}))_{v \in \mathcal{S} \setminus \{0\}}$ and $\text{NC}_a^{\mathcal{S}}(\mathbf{y}) = (\text{NC}_a^v(\mathbf{y}))_{v \in \mathcal{S} \setminus \{0\}}$ for any $\mathcal{S} \subseteq \mathcal{Y}$. The second assumption is as follows.

Assumption 2 (Consideration). For each $a \in \mathcal{A}$, $\mathbf{y} \in \mathcal{Y}^A$, and $v \neq 0$, we have that

(i) $Q_a(v \mid \mathbf{y}, \mathcal{NC}_a) > 0$;

(ii) $Q_a(v \mid \mathbf{y}, \mathcal{NC}_a) \equiv Q_a(v \mid \text{NC}_a^v(\mathbf{y}))$; and

(iii) Under Assumption 2(ii), $Q_a(v \mid 1) / Q_a(v \mid 0)$ differs from 1 and $Q_a(v \mid 2) / Q_a(v \mid 1)$.

Assumption 2(i) states that the probability of considering each option is strictly positive regardless of how many peers have selected it. This assumption captures the idea that agents can eventually pay attention to an alternative for various reasons that are outside the control of our model (e.g., watching an ad on television or receiving a coupon). We allow alternatives to be considered with probability 1,[12] capturing a form of persistence in consideration sets. We allow even richer forms of evolution of consideration sets by introducing history dependence to the model in Section 4.1. Assumption 2(ii) says that the probability of considering a specific option depends on the number but not the identity of the consideration peers that currently selected it. Assumption 2(iii) is a variability requirement stating that the choices made by consideration peers have an effect on consideration probabilities. It rules out constant or exponential probability functions of the number of peers (i.e., $\ln Q_a(v \mid \cdot)$ is nonlinear) around the origin. This restriction could be imposed at any other point in the support. This assumption allows for different levels of satiation. For example, consideration may change only when the number of peers picking the option achieves a given threshold (e.g., 10 agents, 20 agents, etc.). Assumption 2(iii) is not fully needed for all our results. For instance, nonexponential probabilities are needed to identify $\mathcal{NC}_a$, but $\mathcal{N}_a$ and $\mathcal{NR}_a$ only require consideration probabilities to vary with the choices made by the peers of Agent a.

Example 1 (continued). Suppose that the mean attention of Agent 1 towards alternative v, $c_{1,v}$, is such that $c_{1,v}(\mathbf{y}, \{2,3\}) = \text{NC}_1^v(\mathbf{y})$ with $\text{NC}_1^v(\mathbf{y}) = \mathbb{1}(y_2 = v) + \mathbb{1}(y_3 = v)$. Thus, $Q_1(v \mid \text{NC}_1^v(\mathbf{y})) = F_{\varepsilon_{1,v}}(\mathbb{1}(y_2 = v) + \mathbb{1}(y_3 = v))$. Assumption 2 holds if, for instance,

---

[12]Note that, due to its multiplicative structure, the ex-ante probability of facing a given consideration set is between 0 and 1 and adds up to 1 when we sum across all consideration sets.

the probability of considering an option increases with the number of consideration peers that select that alternative at a decreasing rate, e.g., $F_{\varepsilon_{1,v}}(0) = 1/8$, $F_{\varepsilon_{1,v}}(1) = 1/2$, and $F_{\varepsilon_{1,v}}(2) = 3/4$. It also holds if the rate of change is initially increasing, as in the well-known S-shaped curve for network effects, where a phase of increasing returns is then followed by diminishing returns. It only rules out a constant rate of change at some point. □

Let $\mathbf{0}_v^1$ denote a vector obtained by replacing the v-th component of the zero vector $\mathbf{0}$ by 1. The third assumption restricts the preference part of the decision process.

Assumption 3 (Preferences)**.** For each $a \in \mathcal{A}$, $\mathbf{y} \in \mathcal{Y}^A$, $\mathcal{C} \subseteq \mathcal{Y}$, and $v \in \mathcal{C} \setminus \{0\}$,

(i) $R_a(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C}^*) > 0$ for some $\mathcal{C}^*$ such that $C_a(\mathcal{C}^*|\mathbf{y}, \mathcal{NC}_a, \mathcal{Y}) > 0$;

(ii) $R_a(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C}) \equiv R_a\left(v \mid NR_a^{\mathcal{C}}(\mathbf{y}), \mathcal{C}\right)$; and

(iii) $R_a(v \mid \mathbf{0}_v^1, \mathcal{C}) - R_a(v \mid \mathbf{0}, \mathcal{C}) \neq 0$, and its sign does not depend on $\mathcal{C}$.

Assumption 3(i) requires each alternative to be picked with a positive probability at least in one consideration set that can be realized. Together with Assumption 2(i), it implies that every alternative can be picked with a positive probability. This assumption allows for both random and deterministic choice rules. Assumption 3(ii) states that the choice rule depends on the number (but not the identity) of preference peers that selected each of the alternatives in the consideration set. Assumption 3(iii) assumes that the change in the probability of selecting a given alternative due to an additional preference peer selecting this alternative is either positive or negative for all consideration sets that contain the alternative. The effect is required to be strict only around the origin, that is, when all other peers select the default. It rules out that the different changes induced by a preference peer changing her option cancel out. Though this condition can be relaxed, as written, it is easy to interpret and is weaker than assuming either positive or negative peer effects in preferences —as in many existing models. As in Assumption 2(iii), the direction of the peer effect in preferences does not need to be known and can be different for different agents and alternatives.

Example 1 (continued)**.** Suppose that the mean utility of Agent 1 from alternative v given consideration set $\mathcal{C}$ is $u_{1,v,\mathcal{C}}(\mathbf{y}, \{3\}) = \bar{u}_{1,v,\mathcal{C}}(NR_1^v(\mathbf{y}))$ with $NR_1^v(\mathbf{y}) = \mathbb{1}(y_3 = v)$. Thus,

given that only Agent 3 affects the preferences of Agent 1, the choice rule is expressed as

$$R_1\left(v \mid NR_1^{\mathcal{C}}(\mathbf{y}), \mathcal{C}\right) = \frac{\exp\left(\bar{u}_{1,v,\mathcal{C}}\left(\mathbb{1}\left(y_3 = v\right)\right)\right)}{\sum_{v' \in \mathcal{C}} \exp\left(\bar{u}_{1,v',\mathcal{C}}\left(\mathbb{1}\left(y_3 = v'\right)\right)\right)}.$$

Assumption 3 holds if $\bar{u}_{1,v,\mathcal{C}}(1) > \bar{u}_{1,v,\mathcal{C}}(0)$ or $\bar{u}_{1,v,\mathcal{C}}(1) < \bar{u}_{1,v,\mathcal{C}}(0)$ for all $\mathcal{C}$. This requires peer effects in preferences to be either positive or negative for each v and all $\mathcal{C}$. □

The fourth assumption imposes some restrictions on the network of each person.

Assumption 4 (Network)**.** For each a, if $|\mathcal{NCR}_a| \geq 1$ then $|\mathcal{NC}_a \setminus \mathcal{NR}_a| + |\mathcal{NR}_a \setminus \mathcal{NC}_a| \geq 1$.

Assumption 4 is an exclusion restriction. It states that if the agent has a peer that affects both consideration and preferences, then the agent also has at least another peer that affects either only consideration or only preferences. The choice made by such a peer provides an exclusion restriction. Without Assumption 4 the scenario where $\mathcal{NC}_a = \mathcal{NR}_a$ is observationally equivalent to the one in which $\mathcal{NC}_a = \emptyset$. Note that we do not need the two sets of agents to be nonempty. If only one of the peer-effect mechanisms operates in practice, our results will allow us to state whether the interdependencies in choices are due to peer effect in preferences or in consideration.

Example 1 (continued)**.** Agent 1 is the only agent for whom $|\mathcal{NCR}_1| = 1 \geq 1$. But we also have that $|\mathcal{NC}_1 \setminus \mathcal{NR}_1| + |\mathcal{NR}_1 \setminus \mathcal{NC}_1| = 1 \geq 1$. Thus, Assumption 4 is satisfied. □

2.3. Equilibrium Behavior

In this subsection, we define a notion of equilibrium in the network system, i.e., the invariant distribution in the Markov process, and establish its existence and uniqueness.

The i.i.d. Poisson alarm clocks, which determine the revision process, guarantee that each time, at most, one agent revises her selection almost surely. Thus, the transition rates between choice configurations that differ in more than one agent changing the current selection are zero. This facilitates identification as there are fewer terms to recover (Blevins, 2017, 2018). It also rules out some mechanisms for multiple equilibria in the data-generating

15

process. Formally, the transition rate from choice configuration $\mathbf{y}$ to any different one $\mathbf{y}'$ is

$$\mathrm{w}\left(\mathbf{y}' \mid \mathbf{y}\right) = \begin{cases} 0 & \text{if } \sum_{a \in \mathcal{A}} \mathbb{1}\left(y'_a \neq y_a\right) > 1 \\ \sum_{a \in \mathcal{A}} \lambda_a \, \mathrm{P}_a\left(y'_a \mid \mathbf{y}\right) \mathbb{1}\left(y'_a \neq y_a\right) & \text{if } \sum_{a \in \mathcal{A}} \mathbb{1}\left(y'_a \neq y_a\right) = 1 \end{cases}.$$

In the statistical literature on continuous-time Markov processes, these transition rates are the off-diagonal terms of the *transition rate matrix* (also known as the *infinitesimal generator matrix*). The diagonal terms are simply given by $\mathrm{w}\left(\mathbf{y} \mid \mathbf{y}\right) = -\sum_{\mathbf{y}' \in \mathcal{Y}^A \setminus \{\mathbf{y}\}} \mathrm{w}\left(\mathbf{y}' \mid \mathbf{y}\right)$.

We indicate by $\mathcal{W}$ the transition rate matrix. In our model, the number of possible choice configurations is $(Y+1)^A$. Thus, $\mathcal{W}$ is a $(Y+1)^A \times (Y+1)^A$ matrix. To avoid ambiguity in the exposition, we let the choice configurations be ordered according to the lexicographic order. Formally, let $\iota\left(\mathbf{y}\right) \in \left\{1, 2, \ldots, (Y+1)^A\right\}$ be the position of $\mathbf{y}$ according to the lexicographic order. Then, $\mathcal{W}_{\iota(\mathbf{y})\iota(\mathbf{y}')} = \mathrm{w}\left(\mathbf{y}' \mid \mathbf{y}\right)$.

The system in equilibrium is characterized by an invariant distribution $\mu : \mathcal{Y}^A \to (0,1)$, with $\sum_{\mathbf{y} \in \mathcal{Y}^A} \mu\left(\mathbf{y}\right) = 1$, of the dynamic process with transition rate matrix $\mathcal{W}$. This equilibrium behavior relates to the transition rate matrix in a linear fashion $\mu \mathcal{W} = \mathbf{0}$.

The next proposition shows equilibrium existence. It also states that the same conditions guarantee the equilibrium is unique and has full support on the choice configurations. The full support feature is important for us as we rely on choice variation to recover the model.

Proposition 2.1. *If Assumptions 1, 2(i), and 3(i) hold, then there exists a unique equilibrium μ with full support.*

3. Empirical Content of the Model

This section offers restrictions under which the researcher can uniquely recover the set of connections, $\mathcal{NC}_a$ and $\mathcal{NR}_a$, the consideration mechanism, Q_a, the choice rule, R_a, and the Poisson alarm clock, λ_a, for every Agent a. We separate the identification analysis in two parts. We first show how to recover the set of connections, choice rules, and consideration probabilities from the CCPs, $\mathrm{P} = (\mathrm{P}_a)_{a \in \mathcal{A}}$. We then show how to recover P.

Remark 4. In Online Appendix B, we run simulations from our motivating example to show that we can estimate CCPs and parts of the model nonparametrically using our identification ideas step-by-step. We also show that estimation becomes harder with more agents and/or alternatives. In the application, we use a parametric estimator to circumvent these difficulties.

3.1. Identification of the Model From P

The identification strategy we offer builds on a sequence of steps. We initially recover the network structure in three stages: First, we recover the reference group of every agent. Second, we identify whether a given peer affects consideration only or preferences. Lastly, we show how to distinguish between a peer that affects preferences only (preference-only peer) and a peer that affects consideration and preferences (consideration-preference peer). We finally use this information to recover consideration probabilities and choice rules.

Along the analysis we assume a mild "regularity condition." We briefly discuss the role of this condition below and formalize it in Online Appendix A.

Network The key idea in recovering the network is that changes in the choices made by the peers of a given agent induce variation in her CCPs *and* that different types of peers induce a different type of variation. To see this, note that P_a can be rewritten as

$$\mathrm{P}_a\left(v \mid \mathbf{y}\right) = \mathrm{Q}_a\left(v \mid \mathrm{NC}_a^v\left(\mathbf{y}\right)\right) \times \mathrm{D}_a\left(v \mid \mathrm{NR}_a^{\mathcal{Y}}\left(\mathbf{y}\right), \mathrm{NC}_a^{\mathcal{Y}\setminus\{v\}}\left(\mathbf{y}\right)\right),$$

where the second term in the right-hand side is given by

$$\mathrm{D}_a\left(v \mid \mathrm{NR}_a^{\mathcal{Y}}\left(\mathbf{y}\right), \mathrm{NC}_a^{\mathcal{Y}\setminus\{v\}}\left(\mathbf{y}\right)\right) = \sum\nolimits_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}_a^{\mathcal{C}\cup\{v\}}\left(\mathbf{y}\right), \mathcal{C}\cup\{v\}\right) \mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}_a^{\mathcal{Y}\setminus\{v\}}(\mathbf{y}), \mathcal{Y}\setminus\{v\}\right).$$

In words, the observed probability that v is picked equals the product of the probabilities that it is considered and that it is picked when considered. Note that consideration-only peers who select option v enter only the first term. In addition, consideration and/or preference peers who select alternatives other than v affect only the second term. These two observations allow us to use certain changes in $\ln \mathrm{P}_a$ to identify the network.

Define $\Delta_{a'}^v$ as a linear operator that indicates the increment of a function when the choice

of Agent a' in $\mathbf{y}$ changes to v. Formally, given $f : \mathcal{Y}^A \to \mathbb{R}$, let $\Delta^v_{a'} f(\mathbf{y}) = f(\mathbf{y}^v_{a'}) - f(\mathbf{y})$, where $\mathbf{y}^v_{a'}$ denotes the vector in which the a'-th component of $\mathbf{y}$ is replaced by v.

We first identify the reference group of Agent a by using changes in her CCPs. Intuitively, Agent a' is in the reference group of Agent a if changing her choice in the choice configuration affects the decision of Agent a. Specifically, the implied difference in $\ln \mathrm{P}_a$ is given by

$$\Delta^v_{a'} \ln \mathrm{P}_a (v \mid \mathbf{0}) = \Delta^v_{a'} \ln \mathrm{Q}_a (v \mid \mathrm{NC}^v_a (\mathbf{0})) + \Delta^v_{a'} \ln \mathrm{D}_a \left(v \mid \mathrm{NR}^{\mathcal{Y}}_a (\mathbf{0}), \mathrm{NC}^{\mathcal{Y} \backslash \{v\}}_a (\mathbf{0}) \right), \quad (1)$$

where the zero vector $\mathbf{0}$ denotes the configuration where everyone picks the default. Each term in Equation (1) relates to one mechanism of peer effects: the first term reflects (if present) the peer effect in consideration. The second term captures the peer effect in preferences.

When peer effects in consideration and preferences are of the same sign, then, under Assumptions 1 - 3, $\Delta^v_{a'} \ln \mathrm{P}_a (v \mid \mathbf{0}) \neq 0$ if and only if Agent a' is in the reference group of Agent a (i.e., $a' \in \mathcal{N}_a$). When the interaction effects are of different signs, the "if" part requires a "regularity condition" that we discuss in detail in Online Appendix A. This extra condition rules out the possibility that peer effects in consideration and preferences be of opposite signs *and* of equal magnitude. We next state that under all our restrictions the reference groups (even if they are empty) can be recovered from the CCPs.

Proposition 3.1. *If Assumptions 1 - 3 hold, then $\mathcal{N}_a$ is identified for each $a \in \mathcal{A}$.*

Example 1 (continued)**.** The probability that Agent 1 selects option 1 satisfies

$$\ln \mathrm{P}_1 (1 \mid \mathbf{y}) = \ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}^1_1 = \mathbb{1} \left(y_2 = 1 \right) + \mathbb{1} \left(y_3 = 1 \right) \right)$$
$$+ \ln \mathrm{D}_1 \left(1 \mid \mathrm{NR}^{\{1,2\}}_1 = \left(\mathbb{1} \left(y_3 = 1 \right), \mathbb{1} \left(y_3 = 2 \right) \right), \mathrm{NC}^2_1 = \mathbb{1} \left(y_2 = 2 \right) + \mathbb{1} \left(y_3 = 2 \right) \right).$$

Take $\mathbf{y} = \mathbf{0}$. The changes in $\ln \mathrm{P}_1$ when we change the other agents choices from 0 to 1 are

$$\Delta^1_2 \ln \mathrm{P}_1 (1 \mid \mathbf{0}) = \ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}^1_1 = 1 \right) - \ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}^1_1 = 0 \right) \neq 0,$$
$$\Delta^1_3 \ln \mathrm{P}_1 (1 \mid \mathbf{0}) = \ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}^1_1 = 1 \right) - \ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}^1_1 = 0 \right)$$
$$+ \ln \mathrm{D}_1 \left(1 \mid \mathrm{NR}^{\{1,2\}}_1 = (1,0), \mathrm{NC}^2_1 = 0 \right) - \ln \mathrm{D}_1 \left(1 \mid \mathrm{NR}^{\{1,2\}}_1 = (0,0), \mathrm{NC}^2_1 = 0 \right) \neq 0,$$
$$\Delta^1_4 \ln \mathrm{P}_1 (1 \mid \mathbf{0}) = 0.$$

Proposition 3.1 identifies the peers of Agent 1 as those for which these changes differ from 0. Following this idea, we correctly recover $\mathcal{N}_1 = \{2, 3\}$. $\qquad\square$

Next, we identify whether Agent a' in $\mathcal{N}_a$ affects Agent a's preferences or consideration only. Note that differences in $\ln \mathrm{P}_a$ allow us to recover the reference groups, but these differences are silent about the mechanism by which the interactions happen. To see why, note that (for instance) a nonzero $\Delta_{a'}^v \ln \mathrm{P}_a (v \mid \mathbf{0})$ could be generated from the first summand in Equation (1) via Q_a and/or from the second summand via R_a. But these two terms differ in that the second summand varies with the number of peers that select alternatives that are *different* from v, while the first term does not. Thus, the two mechanisms can be set apart by a second shift in $\ln \mathrm{P}_a (v \mid \mathbf{0})$. Let $a', a'' \in \mathcal{N}_a$ and $w \in \mathcal{Y} \setminus \{0\}$ with $w \neq v$. Since Δ_a^v is a linear operator, we can define a double difference as follows

$$\Delta_{a''}^w \Delta_{a'}^v \ln \mathrm{P}_a (v \mid \mathbf{0}) \equiv \Delta_{a''}^w \left[\ln \mathrm{P}_a (v \mid \mathbf{0}_{a'}^v) - \ln \mathrm{P}_a (v \mid \mathbf{0}) \right] = \Delta_{a''}^w \ln \mathrm{P}_a (v \mid \mathbf{0}_{a'}^v) - \Delta_{a''}^w \ln \mathrm{P}_a (v \mid \mathbf{0})$$
$$= \left[\ln \mathrm{P}_a (v \mid (\mathbf{0}_{a'}^v)_{a''}^w) - \ln \mathrm{P}_a (v \mid \mathbf{0}_{a'}^v) \right] - \left[\ln \mathrm{P}_a (v \mid \mathbf{0}_{a''}^w) - \ln \mathrm{P}_a (v \mid \mathbf{0}) \right].$$

Specifically, we have that

$$\Delta_{a''}^w \Delta_{a'}^v \ln \mathrm{P}_a (v \mid \mathbf{0}) = \Delta_{a''}^w \Delta_{a'}^v \ln \mathrm{Q}_a (v \mid \mathrm{NC}_a^v (\mathbf{0})) + \Delta_{a''}^w \Delta_{a'}^v \mathrm{D}_a \left(v \mid \mathrm{NR}_a^{\mathcal{Y}} (\mathbf{0}), \mathrm{NC}_a^{\mathcal{Y} \setminus \{v\}} (\mathbf{0}) \right).$$

Note that $\Delta_{a''}^w \Delta_{a'}^v \ln \mathrm{Q}_a (v \mid \mathrm{NC}_a^v (\mathbf{0})) = 0$ since $\mathrm{Q}_a (v \mid \cdot)$ does not depend on the number of peers who picked w. Also, if a' is a consideration-only peer ($a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$), then

$$\Delta_{a'}^v \mathrm{D}_a \left(v \mid \mathrm{NR}_a^{\mathcal{Y}} (\mathbf{0}), \mathrm{NC}_a^{\mathcal{Y} \setminus \{v\}} (\mathbf{0}) \right) = 0.$$

As a result, $\Delta_{a''}^w \Delta_{a'}^v \ln \mathrm{P}_a (v \mid \mathbf{0}) = 0$ if Agent a' is a consideration-only peer. A key observation is that if Agent a' affects Agent a's preferences, then the second summand in Equation (1) will not disappear after switching Agent a'' from 0 to w. In summary, under Assumptions 2-3 and the regularity condition, $a' \in \mathcal{NR}_a$ if and only if

$$\Delta_{a''}^w \Delta_{a'}^v \ln \mathrm{P}_a (v \mid \mathbf{0}) \neq 0 \text{ for some } a'' \in \mathcal{N}_a.$$

Thus, by checking the double difference for each agent in the reference group of Agent a, we can divide her reference group into consideration-only peers and preference peers (who may or may not affect consideration). This identification strategy requires $Y \geq 2$ and $|\mathcal{N}_a| \geq 2$.

Note that we can also separate the preference peers in two sets by the magnitude of the changes in CCPs. Specifically, for $a' \in \mathcal{NCR}_a$ and $a'' \in \mathcal{NR}_a \setminus \mathcal{NC}_a$, we have that

$$\Delta_{a'}^v \ln \mathrm{P}_a \left(v \mid \mathbf{0} \right) \neq \Delta_{a''}^v \ln \mathrm{P}_a \left(v \mid \mathbf{0} \right).$$

This allows us to separate the preference peers into two groups that we define as $\mathcal{M}'$ and $\mathcal{M}''$. Although we know that one of these sets is $\mathcal{NCR}_a$ and the other is $\mathcal{NR}_a \setminus \mathcal{NC}_a$, without further restrictions, we cannot tell which is which. We address this issue in the end.

Proposition 3.2. *Suppose Assumptions 1 - 3 hold. For any $a \in \mathcal{A}$, if $Y \geq 2$ and $|\mathcal{N}_a| \geq 2$, then $\mathcal{NC}_a \setminus \mathcal{NR}_a$ and $\mathcal{NR}_a = \mathcal{M}' \cup \mathcal{M}''$ are identified. Also, $\mathcal{NCR}_a \in \{\mathcal{M}', \mathcal{M}''\}$.*

Example 1 (continued)**.** Recall that we identified that $\mathcal{N}_1 = \{2, 3\}$. We now want to know what type of peers Agents 2 and 3 are. Next we display two double differences

$$\Delta_3^2 \Delta_2^1 \ln \mathrm{P}_1 \left(1 \mid \mathbf{0} \right) = 0,$$
$$\Delta_2^2 \Delta_3^1 \ln \mathrm{P}_1 \left(1 \mid \mathbf{0} \right) = \left[\ln \mathrm{D}_1 \left(1 \mid \mathrm{NR}_1^{\{1,2\}} = (1, 0), \mathrm{NC}_1^2 = 1 \right) - \ln \mathrm{D}_1 \left(1 \mid \mathrm{NR}_1^{\{1,2\}} = (0, 0), \mathrm{NC}_1^2 = 1 \right) \right]$$
$$- \left[\ln \mathrm{D}_1 \left(1 \mid \mathrm{NR}_1^{\{1,2\}} = (1, 0), \mathrm{NC}_1^2 = 0 \right) - \ln \mathrm{D}_1 \left(1 \mid \mathrm{NR}_1^{\{1,2\}} = (0, 0), \mathrm{NC}_1^2 = 0 \right) \right] \neq 0.$$

In the first line, we first change the choice of Agent 2 from 0 to 1. Since Agent 2 is a consideration-only peer, this difference does not depend on the number of other peers selecting options different from 1. Thus, when we further change the choice of Agent 3 from 0 to 2, the result is 0, and we conclude that Agent 2 is a consideration-only peer.

In the second line, we first change the choice of Agent 3 from 0 to 1. Since Agent 3 is a consideration-preference peer of Agent 1, this difference depends on the choices made by other peers selecting other alternatives. Thus, when we further change the choice of Agent 2 from 0 to 2, the result differs from 0, and we conclude that Agent 3 is a preference peer. $\square$

Finally, we identify the set of consideration-preference peers (i.e., $\mathcal{NCR}_a$) from the group of peers that affect preferences. We discuss identification with and without consideration-only

peers separately. By Assumption 4, if $\mathcal{NCR}_a$ is nonempty, then there exists a peer that is either a consideration-only or preference-only peer. Assume that we have already identified two peers such that Agent a' is a consideration-only peer (i.e., $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$) and Agent a'' affects preferences (i.e., $a'' \in \mathcal{NR}_a$). Note that

$$\Delta_{a''}^v \Delta_{a'}^v \ln \mathrm{P}_a \left(v \mid \mathbf{0} \right) = \Delta_{a''}^v \Delta_{a'}^v \ln \mathrm{Q}_a \left(v \mid \mathrm{NC}_a^v \left(\mathbf{0} \right) \right).$$

This is so because, since Agent a' only affects consideration, the second term in Equation (1) is zero. Thus, if Assumption 2(iii) holds, $a'' \in \mathcal{NCR}_a$ if and only if $\Delta_{a''}^v \Delta_{a'}^v \ln \mathrm{P}_a \left(v \mid \mathbf{0} \right) \neq 0$.

Suppose next that there is no consideration-only peer. We can implement a similar idea by *replicating* the consideration-only peer behavior with a consideration-preference peer and a preference-only one. Note that these two peers can be identified by Proposition 3.2. Pick some Agent $a' \in \mathcal{M}'$ and Agent $a'' \in \mathcal{M}''$. We have that

$$\ln \mathrm{P}_a \left(v \mid \mathbf{0}_{a'}^v \right) - \ln \mathrm{P}_a \left(v \mid \mathbf{0}_{a''}^v \right) = (-1)^{\mathbb{1}\left(a' \notin \mathcal{NCR}_a \right)} \left(\ln \mathrm{Q}_a \left(v \mid 1 \right) - \ln \mathrm{Q}_a \left(v \mid 0 \right) \right).$$

That is, when we switch a consideration-preference peer from the default to alternative v, $\ln \mathrm{P}_a \left(v \mid \mathbf{0} \right)$ changes via two effects, namely, preference and consideration. Importantly, the effect via preferences coincides with the one we get by switching a preference-only peer from the default to alternative v. Thus, by subtracting the two effects, we can recover the effect (up to sign) of switching a consideration-only peer from the default to option v.

Finally, take another Agent a''' from either $\mathcal{M}'$ or $\mathcal{M}''$ and implement a double difference to the previous expression by changing the alternative of Agent a''' from the default to v. As before, we identify whether Agent a''' is a consideration-preference or preference-only peer by checking whether this double change is different from zero:[13]

$$\Delta_{a'''}^v \left[\ln \mathrm{P}_a \left(v \mid \mathbf{0}_{a'}^v \right) - \ln \mathrm{P}_a \left(v \mid \mathbf{0}_{a''}^v \right) \right] \neq 0 \iff a''' \in \mathcal{NCR}_a.$$

This information allows us to know whether $\mathcal{NCR}_a = \mathcal{M}'$ or $\mathcal{NCR}_a = \mathcal{M}''$.

The last proposition offers final conditions for all the parts of the network to be identified.

---

[13]This procedure requires at least three peers in $\mathcal{N}_a$.

Proposition 3.3. *Suppose Assumptions 1 - 4 hold. Suppose also that $\mathcal{NC}_a \backslash \mathcal{NR}_a$ is identified (or known) and $|\mathcal{N}_a| \geq 3 - |\mathcal{NC}_a \backslash \mathcal{NR}_a|$. Then, $\mathcal{NC}_a$ and $\mathcal{NR}_a$ are identified.*

Example 1 (continued)**.** We learned earlier that Agent 2 is a consideration-only peer of Agent 1. We also learned that Agent 3 is either a preference-only peer or affects both preferences and consideration of Agent 1. We next establish whether $\mathcal{NC}_1 = \{2\}$ or $\mathcal{NC}_1 = \{2, 3\}$. Since Agent 2 is a consideration-only peer, we have that

$$\Delta_2^1 \ln \mathrm{P}_1 (1 \mid \mathbf{0}) = \ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}_1^1 = 1 \right) - \ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}_1^1 = 0 \right).$$

Changing the alternative of Agent 3 from 0 to 1, we obtain that by Assumption 2(iii)

$$\begin{aligned}\Delta_3^1 \Delta_2^1 \ln \mathrm{P}_1 (1 \mid \mathbf{0}) &= \left[\ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}_1^1 = 2 \right) - \ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}_1^1 = 1 \right) \right] \\ &\quad - \left[\ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}_1^1 = 1 \right) - \ln \mathrm{Q}_1 \left(1 \mid \mathrm{NC}_1^1 = 0 \right) \right] \neq 0.\end{aligned}$$

Thus, we identify that Agent 3 is also a consideration peer. That is, $\mathcal{NC}_1 = \{2, 3\}$. $\qquad\square$

To sum up, the reference group of Agent a is identified by checking the variation in $\ln \mathrm{P}_a$ as we switch other agents from the default alternative to a specific v. If, in doing so, we identify that the agent has two or more peers, we can recover the consideration-only peers by using the additive separability of $\ln \mathrm{P}_a (v \mid \mathbf{y})$ in $\mathrm{Q}_a (v \mid \mathrm{NC}_a^v (\mathbf{y}))$. Finally, if we identify at least one consideration-only peer, we can use her as a baseline to identify all other types of peers. Otherwise, we *create* such a peer by mixing the behavior of a consideration-preference peer with the one of a preference-only peer and use the behavior of the *constructed* peer as a baseline to complete the network identification. In this case, we need at least three peers.

We finally remark that while no restriction on the number of options is needed to recover the reference group of a given agent, we assume $Y \geq 2$ to divide this set in consideration and preference peers. We next use the initial example to see why this requirement is needed and to state what can be done when it fails, i.e. $Y = 1$.

Example 1 (continued)**.** Let us keep the network but assume $\mathcal{Y} = \{0, 1\}$. Then

$$\ln \mathrm{P}_1 (1 \mid \mathbf{y}) = \ln \mathrm{Q}_1 \left(1 \mid \mathbb{1} \left(y_2 = 1 \right) + \mathbb{1} \left(y_3 = 1 \right) \right) + \ln \mathrm{R}_1 \left(1 \mid \mathbb{1} \left(y_3 = 1 \right), \{0, 1\} \right).$$

In this set-up,

$$\Delta_2^1 \ln P_1 (1 \mid \mathbf{0}) = \ln Q_1 (1 \mid 1) - \ln Q_1 (1 \mid 0),$$

$$\Delta_3^1 \ln P_1 (1 \mid \mathbf{0}) = \ln Q_1 (1 \mid 1) - \ln Q_1 (1 \mid 0) + \ln R_1 (1 \mid 1, \{0, 1\}) - \ln R_1 (1 \mid 0, \{0, 1\}),$$

$$\Delta_4^1 \ln P_1 (1 \mid \mathbf{0}) = 0.$$

Thus, we can learn that $\mathcal{N}_1 = \{2, 3\}$. But now, double differences will not allow us to state whether each of these agents is a consideration and/or a preference peer. $\qquad\square$

The above example shows that our identification strategy fails to separate the type of peers when $Y = 1$. In this case, there are other sets of assumptions that we could invoke to restore identification. Among them, partial knowledge of the network structure and knowledge of the sign of peer effects would allow identification, as we illustrate next.

Example 1 (continued). Recall that $Y = 1$. Suppose we have partial knowledge of the network structure. In particular, suppose we know that $\mathcal{NR}_1 = \{3\}$. Since we can still learn that $\mathcal{N}_1 = \{2, 3\}$, we conclude that $\mathcal{NC}_1 \backslash \mathcal{NR}_1 = \{2\}$. Thus, to recover the complete network, we only need to learn whether Agent 3 is a preference-only peer or a consideration-preference peer of Agent 1. As before, the fact that

$$\begin{aligned}
\Delta_3^1 \Delta_2^1 \ln P_1 (1 \mid \mathbf{0}) = &\left[\ln Q_1 \left(1 \mid \mathrm{NC}_1^1 = 2 \right) - \ln Q_1 \left(1 \mid \mathrm{NC}_1^1 = 1 \right) \right] \\
&- \left[\ln Q_1 \left(1 \mid \mathrm{NC}_1^1 = 1 \right) - \ln Q_1 \left(1 \mid \mathrm{NC}_1^1 = 0 \right) \right] \neq 0
\end{aligned}$$

allows us to conclude that Agent 3 is a consideration-preference peer.

When $Y = 1$ we can also recover the network structure under sign restrictions, which we have not imposed so far. If we assume that peer effects in consideration and preferences are of opposite signs, then we could dispense with the assumption that either $\mathcal{NC}_a$ or $\mathcal{NR}_a$ is known. This situation might apply to vaccines. Arguably, a person becomes aware of a vaccine if more of her friends are getting shots. Also, if more friends get vaccinated, then the chances of getting sick reduce, and this reduces the incentives to get the vaccine. Thus, the peer effects in consideration and preferences are positive and negative, respectively.[14] $\square$

---

[14]In a different model, a similar idea has been used by Agranov et al. (2021) to explain some data on

Consideration Mechanisms and Choice Rules We first state that if we know the network structure, and each agent has at least one consideration-only peer —or such a peer can be constructed from consideration-preference and preference-only peers, as we do above— then we can recover ratios of consideration probabilities. To show this claim, let $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$. Since Agent a' only affects consideration, we can shift Agent a''s choice from the default to v and recover some information about the peer effect in consideration. Specifically, we have

$$\Delta_{a'}^v \ln \mathrm{P}_a(v \mid \mathbf{0}) = \ln \mathrm{Q}_a(v \mid 1) - \ln \mathrm{Q}_a(v \mid 0).$$

Thus, we can identify $\mathrm{Q}_a(v \mid 1) / \mathrm{Q}_a(v \mid 0)$. If $\mathcal{NC}_a \setminus \mathcal{NR}_a$ is empty, but $\mathcal{NR}_a \setminus \mathcal{NC}_a$ is not, we can use preference-only peers in a similar way. In particular, suppose $a' \in \mathcal{NCR}_a$ and $a'' \in \mathcal{NR}_a \setminus \mathcal{NC}_a$. Then, $\ln \mathrm{P}_a(v \mid \mathbf{0}_{a'}^v) - \ln \mathrm{P}_a(v \mid \mathbf{0}_{a''}^v) = \ln \mathrm{Q}_a(v \mid 1) - \ln \mathrm{Q}_a(v \mid 0)$. By applying the same ideas to different initial configurations, we can identify ratios of consideration probabilities as we formally state next.

Proposition 3.4. *Let $\mathcal{NC}_a$ and $\mathcal{NR}_a$ be known and Assumptions 1 - 4 hold. Then*

$$\mathrm{Q}_a(v \mid n+1) / \mathrm{Q}_a(v \mid n)$$

is identified from P_a for each n from 0 to $|\mathcal{NC}_a| - 1$. (We use the convention that if $|\mathcal{NC}_a| = 0$, then the set "from 0 to -1" is empty.)

Remark 5. Proposition 3.4 is valid for a substantially more general consideration set model. For example, the assumption that each alternative is added to the consideration sets independently from other alternatives (Assumption 1) can be dropped. Indeed, by definition, $\mathrm{P}_a(v \mid \mathbf{y}) = \mathrm{Q}_a(v \mid \mathrm{NC}_a^v(\mathbf{y})) \mathrm{Pr}_a(v \mid \mathbf{y}, v \text{ is considered})$, where the second term is the conditional probability that v is picked conditional on being considered. Thus, variation in the choices made by $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$ would identify Q_a up to scale. Note, however, that in this case, knowing Q_a is not enough to identify C_a since Q_a does not convey information about the probability of several items being considered simultaneously.

COVID-19 vaccine uptake.

We next show that we can also recover some counterfactual objects of interest. Adding some restrictions, these counterfactuals will allow us to recover the choice rules. Define

$$\mathrm{P}_a^* \left(v \mid \mathbf{y}, \mathcal{Y} \setminus \mathcal{Z} \right) = \sum_{\mathcal{C} \subseteq \mathcal{Y} \setminus \mathcal{Z}} \mathrm{R}_a \left(v \mid \mathrm{NR}_a^{\mathcal{C}} \left(\mathbf{y} \right), \mathcal{C} \right) \mathrm{C}_a \left(\mathcal{C} \mid \mathrm{NC}_a^{\mathcal{Y} \setminus \mathcal{Z}} (\mathbf{y}), \mathcal{Y} \setminus \mathcal{Z} \right)$$

for each $\mathcal{Z} \subseteq \mathcal{Y} \setminus \{0\}$. That is, $\mathrm{P}_a^* \left(v \mid \mathbf{y}, \mathcal{Y} \setminus \mathcal{Z} \right)$ is the counterfactual probability of selecting alternative v under choice configuration $\mathbf{y}$ when we restrict the set of available options or the menu from $\mathcal{Y}$ to $\mathcal{Y} \setminus \mathcal{Z}$. It tells us what happens to the CCPs when we remove set $\mathcal{Z}$ from the original menu. Note that, by definition, $\mathrm{P}_a^* \left(v \mid \mathbf{y}, \mathcal{Y} \right) = \mathrm{P}_a \left(v \mid \mathbf{y} \right)$.

To fix the ideas behind the next result, consider the setting with $\mathcal{A} = \{a, a'\}$, $\mathcal{Y} = \{0, v, v'\}$, $\mathcal{NR}_a = \emptyset$, and $\mathcal{NC}_a = \{a'\}$. Take $\mathbf{y}$ such that $y_{a'} = 0$ (y_a can be arbitrary). Recall that $\mathbf{y}_{a'}^{v'}$ denotes a configuration where the a'-th component of $\mathbf{y}$ is replaced by v'. Since

$$\mathrm{P}_a^*(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\}) = \mathrm{Q}_a \left(v \mid 0 \right) \mathrm{R}_a \left(v \mid 0, \{0, v\} \right),$$

we have that

$$\mathrm{P}_a \left(v \mid \mathbf{y} \right) = \mathrm{Q}_a \left(v' \mid 0 \right) \mathrm{Q}_a \left(v \mid 0 \right) \mathrm{R}_a \left(v \mid (0, 0), \{0, v, v'\} \right) + \left[1 - \mathrm{Q}_a \left(v' \mid 0 \right) \right] \mathrm{P}_a^*(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\}).$$

This is the observed probability of Agent a choosing option v given that her peer a' previously chose the default. Moreover, by switching a''s choice from the default to v', we have

$$\mathrm{P}_a \left(v \mid \mathbf{y}_{a'}^{v'} \right) = \mathrm{Q}_a \left(v' \mid 1 \right) \mathrm{Q}_a \left(v \mid 0 \right) \mathrm{R}_a \left(v \mid (0, 0), \{0, v, v'\} \right) + \left(1 - \mathrm{Q}_a \left(v' \mid 1 \right) \right) \mathrm{P}_a^*(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\}).$$

Note that we used the fact that since Agent a' only affects Agent a's consideration probability, but not the preference, the variation of Agent a''s choice in the choice configuration provides variation in the consideration probability but not in the choice rule. That is, $\mathrm{R}_a \left(v \mid (0, 0), \{0, v, v'\} \right)$ does not vary when a' switches from the default to a different alternative. Moreover, we also used the fact that $\mathrm{P}_a^*(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\}) = \mathrm{P}_a^*(v \mid \mathbf{y}_{a'}^{v'}, \mathcal{Y} \setminus \{v'\})$, which follows from v' being excluded from the menu and, thus, switching to it does not change the probability of picking v.

Solving this system of two equations with respect to $P_a^*(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\})$, we obtain that

$$P_a^*(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\}) = \frac{P_a\left(v \mid \mathbf{y}_{a'}^{v'}\right) - t_{v'}\,P_a\left(v \mid \mathbf{y}\right)}{1 - t_{v'}},$$

where $t_{v'} = Q_a\left(v' \mid 1\right) / Q_a\left(v' \mid 0\right) \neq 1$ can be identified using Proposition 3.4. It follows that we can recover the counterfactual CCP $P_a^*(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\})$ for any $\mathbf{y}$ for which the alternative corresponding to one of the consideration-only peers is equal to 0 (i.e., $y_{a'} = 0$). Essentially, we just used a consideration-only peer to exclude one alternative from the menu. Applying the same argument to these new counterfactual CCPs, we can exclude two alternatives as long as we have two consideration-only peers. Again, we can use any initial $\mathbf{y}$ as long as the components that correspond to any two consideration-only peers are set to 0. That is, we can exclude any set of nondefault alternatives if its cardinality is smaller than $|\mathcal{NC}_a \setminus \mathcal{NR}_a|$.

The next result formalizes and extends this argument.

Proposition 3.5. *Suppose $\mathcal{NC}_a$ and $\mathcal{NR}_a$ are known, and Assumptions 1 - 3 are satisfied. Then $P_a^*\left(v \mid \mathbf{y}, \mathcal{Y} \setminus \mathcal{Z}\right)$ is identified from P_a for every $\mathcal{Z} \subseteq \mathcal{Y} \setminus \{0\}$ such that $|\mathcal{Z}| \leq |\mathcal{NC}_a \setminus \mathcal{NR}_a|$ and each $\mathbf{y}$ for which at least $|\mathcal{Z}|$ of its components corresponding to any peers in $\mathcal{NC}_a \setminus \mathcal{NR}_a$ are 0.*

Proposition 3.5 addresses an important counterfactual prediction: What would happen if some alternatives were removed or become unavailable? Note the identification of these counterfactual CCPs does not require knowledge of either Q_a or R_a. We only use ratios of Q_as. It follows from these ideas that (in our setting) variation in the choices of consideration-only peers is equivalent to menu variation in the stochastic choice literature (Aguiar et al., 2023). In particular, if one has enough consideration-only peers, we can identify the counterfactual CCPs for binary menus $P_a^*(v \mid \mathbf{y}, \{0, v\}) = Q_a\left(v \mid NC_a^v\left(\mathbf{y}\right)\right) R_a\left(v \mid NR_a^v\left(\mathbf{y}\right), \{0, v\}\right)$. Hence, if either $Q_a\left(v \mid NC_a^v\left(\mathbf{y}\right)\right)$ or $R_a\left(v \mid NR_a^v\left(\mathbf{y}\right), \{0, v\}\right)$ is known, we can recover $Q_a\left(v \mid \cdot\right)$ (by Proposition 3.4) and then $R_a\left(v \mid NR_a^v\left(\mathbf{y}\right), \{0, v\}\right)$ from our recent ideas. Applying the same argument to menus of size three, we can identify R_a for sets of size three, and so on.

Proposition 3.6. *Suppose the assumptions of Proposition 3.5 are satisfied. If, in addition, we have that $|\mathcal{NC}_a \setminus \mathcal{NR}_a| \geq Y - 1$ and, for each $v \neq 0$, either $Q_a\left(v \mid n_1\right)$ or $R_a\left(v \mid n_2, \{0, v\}\right)$*

is known for some n_1 and n_2 in the support, then Q_a and R_a are identified from P_a.

The assumption that either $Q_a(v \mid n_1)$ or $R_a(v \mid n_2, \{0, v\})$ is known for some n_1 and n_2 in the support can be satisfied in different settings. For example, it is satisfied if the default is never picked when it is part of a binary menu with some alternative v and when all preference peers pick v (i.e., $R_a(v \mid |\mathcal{NR}_a|, \{0, v\}) = 1$). This may happen when one decides whether to use a particular social media with the default choice being not to use any. It would be reasonable to think that if all the friends of a given person are using this social media, then the person will use it for sure. Another example when the assumption is satisfied is when the alternative is considered with probability 1 if enough (or all) consideration peers pick the alternative (i.e., $Q_a(v \mid |\mathcal{NC}_a|) = 1$). In the case of online games, this is the same as to say that a player considers for sure a game when all her peers have just played it.

3.2. Identification of P

This section studies identification of the CCPs, P, and the rates of the Poisson alarm clocks from two different datasets. In Dataset 1, the researcher observes the precise moment at which an agent revises her strategy and the configuration of choices at that time. In Dataset 2, the researcher observes the configuration of choices at fixed time intervals.

Assume the researcher observes agents' choices at time intervals of length Δ and can consistently estimate $\Pr\left(\mathbf{y}^{t+\Delta} = \mathbf{y}' \mid \mathbf{y}^t = \mathbf{y}\right)$ for each pair $\mathbf{y}', \mathbf{y} \in \mathcal{Y}^A$. We capture these transition probabilities by a matrix $\mathcal{P}(\Delta)$.[15] Let $e^{(\Delta \mathcal{W})}$ be the matrix exponential of $\Delta \mathcal{W}$. Then $\mathcal{P}(\Delta)$ relates to transition rate matrix $\mathcal{W}$ in Section 2.3 by $\mathcal{P}(\Delta) = e^{(\Delta \mathcal{W})}$.

The two datasets we consider differ regarding Δ: In Dataset 1, the time interval is very small, i.e., the researcher knows $\lim_{\Delta \to 0} \mathcal{P}(\Delta)$. This ideal dataset registers agents' choices at the exact time at which any given agent revises her choice. With the proliferation of online platforms and scanners, this kind of data is often available. In Dataset 2, the time interval is of arbitrary size, i.e., the researcher knows $\mathcal{P}(\Delta)$. In both cases, the identification question is whether (or under what extra restrictions) we can recover $\mathcal{W}$ from the transition

---

[15]Here again, we assume that the choice configurations are ordered according to the lexicographic order when we construct $\mathcal{P}(\Delta)$.

probabilities in $\mathcal{P}(\Delta)$ which are identified and estimated from the data directly.

Proposition 3.7 (Dataset 1). *If Assumptions 1, 2(i), and 3(i) hold, then the CCPs* P *and the rates of the Poisson alarm clocks* $(\lambda_a)_{a \in \mathcal{A}}$ *are identified from Dataset 1.*

The proof of Proposition 3.7 follows because when the time interval between the observations goes to zero, we can recover $\mathcal{W}$. At least two known cases produce the same result without assuming $\Delta \to 0$. One of them requires the length of the interval Δ to be below a threshold $\overline{\Delta}$. The issue of this approach is that the value of the threshold depends on details of the model that are unknown to the researcher. The second case requires the researcher to observe the dynamic system at two different intervals Δ_1 and Δ_2 that are not multiples of each other (see, for example, Blevins, 2017 and the literature therein). The following proposition, based on Theorem 1 in Blevins (2018), offers a third case in which the transition rate matrix can be identified from Dataset 2.

Proposition 3.8 (Dataset 2). *If Assumptions 1, 2(i), and 3(i) hold, and* $\mathcal{W}$ *has distinct eigenvalues that do not differ by an integer multiple of* $2\pi i/\Delta$, *where i denotes the imaginary unit, then* P *and* $(\lambda_a)_{a \in \mathcal{A}}$ *are generically identified from Dataset 2.*

The restriction on eigenvalues of $\mathcal{W}$ is a regularity condition that is generically satisfied.[16] The key element in proving Proposition 3.8 is that the transition rate matrix in our model is rather parsimonious since, at any given time, only one agent revises her selection with a nonzero probability. Thus, the transition rate matrix $\mathcal{W}$ has many zeros in known locations.

4. Extensions

4.1. History Dependence and Own Past Choices

We have assumed that the choices made by a given agent are only affected by the current aggregate choices made by her peers and ignore her own past choices. We next extend the model by allowing that both the consideration and preferences of a given agent depend

---

[16]See Blevins (2017) for a discussion of this assumption.

on the history of her own choices and those of her peers. As consideration probabilities can be 1, the dependence on past choices allows nontrivial dynamics in consideration sets. For instance, they may not change much over long periods. Thus, this extension allows us to accommodate, among others, persistence in consideration sets and choices —see the discussion after Assumption 2 in Section 2. We use these ideas in our empirical application.

There are many ways in which history can be embedded into the model. We propose here a possibility that allows us to model an interesting situation (described below) and requires minimal extra notation. Let $\{t_k\}_{k=1}^{+\infty}$ be an (increasing) sequence of random time periods in which the clocks of different agents went off. Let $\mathbf{y}_{t_k}$ denote the configuration of choices in the network at k-th time period (at this moment the alarm clock of some agent went off). As a result, we can encode the whole history of choice configurations before moment t as $h_t = (\mathbf{y}_{t_k})_{t_k < t}$. Next, assume that the choice rules and consideration probabilities depend not only on choices made by peers at the moment at which the choice is revised but also on the whole history of choices h_t. Hence, given the history of choice configurations h_t, the probability that alternative v is picked by Agent a at time t would be

$$\mathrm{P}_a\left(v \mid \mathbf{y}_t, h_t\right) = \sum_{\mathcal{C} \subseteq \mathcal{Y}} \mathrm{R}_a\left(v \mid \mathbf{y}_t, h_t, \mathcal{NR}_a, \mathcal{C}\right)$$
$$\prod_{v' \in \mathcal{C}} \mathrm{Q}_a\left(v' \mid \mathbf{y}_t, h_t, \mathcal{NC}_a\right) \prod_{v' \in \mathcal{Y} \setminus \mathcal{C}}\left(1 - \mathrm{Q}_a\left(v' \mid \mathbf{y}_t, h_t, \mathcal{NC}_a\right)\right).$$

None of our previous results use variation beyond the choices made by connected agents at the moment of making a decision. Hence, if we condition on the choice made by Agent a, y_{at}, and the history h_t of choices, then we can establish the identification of all parts of the model from P_a using our previous ideas —thus, we omit the proof of the next result.

Proposition 4.1. *Suppose Assumptions 1 - 4 are satisfied conditional on y_{at} and the history h_t for all possible y_{at} and h_t. Also, let us extend the definition of P_a^* to allow for dependence on y_{at} and h_t. Then, all propositions from Section 3.1 are still valid.*

Proposition 4.1 takes as an input the CCPs that (now) depend on the histories of choices made by everyone in the network, i.e., it is implicitly assumed that P_a is identified. Since we only observe choices made by agents from one network, it would be impossible to identify

the CCPs conditional on *all* histories without further assumptions. To address this difficulty, we could restrict the length of the history that affects P_a.[17] This can be done by assuming that there exists finite $K \geq 1$ such that Q_a and R_a depend only on the first K components of h_t. Under this restriction, we would get that for any $k > K$, $(\mathbf{y}_{t_{k'}})_{k'=1,\ldots,k}$, v, and a

$$P_a \left(v \mid \mathbf{y}_{t_k}, (\mathbf{y}_{t_{k'}})_{k'=1,\ldots,k-1} \right) = P_a \left(v \mid \mathbf{y}_{t_k}, (\mathbf{y}_{t_{k'}})_{k'=k-K,\ldots,k-1} \right).$$

Hence, P can be recovered from Dataset 1.[18]

To motivate the details of this extension let us go back to the example of the online platform that offers video games to a set of players. As we mentioned in the introduction, these platforms often allow agents to form social networks and make the last purchased or played by peers game visible to the agent. Based on the history of acquired games, the platforms could also share with their subscribers the last few games that were acquired, and the identity of the players that acquired them. One could argue that recently acquired games could receive further attention by the subscribers of the platform. Recent games played by a subscriber could also have a special effect on her consideration set.

Example 1 (continued). Let $a^*(t_{k-1})$ be the agent that made a choice at the t_{k-1} moment. Hence, $\mathbf{y}_{a^*(t_{k-1}),t_k}$ is the choice that Agent $a^*(t_{k-1})$ made. Assume that consideration of an alternative (in addition to the previous arguments) depends on whether that alternative is the most recent choice made by a consideration peer and the alternative of the agent in the current choice configuration. That is,

$$Q_a \left(v \mid \mathbf{y}_{t_k}, h_{t_k}, \mathcal{NC}_a \right) = Q_a \left(v \mid y_{a,t_k}, \text{NC}_a^v(\mathbf{y}_{t_k}), \mathbb{1}\left(\mathbf{y}_{a^*(t_{k-1}),t_k} = v \right) \mathbb{1}\left(a^*(t_{k-1}) \in \mathcal{NC}_a \right) \right).$$

Thus, histories of length $K = 1$ affect the CCPs. For $\mathbf{y}_{t_k} = (1,1,2,0)$ and $\mathbf{y}_{t_{k-1}} = (1,0,2,0)$, we have that $a^*(t_{k-1}) = 2$ and $\mathbf{y}_{a^*(t_{k-1}),t_k} = 1$. Thus, the consideration probabilities of Agent

---

[17]CCPs could also be identified even if they depend on the whole choice history if one requires the impact of the remote past to decay sufficiently fast with time (see, Härdle et al., 1997, Bierens, 1996, and Truquet, 2023 for examples).

[18]Alternatively, one can also restrict the history in terms of the length of the time period rather than the number of actions. We could assume that there exists $\bar{t} > 0$ such that $P_a \left(v \mid \mathbf{y}_{t_k}, (\mathbf{y}_{t_{k'}})_{k'=1,\ldots,k-1} \right) = P_a \left(v \mid \mathbf{y}_{t_k}, (\mathbf{y}_{t_{k'}})_{\{k' : k' < k,\, t_k - t_{k'} < \bar{t}\}} \right).$

1 are $Q_1(1 \mid 1, 1, 1)$ and $Q_1(2 \mid 0, 1, 0)$ since option 1 (and not 2) was chosen at t_{k-1}. $\square$

4.2. Nonobservable Default

In many settings, the decision to choose the default alternative is often not observed. For example, if the default is "do nothing," then at any point in time that there is no change in the behavior of a given agent, we do not know whether she woke up and decided to do nothing or she did not have an opportunity to make a new decision. When this happens, even in continuous-time data setting (Dataset 1), there is no hope to separately identify λ_a and P_a. Therefore, some form of normalization is required. In our empirical application, we find it convenient to assume that $\lambda_a = 1$. This implies that, on average, agents have an opportunity to make a choice once per unit of time. Once λ_a is normalized, we can identify the CCPs P_a from the data directly, with which we can follow the identification results for network structure, consideration probabilities, and choice rules.

5. Application

We investigate peer effects in consideration and payoffs on the expansion decisions of the two dominant tea chains in the high-end tea industry in China. We have three goals: First, we showcase our identification strategy and provide a practical estimation procedure. Second, we show that ignoring the presence of limited consideration might mislead our estimates of profitability of different markets. Third, we quantify the direct effect of limited consideration and peer effect in consideration on the dynamics of market structure.

The tea beverage industry in China has seen a rapid expansion with its overall revenue increasing from 42.2 to 83.1 billion yuan from 2017 to 2020. This industry is divided into three segments: high-end, middle, and low-end. We study the two leading tea firms in the high-end segment —Heytea and Nayuki— which did not accept franchising before 2022. We acquired city-level store registration data from a commercial provider that sources records from the National Enterprise Credit Information Publicity System (CnOpenData, 2021).

31

This dataset allows us to determine when the enterprise enters and exits the market, enabling us to construct the cumulative number of stores in each city. We supplement the registration data with regional information from the China City Statistical Yearbook (National Bureau of Statistics of China, 2016-2021).[19] To avoid any changes in demand caused by the COVID-19 pandemic, we restrict our sample until the end of 2020. By then, Nayuki had 485 stores in 57 cities, while Heytea had 729 stores in 46 cities.

5.1. Empirical Model

We first describe the model of firm expansion decisions and then introduce the specifications for consideration and payoffs. We define a market at the level of a prefecture-level city, which ranks below a province and above a county in China's administrative structure. Thus markets are geographically isolated from each other. We collect all unknown primitives by θ.

Choice Set, Agents, and Peers There are finite sets of firms, $\mathcal{F}$, and markets to expand to, $\mathcal{M}$. Every firm f decides whether to open a store ($v = 1$) in market m or not ($v = 0$). We call a pair $a = (f, m) \in \mathcal{F} \times \mathcal{M}$ an agent —knowing the firm and the market identifies the agent and vise versa. Thus, $\mathcal{A} = \mathcal{F} \times \mathcal{M}$ and $\mathcal{Y} = \{0, 1\}$.[20] The set of markets in which these firms can open a new store is quite large, and the data correspond to a time when these firms were relatively new in the industry. We argue that managers might circumscribe the set of markets they consider at a given time to facilitate the decision process. At the moment of deciding whether to open a store, the attention that firm f pays to market m depends on its own and competitor's past choices in market m; it also depends on past openings at "neighboring" markets. Formally, $\mathcal{NC}_a$ and $\mathcal{NR}_a$ are the sets of pairs of firms and markets that affect consideration and payoffs, respectively, of firm f in market m.

As we just described, each firm decides whether to open a new store in each market. At the end of the analysis of network identification in Section 3.1, we modified Example 1 to show that with only one non-default option we cannot rely on double differences to

---

[19]Online Appendix C.1 offers more details on the dataset we use.

[20]We abuse notation a bit since $\mathcal{A}$ was previously defined as a set of the form $\{1, 2, \ldots, A\}$. To be consistent with the initial notation, we can take any one-to-one mapping $\tilde{a} : \mathcal{F} \times \mathcal{M} \to \{1, 2, \ldots, |\mathcal{F} \times \mathcal{M}|\}$ and define an agent as $a = \tilde{a}(f, m)$.

recover all parts of the network structure, and thereby need to assume partial knowledge of it. We follow the literature (e.g., Arcidiacono et al., 2016) and assume the marginal profit of firm f in market m from opening a new store is only affected by both its own and its competitors' openings in market m. Formally, $(f', m') \in \mathcal{NR}_{(f,m)}$ if and only if $m = m'$. After constructing $\mathcal{NR}_a$ in this way, we estimate the consideration network from the data. A potential concern with this assumption is the possibility of spillover effects across markets on profits. For example, the profits in different markets may be correlated through shipping cost savings from the distribution chain (see, for instance, Jia, 2008, Holmes, 2011, Zheng, 2016, Houde et al., 2023). We believe this is unlikely in our application as Nayuki relies on third-party logistics for shipping instead of building its own distribution centers. Moreover, its financial report states that storage and shipping accounted for approximately only 1.9 percent of total revenues in 2020.[21][22] Another potential source of spillover effects is information aggregation: the more stores a firm has in the area, the more information it has about the profitability of a particular market. In Online Appendix C.3, we re-estimate the model allowing the number of own stores in the nearby markets to affect firm's profitability in the focal market —keeping the assumption that rival's stores in the nearby markets only affect consideration. We show the main results are robust to these potential spillover effects.

Observable Characteristics Every market m at every moment of time t is characterized by observed market characteristics S_{mt} (e.g., GDP and population density) that include a constant. Let N_{at} denote the number of stores of Agent a (i.e., the number of stores of firm f in market m). Also define $S_t = (S_{mt})_{m \in \mathcal{M}}$ and $N_t = (N_{at})_{a \in \mathcal{A} = \mathcal{F} \times \mathcal{M}}$.

Market Consideration In our application, there are 71 markets where at least one firm opened a store by the end of the measurement. As we already justified, we allow firms to consider only a subset of markets when making a decision. Given the numbers of stores each firm had in the market at time t, N_t, the probability that firm f considers opening a new

---

[21]http://www.cn156.com/cms/scm/105854.html, Assessed August 2025.

[22]We had not found any reliable sources for Heytea. But, being similar in other respects, we suspect the same argument applies to Heytea.

store in market m at time t is

$$Q_a(1 \mid N_t, S_t, \mathcal{NC}_a) = F_{\tilde{\varepsilon}}\left(\bar{\tilde{\pi}}_{at}(S_t, N_t; \theta)\right),$$

where $F_{\tilde{\varepsilon}}$ is a known c.d.f.; θ is the vector of unknown parameters; and $\bar{\tilde{\pi}}_{at}(S_t, N_t; \theta)$ is the mean attention index, which is known up to θ. We allow the current market features of market m (including the market characteristics and all firms' number of stores) to affect the attention index of Agent a. Moreover, we allow the market structure of Agent a's neighborhood markets to affect her attention to market m.

Payoff from a New Store Conditional on a market being considered, the firm decides whether to open at least one new store in that market based on its marginal profit π_{at}. This marginal profit captures not just the instantaneous (one period) profitability of an extra store, but the expected profitability of the store in the long run.[23] The probability of opening a new store in market m by firm f at time t conditional on it being considered is

$$R_a(1 \mid N_t, S_t, \mathcal{NR}_a, \{0, 1\}) = F_{\varepsilon}\left(\bar{\pi}_{at}(S_t, N_t; \theta)\right),$$

where F_{ε} is a known c.d.f. and $\bar{\pi}_{at}(S_t, N_t; \theta)$ is the mean marginal profit —known up to θ.

Model Implied CCP Altogether, the probability that firm f opens a new store in market m (with $a = (f, m)$) is

$$\mathrm{Pr}_a(1 \mid N_t, S_t; \theta) = F_{\tilde{\varepsilon}}\left(\bar{\tilde{\pi}}_{at}(S_t, N_t; \theta)\right) F_{\varepsilon}\left(\bar{\pi}_{at}(S_t, N_t; \theta)\right),$$

which completely characterizes the probability of observing a new store in a given market by a given firm conditional on the history and the market characteristics. When evaluated at the true parameter value θ_0, Pr_a matches the CCP P_a, i.e., $\mathrm{P}_a(1 \mid N_t, S_t) = \mathrm{Pr}_a(1 \mid N_t, S_t; \theta_0)$.

The vector of parameters θ contains the parameters entering $\bar{\tilde{\pi}}_{at}$ and $\bar{\pi}_{at}$ that relate to both the covariates and the consideration network structure $\mathcal{NC}_a$, $a \in \mathcal{A}$. Note that

---

[23]We do not explicitly model forward-looking behavior to focus on the peer effect in limited consideration on firms' decisions. Incorporating forward-looking behavior can be computationally intensive and requires additional assumptions about players' expectations and beliefs, which is beyond the scope of this paper.

$\mathcal{NR}_a = \{(f', m') \; : \; f' \neq f, m = m'\}$ is assumed to be known and, thus, is not a part of θ.

5.2. Estimation

Data The data we have consist of three objects: (i) the exact date of store openings $\{t_k\}_{k=1}^K$; (ii) the state of the market structure $\{N_{at_k}\}_{a \in \mathcal{A}, k=1,\dots,K}$ sampled from a continuous time over interval $[0, t_K]$, where N_{at_k} is the number of stores owned by firm f in market m immediately prior to k-th change at time t_k —the last date of measurements coincides with the last day in which any action was observed; and (iii) observable market characteristics $\{S_{m,t_k}\}_{a \in \mathcal{A}, k=1,\dots,K}$.

Likelihood Function Our identification argument is constructive and can be used to estimate the model nonparametrically. However, for small and moderate-sized samples, we suggest using the parametric maximum likelihood estimator of CCPs P_a, as it is reasonably flexible in allowing market-specific consideration network links and efficiently uses all variations across markets (see Online Appendix B for details on the nonparametric estimator). We also add the network links in the parametrization of CCPs to estimate the CCPs, these network links, and the other consideration and payoff parameters in one step —instead of first estimating the CCPs and then applying our identification argument to estimate the rest of the model. In particular, we construct from the data a state vector $r_{t_k} = (r_{at_k})_{a \in \mathcal{A}}$, where r_{at_k} indicates whether there is a change in the number of stores of firm f in market m at time t_k, i.e., $r_{at_k} = \mathbb{1}\left(N_{at_{k+1}} > N_{at_k} \right)$. The probability of observing r_{t_k}, given the data and model parameters θ conditional on an alarm clock going off, is

$$p(r_{t_k}, S_{t_k}, N_{t_k}; \theta) = \prod_{a:r_{at_k}=1} \mathrm{Pr}_a(1 \mid N_t, S_t; \theta) \times \prod_{a:r_{at_k}=0} \left[1 - \mathrm{Pr}_a(1 \mid N_t, S_t; \theta) \right].$$

Hence, the probability that no new stores are opened in any market by any firm, given market characteristics and number of stores already opened (probability of picking the default), is

$$p_0(S_{t_k}, N_{t_k}; \theta) = \prod_{a \in \mathcal{A}} \left[1 - \mathrm{Pr}_a(1 \mid N_{at}, S_{at}; \theta) \right].$$

Finally, given that the arrival process is exponential, the log-likelihood of observing the

data given θ and normalizing $\lambda_a = 1$ (the choice of "doing nothing" is not observed; see Section 4.2 for details on this normalization) is

$$\hat{L}(\theta) = \sum_{k=1}^{K} -(t_{k+1} - t_k)\lambda(1 - p_0(S_{t_k}, N_{t_k}; \theta)) + \ln(\lambda p(r_{t_k}, S_{t_k}, N_{t_k}; \theta)).$$

The maximum likelihood estimator of θ, $\hat{\theta}$ is defined as the maximizer of $\hat{L}$ over a parameter space Θ.[24] The estimator $\hat{\theta}$ leads to an estimator of CCPs

$$\hat{P}_a(1 \mid N_t, S_t) = \Pr_a(1 \mid N_t, S_t; \hat{\theta}).$$

The construction of confidence sets for the parameters and estimated CCPs would require taking into account the estimation error in the estimated network. We leave this difficult problem for future research. An important feature stemming from Assumptions 1, 2(i), and 3(i), as shown by Proposition 2.1, is that our model has a single equilibrium or invariant distribution, i.e., the multiplicity of equilibria in the data-generating process is not an issue in our estimation.

Parameterization We assume that $F_{\tilde{\varepsilon}}$ and F_{ε} are Logistic c.d.f. Given that our sample size is small relative to the number of agents —there are 598 different dates in which we observed firms opening a store[25] for 71×2 agents— we use the following second-degree polynomial parameterization to flexibly approximate mean marginal profits and mean attention index:[26]

$$\bar{\pi}_{at}(S_t, N_t; \theta) = S'_{mt}\beta_f + \sum_{f'} \left[N_{(f',m)t}\alpha_{f,f'} + N^2_{(f',m)t}\gamma_{f,f'} \right],$$

$$\bar{\bar{\pi}}_{at}(S_t, N_t; \theta) = S'_{mt}\tilde{\beta}_f + \sum_{f'} \left[N_{(f',m)t}\tilde{\alpha}_{f,f'} + N^2_{(f',m)t}\tilde{\gamma}_{f,f'} \right] +$$

$$+ \sum_{f'} \left[\tilde{\delta}_{f,f'} \sum_{a'':f''=f'} \delta_{m,m''}N_{a''t} + \tilde{\eta}_{f,f'} \left(\sum_{a'':f''=f'} \delta_{m,m''}N_{a''t} \right)^2 \right].$$

---

[24]Checking all possible network structures is not feasible in our application. We use a variation of a greedy algorithm. See Online Appendix C.2 for further details.

[25]For some days, we observe multiple agents opening a store. As we mentioned earlier, the identification results in Section 3.1 are still valid if agents have perfectly synchronized clocks.

[26]We have estimated the model under several alternative specifications. The results are qualitatively the same and are available upon request.

The parameter $\delta_{m,m''} \in \{0,1\}$ captures the consideration network structure. It is equal to 1 if stores in market m'' affect consideration in market m and it is 0 otherwise. Note that estimation of the consideration network and probabilities crucially relies on variation of the number of openings by the two firms in nearby markets. For example, $\hat{\delta}_{m,m''} = 1$ if allowing the variation of openings in market m'' to affect the CCPs of market m increases the value of the likelihood function. That is, as in our identification strategy, the estimates add market m'' to the consideration network of market m if changing the number of stores in market m'' affects firms' CCPs in market m. Therefore, it is the variation in openings that allows us to pin down the values of the parameters that define the different parts of the model in the maximum likelihood estimation.

The mean marginal profit has two parts: the first one captures the impact of the observable market characteristics, and the second part captures the impact of the number of stores all firms have in market m.[27] The mean attention has an additional part capturing the peer effect in consideration from markets different from m, i.e., we allow firm-specific peer effects.[28]

The parameterization we use imposes two restrictions that are not needed for identification but reduce the computational burden: First, the payoff and consideration parameters are firm-specific and do not change across markets. Second, the consideration network link parameters $\delta_{m,m''}$ (that capture $\mathcal{NC}_a$) vary across markets but not across firms.

Remark 6. Since we use the total number of stores opened by each firm in every market as the determinant of consideration and expansion probabilities, formally we have a model with infinite history dependence. This, however, does not constitute any issues in our application, since the mean attention and mean marginal profits take known parametric forms.

---

[27]The fully structural model of marginal profits should contain information on fixed and marginal costs and prices among many other things. We specify the marginal profit function in the reduced form because of the availability of the data and to simplify the analysis.

[28]We use a second-degree polynomial in our approximation to capture potential nonmonotonicities. For instance, in markets with few stores, firms may act as complements, while, in more mature markets, firms may become substitutes. This would result in marginal profits that are not monotonic in competitors' stores.

5.3. Estimation Results

Network Structure The estimated consideration network has 266 directed links (i.e., the adjacency matrix is not symmetric and has 266 nonzero elements). The initial network we used in optimization uses spatial information about markets (see Online Appendix C.2) and allows up to 563 links, so we use the likelihood value to close down almost 300.

Consideration We calculated the consideration probabilities for every market and each firm using the numbers of stores and covariate values observed at the beginning (i.e., when Heytea started operating) and at the end of the measurements. Figure 1 shows the fraction of markets as a function of consideration probabilities for Heytea and Nayuki, respectively. Both firms display substantial limited consideration at the beginning: the averages (standard deviations) across markets of the consideration probabilities are 0.005 (0.006) and 0.029 (0.03) for Heytea and Nayuki, respectively. By the end of 2020, as a consequence of increases in the number of stores that the firms have in different markets, the consideration probabilities became substantially larger, with Nayuki becoming an almost full consideration firm: the averages (standard deviations) across markets of the consideration probabilities are 0.027 (0.04) and 0.82 (0.38) for Heytea and Nayuki, respectively.[29] We next show that incorporating limited attention is important for obtaining accurate estimates of profitability of markets and that it affects market structure —which impacts consumer welfare.

Marginal Profits Estimates We analyze the probabilities of opening a new store across markets (conditional on being considered). To quantify the effect of adding limited consideration to the expansion decision, we also estimated the marginal profit parameters assuming that all markets are considered. We refer to the former as limited consideration estimates and to the latter as full consideration estimates. The results of the estimation are presented in Figure 2. The difference between the limited and full consideration expansion probabilities is striking. In the beginning, the full consideration model would substantially

---

[29]As anecdotal evidence for our findings, the news reported that Nayuki implemented a nationwide city expansion plan. Beginning in late 2017, expansion moved beyond the province of Guangdong, rapidly extending into South China, Central China, East China, and other regions (https://news.qq.com/rain/a/20230109A03FOE00?utm.com, Accessed August 2025). In contrast, Heytea announced in 2020 its plan to continue focusing primarily on first-tier and provincial capital cities in China (https://news.sina.cn/gn/2020-07-14/detail-iivhvpwx5274609.d.html, Accessed August 2025).
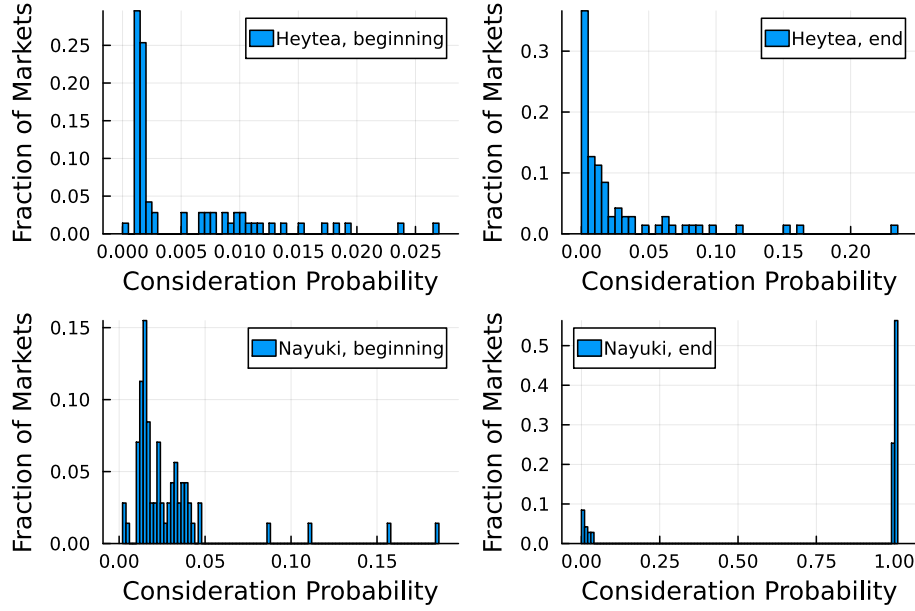
Figure 1 – **Normalized histogram of consideration probabilities for both firms at the data's beginning and end.**

underestimate the expansion probabilities (all points are below the 45°-line). By the end of 2020, while the discrepancy almost disappeared for Nayuki, which is not surprising given that it became an almost full consideration firm, the expansion probabilities for Heytea are still heavily underestimated. Thus, ignoring limited consideration leads to completely misleading estimates about the profitability of different markets. Qualitatively, this difference is explained by the fact that the full consideration model attributes "not-opening" a new store to negative marginal profits instead of limited consideration.

5.4. Counterfactuals

We evaluate the effect of limited consideration on market structure by comparing the fraction of monopolistic, duopolistic, and markets that are not served across time between our limited consideration model and a situation in which firms were fully attentive, i.e., the firms consider all 71 markets but marginal profit functions are kept the same as in our model estimates.

The simulation starts with zero stores and then creates expansion decisions for about 500 days. Figure 3 depicts the fraction of monopolistic, duopolistic, and markets not served by any firm as a function of time in the estimated limited consideration setting and the one
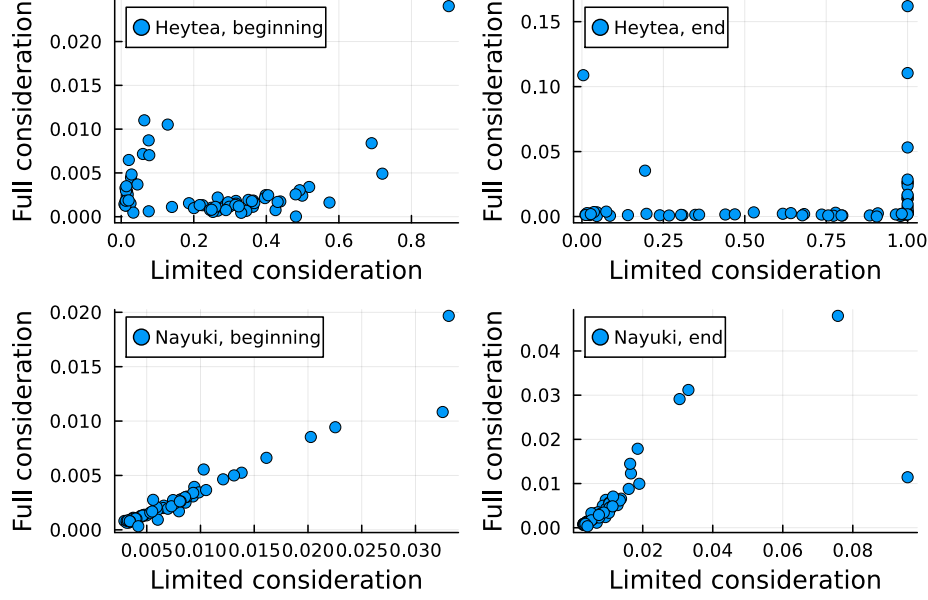
Figure 2 – **Limited consideration vs. full consideration expansion probabilities.**

that imposes full consideration. Limited consideration has a large effect on the dynamics of market structure. With full consideration, almost all markets would be served rather faster. For instance, in less than a year both firms would be present in about half of the markets.
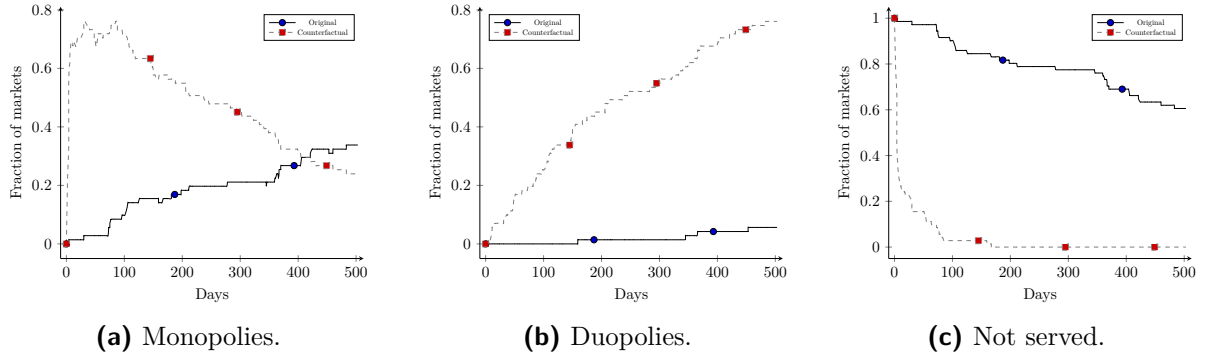

(a) Monopolies.　　　　　　**(b)** Duopolies.　　　　　　**(c)** Not served.

Figure 3 – **Counterfactual Scenario 1. Fraction of monopolistic, duopolistic, and markets that are not served over time.**

In the second counterfactual, we remove connections across markets in consideration. Specifically, we shut down the effect of own and opponent stores in the *neighboring* markets on consideration.[30] The market penetration is not affected much. As part (b) of Figure 4

---

[30]We abstract away from any potential dependence between the network formation and the expansion decision processes and assume that the network is an exogenously given fixed parameter. This assumption is important for any form of counterfactual analyses that involves changes in the network structure.

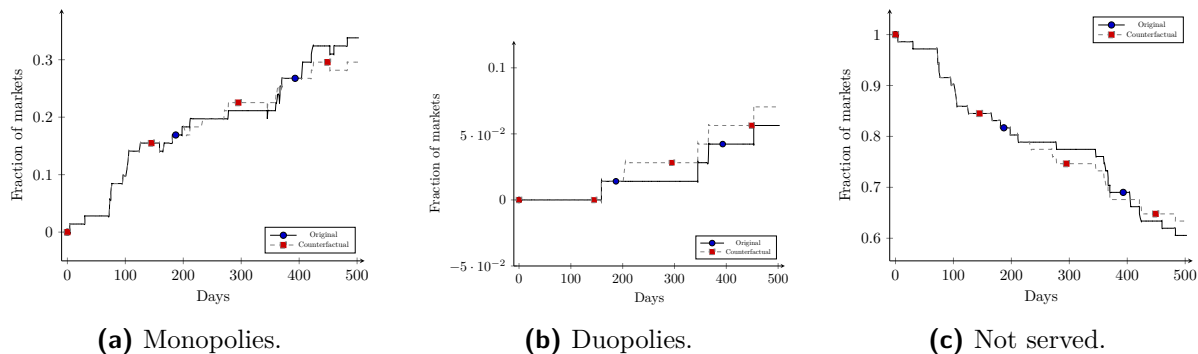shows peer-effects on consideration from the neighboring markets speed up competition.

(a) Monopolies. **(b)** Duopolies. **(c)** Not served.

Figure 4 – **Counterfactual Scenario 2. Fraction of monopolistic, duopolistic, or markets that are not served over time.**

6. Conclusion

This paper offers a model of interactions in which different types of peers affect the choices of a given agent via different mechanisms. We show that these peer effect mechanisms have different behavioral implications in the data. This allows us to recover the set of connections between the agents and the type of interactions between them. The choices of different types of peers act as dual exclusion restrictions and allow us to recover the consideration probabilities and the random preferences. We apply the model to data on tea chains expansions in China. The empirical application adds to the literature on boundedly rational firms. While studying a rather general model, we leave a few interesting variants for future research, such as forward-looking behavior or the possibility that each agent ends up making decisions with the purpose of affecting the consideration set of others. We believe this set up could lead to a new model of endogenous social norms.

References

Abaluck, J. and A. Adams-Prassl (2021) "What do consumers consider before they choose? Identification from asymmetric demand responses," *The Quarterly Journal of Economics*,

136 (3), 1611–1663.

Agranov, M., M. Elliott, and P. Ortoleva (2021) "The importance of Social Norms against Strategic Effects: The case of Covid-19 vaccine uptake," *Economics Letters*, 206, 109979.

Aguiar, V. H. (2017) "Random categorization and bounded rationality," *Economics Letters*, 159, 46–52.

Aguiar, V. H., M. J. Boccardi, and M. Dean (2016) "Satisficing and stochastic choice," *Journal of Economic Theory*, 166, 445 – 482, https://doi.org/10.1016/j.jet.2016.08.008.

Aguiar, V. H., M. J. Boccardi, N. Kashaev, and J. Kim (2023) "Random utility and limited consideration," *Quantitative Economics*, 14 (1), 71–116.

Aguiar, V. H. and N. Kashaev (2024) "Identification and estimation of discrete choice models with unobserved choice sets," *Journal of Business & Economic Statistics*, 1–25.

Aguirregabiria, V. and P. Mira (2007) "Sequential estimation of dynamic discrete games," *Econometrica*, 75 (1), 1–53.

Allen, R. and J. Rehbeck (2023) "Latent Utility and Permutation Invariance: A Revealed Preference Approach," *Available at SSRN 4340420*.

Arcidiacono, P., P. Bayer, J. R. Blevins, and P. B. Ellickson (2016) "Estimation of dynamic discrete choice models in continuous time with an application to retail competition," *The Review of Economic Studies*, 83 (3), 889–931.

Armstrong, M. and S. Huck (2010) "Behavioral economics as applied to firms: A primer."

Bajari, P., L. C. Benkard, and J. Levin (2007) "Estimating dynamic models of imperfect competition," *Econometrica*, 75 (5), 1331–1370.

Barseghyan, L., M. Coughlin, F. Molinari, and J. C. Teitelbaum (2021a) "Heterogeneous choice sets and preferences," *Econometrica*, 89 (5), 2015–2048.

Barseghyan, L., F. Molinari, and M. Thirkettle (2021b) "Discrete choice under risk with limited consideration," *American Economic Review*, 111 (6), 1972–2006.

Berry, S. T. (1992) "Estimation of a model of entry in the airline industry," *Econometrica*, 60 (4), 889–917.

Bierens, H. J. (1996) *Topics in advanced econometrics: estimation, testing, and specification of cross-section and time series models*: Cambridge university press.

Blevins, J. R. (2017) "Identifying restrictions for finite parameter continuous time models with discrete time data," *Econometric Theory*, 33 (3), 739–754.

——— (2018) "Identification and estimation of continuous time dynamic discrete choice games,"Technical report, Mimeo, Ohio State University.

Blevins, J. R., A. Khwaja, and N. Yang (2018) "Firm expansion, size spillovers and market dominance in retail chain dynamics," *Management Science*, 64 (9), 3971–4470.

Blume, L. E. (1993) "The statistical mechanics of strategic interaction," *Games and Economic Behavior*, 5 (3), 387–424.

——— (1995) "The statistical mechanics of best-response strategy revision," *Games and Economic Behavior*, 11 (2), 111–145.

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides (2011) "Identification of social interactions," in *Handbook of Social Economics*, 1, 853–964: Elsevier.

Blume, L. E., W. A. Brock, S. N. Durlauf, and R. Jayaraman (2015) "Linear social interactions models," *Journal of Political Economy*, 123 (2), 444–496.

Bonaldi, P., A. Hortaçsu, and J. Kastl (2015) "An empirical analysis of funding costs spillovers in the euro-zone with application to systemic risk,"Technical report, NBER.

Borah, A. and C. Kops (2018) "Choice via Social Influence,"Technical report, Mimeo.

Brady, R. L. and J. Rehbeck (2016) "Menu-Dependent Stochastic Feasibility," *Econometrica*, 84 (3), 1203–1223.

Bramoullé, Y., H. Djebbari, and B. Fortin (2020) "Peer effects in networks: A survey," *Annual Review of Economics*, 12, 603–629.

Bresnahan, T. f. and P. C. Reiss (1990) "Entry in monopoly market," *Review of Economic Studies*, 57 (4), 531–553.

Caeyers, B. (2014) "Peer effects in development programme awareness of vulnerable groups in rural Tanzania,"Technical report, CSAE, University of Oxford.

Caplin, A., M. Dean, and J. Leahy (2019) "Rational inattention, optimal consideration sets, and stochastic choice," *The Review of Economic Studies*, 86 (3), 1061–1094.

Cattaneo, M. D., X. Ma, Y. Masatlioglu, and E. Suleymanov (2020) "A random attention model," *Journal of Political Economy*, 128 (7), 2796–2836.

Chambers, C. P., T. Cuhadaroglu, and Y. Masatlioglu (2023) "Behavioral influence," *Journal of the European Economic Association*, 21 (1), 135–166.

Ciliberto, F. and E. Tamer (2009) "Market structure and multiple equilibria in airline markets," *Econometrica*, 77 (6), 1791–1828.

CnOpenData (2021) "Firm registration data for Chinese cities, 2015–2020," `https://www.cnopendata.com`, `https://www.gsxt.gov.cn`, Accessed June, 2024.

Conlon, C. T. and J. H. Mortimer (2013) "Demand estimation under incomplete product availability," *American Economic Journal: Microeconomics*, 5 (4), 1–30.

Crawford, G. S., R. Griffith, and A. Iaria (2021) "A survey of preference estimation with unobserved choice set heterogeneity," *Journal of Econometrics*, 222 (1), 4–43.

Dardanoni, V., P. Manzini, M. Mariotti, and C. J. Tyson (2020) "Inferring cognitive heterogeneity from aggregate choices," *Econometrica*, 88 (3), 1269–1296.

De Paula, A. (2017) "Econometrics of network models," in *Advances in Economics and Econometrics: Theory and Applications, Eleventh World Congress*, 268–323, Cambridge University Press Cambridge.

De Paula, Á., I. Rasul, and P. Souza (2023) "Identifying Network Ties from Panel Data: Theory and an Application to Tax Competition," *The Review of Economic Studies*, forthcoming.

Draganska, M. and D. Klapper (2011) "Choice set heterogeneity and the role of advertising: An analysis with micro and macro data," *Journal of Marketing Research*, 48 (4), 653–669.

Duflo, E. and E. Saez (2003) "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment," *The Quarterly journal of economics*, 118 (3), 815–842.

Durlauf, S. N. and H. P. Young (2001) *Social dynamics*, 4: Mit Press.

Enikolopov, R., A. Makarin, and M. Petrova (2020) "Social media and protest participation: Evidence from Russia," *Econometrica*, 88 (4), 1479–1514.

Gaynor, M., C. Propper, and S. Seiler (2016) "Free to choose? Reform, choice, and consideration sets in the English National Health Service," *American Economic Review*, 106 (11), 3521–57.

Godes, D. and D. Mayzlin (2004) "Using online conversations to study word-of-mouth communication," *Marketing science*, 23 (4), 545–560.

Goeree, M. S. (2008) "Limited information and advertising in the US personal computer industry," *Econometrica*, 76 (5), 1017–1074.

Graham, B. S. (2015) "Methods of identification in social networks," *Annu. Rev. Econ.*, 7 (1), 465–485.

Härdle, W., H. Lütkepohl, and R. Chen (1997) "A review of nonparametric time series analysis," *International statistical review*, 65 (1), 49–72.

Heidhues, P. and B. Kőszegi (2018) "Behavioral industrial organization," *Handbook of Behavioral Economics: Applications and Foundations 1*, 1, 517–612.

Holmes, T. J. (2011) "The diffusion of Wal-Mart and economies of density," *Econometrica*, 79 (1), 253–302.

Horan, S. (2019) "Random consideration and choice: A case study of "default" options," *Mathematical Social Sciences*, 102, 73–84.

Houde, J.-F., P. Newberry, and K. Seim (2023) "Nexus Tax Laws and Economies of Density in E-Commerce: A Study of Amazon's Fulfillment Center Network," *Econometrica*, 91 (1), 147–190, https://doi.org/10.3982/ECTA15265.

Jia, P. (2008) "What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry," *Econometrica*, 76 (6), 1263–1316.

Kashaev, N. ⓡ V. H. Aguiar (2022) "A random attention and utility model," *Journal of Economic Theory*, 204, 105487, https://doi.org/10.1016/j.jet.2022.105487.

Kashaev, N. and N. Lazzati (2019) "Peer effects in random consideration sets," *arXiv preprint arXiv:1904.06742*.

Lazzati, N. (2020) "Co-Diffusion of Technologies in Social Networks," *American Economic Journal: Microeconomics*, 12, 193–228.

Lee, S. Y. (2015) "Interpersonal influence on online game choices," *Computers in Human Behavior*, 45, 129–136.

Lewbel, A., X. Qu, and X. Tang (2023) "Social networks with unobserved links," *Journal of Political Economy*, 131 (4), 898–946.

Lleras, J. S., Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay (2017) "When more is less: Limited consideration," *Journal of Economic Theory*, 170, 70–85.

Lu, Z. (2022) "Estimating multinomial choice models with unobserved choice sets," *Journal of Econometrics*, 226 (2), 368–398, https://doi.org/10.1016/j.jeconom.2021.06.004.

Manresa, E. (2013) "Estimating the structure of social interactions using panel data," *Unpublished Manuscript. CEMFI, Madrid*.

Manski, C. F. (1977) "The structure of random utility models," *Theory and decision*, 8 (3), 229–254.

———— (2000) "Economic analysis of social interactions," *Journal of economic perspectives*, 14 (3), 115–136.

Manzini, P. and M. Mariotti (2014) "Stochastic choice and consideration sets," *Econometrica*, 82 (3), 1153–1176.

Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012) "Revealed Attention," *American Economic Review*, 102 (5), 2183–2205, 10.1257/aer.102.5.2183.

Mehta, N., S. Rajiv, and K. Srinivasan (2003) "Price uncertainty and consumer search: A structural model of consideration set formation," *Marketing Science*, 22 (1), 58–84.

National Bureau of Statistics of China (2016-2021) "China City Statistical Yearbook," Annual series; various years consulted.

Qiu, L., Z. Shi, and A. B. Whinston (2018) "Learning from your friends' check-ins: An empirical study of location-based social networks," *Information Systems Research*, 29 (4), 1044–1061.

Roberts, J. H. and J. M. Lattin (1991) "Development and testing of a model of consideration set composition," *Journal of Marketing Research*, 28 (4), 429–440.

Simon, H. A. (1945) *Administrative behavior*: Macmillan.

———— (1955) "A behavioral model of rational choice," *The Quarterly Journal of Economics*, 99–118.

Truquet, L. (2023) "Strong mixing properties of discrete-valued time series with exogenous covariates," *Stochastic Processes and their Applications*, 160, 294–317.

Zheng, F. (2016) "Spatial competition and preemptive entry in the discount retail industry," *Columbia Business School Research Paper* (16-37).