Peer Effects in Consideration and Preferences*

Nail Kashaev Natalia Lazzati Ruli Xiao†

This version: January 18, 2024

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 only preferences, only 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 the preferences and the consideration mechanisms of each agent. 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 coffee chains and find evidence of limited consideration. We simulate counterfactual predictions and show how limited consideration slows down competition.

JEL codes: C31, C33, D83, O33

Keywords: Peer Effects, Random Preferences, Random Consideration Sets

*This paper subsumes Kashaev and Lazzati (2019). For their useful comments and suggestions we thank Victor Aguiar, Roy Allen, Tim Conley, Steven Durlauf, Yao Luo, Mathieu Marcoux, Rory McGee, Paola Manzini, Marco Mariotti, Salvador Navarro, Áureo de Paula, David Rivers, and Al Slivinski. We would also like to thank the audience at Workshop in Experiments, Revealed Preferences, and Decisions, the Inter University Colloquium on Analytical Research in Economics, the CARE-Colloquium series at University of Kansas, NAWM ES 2021, the Econometrics of Games, Matching and Network at Toulouse, Bristol ESG 2022, MEG 2022, Midwest Economic Theory and International Economics Meetings 2022, 2023 PennState Alumni Conference, CESG 2023, and the seminar presentations at El Colegio de Mexico, Cornell, LSE, MSU, Queen Mary University of London, Rice, UCL, UC Riverside, Washington University in St. Louis, and York.

†Kashaev: Western University; nkashaev@uwo.ca. Lazzati: UC Santa Cruz; nlazzati@ucsc.edu. Xiao: Indiana University; rulixiao@iu.edu.

1. Introduction

It has long been agreed that agents are subject to social influence when making decisions.[1] These interactions have been shown to be important for individuals in areas such as health and education and for firms in a variety of decisions such as opening a new store. It has also been argued that agents might affect the decisions of others in different ways.[2] 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 can affect different aspects of the decision process. First, the choices of peers can affect the subset of options that the agent ends up considering.[3] Second, these choices can affect the preferences of the agent over the set of alternatives. We show that these two mechanisms have different behavioral implications in the data. We exploit them to further show that all parts of the model can be recovered from a sequence of choices and illustrate our ideas with an empirical application of expansion decisions of the two largest coffee chains in China.

In our model, agents are linked through a network. An important aspect of our approach is that peers might have different roles in the decision process of a given agent. Thus, the network defines not just the direction of the connection but also the way in which agents affect each other. In particular, a link specifies whether a peer affects the consideration of alternatives, the ranking of preferences, or both.[4] More to the point, at a randomly given time, an agent gets the opportunity to select a new option out of a finite set of alternatives. As in the consideration set models, the agent does not pay attention to all the available options at the moment of choosing. 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. As in the canonical peer effects models, the choices of

[1] See Durlauf and Young (2001) and the references therein for examples.

[2] See Manski (2000) for a discussion of the channels through which agents can affect each other.

[3] This possibility has been (explicitly or implicitly) discussed by other researchers in specific contexts —e.g., the choices of peers 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, Shi and Whinston, 2018), or an opportunity to protest (Enikolopov, Makarin and Petrova, 2020).

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

peers can also affect the preferences of the agent over alternatives. This model leads to a sequence of choices that evolves through time according to a continuous-time Markov random process.

The model we build 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).[5] An agent can purchase any game. 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 purchasing decision. Platforms often allow agents to form social networks that make the last purchased or played game by peers visible to the agent. This information may help the agent to shrink the subset of games she ends up considering. Moreover, some of these games are played in groups. In these cases, the choices of peers can also directly affect the utility the agent gets from playing a particular game. Our model can help the platform personalize each agent's reference groups to maximize profits or the probability of making a sale.

After we show equilibrium existence and characterize equilibrium behavior, we consider a researcher who observes a sequence of choices made by the network members. We show that all primitives of the model can be uniquely recovered. These primitives include the network structure, the consideration probabilities, and the preferences of every agent. There are three aspects of our nonparametric identification results that deserve special attention. First, we allow for very general forms of heterogeneity across agents with respect to all parts of the model. Second, we recover not just the set of peers for each agent, but also whether each of them affects consideration, preferences, or both. That is, we identify the full network structure. Third, in contrast to most other works on consideration sets, we do not rely on the variation of either exogenous covariates or the set of available options (menus) to identify the model primitives. Instead, we use the choices of peers. One can think of them as excluded covariates that affect different parts of the decision process.[6] Importantly, these excluded covariates vary *endogenously* in the model, and we need to *identify* them in the data. Thus, extra work needs to be done to use peers as means for recoverability.

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

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

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

the choices of others (at the moment of revising her selection). The identification strategy we offer is a two-step procedure. First, we show how to identify the primitives of the model by using these CCPs and the variation in the choices of peers. Second, we study identification of the CCPs from observed data.

The identification strategy we propose for the primitives of the model is novel, simple, and constructive. We first recover the reference group of each agent by using the fact that, in our framework, variation in the choices of peers induces variation in the CCPs of a given agent. For example, if the frequency by which an agent selects an alternative increases when another agent selects that option keeping the choice of everyone else fixed, then we can conclude that the latter agent is a peer of the former one. However, we cannot directly infer whether this increase is coming from the peer effect in consideration or preferences. The separation of the two mechanisms relies on a few key observations that we describe next.

We initially notice that 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 for the alternative, 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. Similarly, while the second probability varies when a preference peer chooses something different from the alternative, the first probability is not affected since it only captures consideration. In other words, an interdependence between alternatives is present in preferences, but not in consideration. This form of separability allows us to discern the peer effect in preferences from the one in consideration via a cross order effect of peers in alternatives in the CCPs.

When the network structure is recovered, as we mentioned before, choices of different types of peers can be used to construct exclusion restrictions to identify consideration mechanism and preferences. For example, variation in choices of peers that only affect consideration (consideration-only peers), can be used to identify changes in consideration probabilities, since the preference part of CCPs does not change when we vary the choices of consideration-only peers. To recover preferences, we first show that variation in choices of 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 be then used to identify preferences.

To identify the CCPs, we consider two datasets: continuous-time data and discrete-time data with arbitrary time intervals. These two datasets coincide in that they provide a long sequence of choices from agents in the network. They differ in the timing at which the researcher observes these choices. In continuous-time datasets, the researcher observes agents' choices in real-time. We can think of this dataset as the "ideal dataset." With the proliferation of online platforms and scanners, this sort of data is available for many applications involving individual decision making. Our empirical application is an example of this type of dataset involving firms' decisions. In continuous-time datasets, the researcher directly recovers the CCPs. In discrete-time datasets, the researcher observes the joint vector of choices at fixed time intervals (e.g., the choices of agents are observed every Monday). In this case, the CCPs are not directly observed, and they need to be inferred from the data. Adding an extra mild condition, we show that the CCPs are also uniquely identified. For this last result, we invoke insights from Blevins (2017, 2018).

We provide several empirically relevant extensions of our baseline model. In particular, we show that our results extend to finite history dependence; we explain how to proceed when one of the choices (e.g., "do nothing") is not observed in the data; we discuss how to identify the CCPs using short panel datasets; and we revise the framework to accommodate the case where multiple agents make choices simultaneously. We also build a random consideration model of bundles and show that all aspects of that model can be identified. Finally, we provide a new identification argument for a general model of random consideration that uses payoff-specific unbounded covariates. To the best of our knowledge, the latter result is new to the literature on random consideration sets and is of independent interest.

To showcase our methodology, we first provide evidence on the possible limited consideration in firms' behaviors and its impact on the expansion decisions of China's top two coffee chains, Starbucks and Luckin. Given the large set of markets (152 markets) to open a new store in, we argue that the firms may be boundedly rational by not considering all of them due to limited knowledge or ability of the "administrative man" (Simon, 1945, 1955).[7] We assume that for a given market the number of stores firms have in the neighboring markets affects the probability of considering that market, but does not affect the profitability of that market. This exclusion restriction allows us to identify the network structure —neighboring markets that affect consideration only (i.e. peer effect in consideration)— and the parameters governing the consideration and profitability.

We show that overlooking limited consideration can lead to misunderstandings about market profitability and, consequently, firm behavior. In particular, our results suggest that Starbucks displays full consideration, while Luckin does not consider all markets. As a result, while a full consideration model would explain the slow expansion of Luckin in some markets by low profitability, our estimation results indicate that these markets are often profitable, but Luckin does not expand in them because they are not considered. We also measure the impact of limited consideration and peer effect in consideration in shaping market structure. This is done by carrying out two counterfactual scenarios where we eliminate one of these factors at a time. We find that limited consideration and peer effect in consideration work in opposite directions. Specifically, while peer effects in consideration increase the rate at which markets are served by both firms, limited consideration slows down the emergence of duopolies.[8] Finally, we consider a counterfactual scenario where we exchange the initial conditions of Luckin and Starbucks. That is, we assume that Luckin has a large number of stores when Starbucks decides to open her first store. This change speeds up the emergence of duopolies, indicating that Starbucks expands substantially faster than Luckin when put in the same conditions. We interpret this results as evidence of Starbucks being less boundedly rational than Luckin. Let us finally remark that while most of the literature on consideration sets focuses on boundedly rational consumers, to the best of our knowledge, we are the first to apply similar ideas to the behavior of firms.

We finally 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

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

[8] Given that Starbucks considers almost all markets, the market penetration is not affected by limited consideration.

peer effect in consideration sets, we can use variation in the choices of peers as the main tool to recover random preferences. The literature on identification of single-agent consideration set models has mainly relied on the variation of the set of available options (menus). It includes Aguiar (2017), Aguiar, Boccardi and Dean (2016), Brady and Rehbeck (2016), Caplin, Dean and Leahy (2019), Cattaneo, Ma, Masatlioglu and Suleymanov (2020), Horan (2019), Kashaev ⓡ Aguiar (2022), Lleras, Masatlioglu, Nakajima and Ozbay (2017), Manzini and Mariotti (2014), and Masatlioglu, Nakajima and Ozbay (2012). (See Aguiar, Boccardi, Kashaev and Kim, 2023 for a comparison of several consideration set models in an experiment.) Other papers have relied on exogenous covariates that shift preferences or consideration sets. The latter include Barseghyan, Molinari and Thirkettle (2021b), Barseghyan, Coughlin, Molinari and Teitelbaum (2021a), Crawford, Griffith and Iaria (2021), Conlon and Mortimer (2013), Draganska and Klapper (2011), Gaynor, Propper and Seiler (2016), Goeree (2008), Mehta, Rajiv and Srinivasan (2003), and Roberts and Lattin (1991). Variation of exogenous covariates has also been used by Abaluck and Adams-Prassl (2021) via an approach that exploits symmetry breaks with respect to the full consideration set model. In the context of the model of random consideration sets of Manski (1977) and Manzini and Mariotti (2014), Abaluck and Adams-Prassl (2021) use unbounded alternative-specific covariates to generate exogenous menu variation that allows them to identify the consideration probabilities. Yu (2023) builds a bundles model with limited consideration but also relies on exogenous variation in observed covariates. Aguiar ⓡ Kashaev (2021), Allen and Rehbeck (2023), Crawford et al. (2021), and Dardanoni, Manzini, Mariotti and Tyson (2020) use repeated choices (i.e., they work with panel data) but do not allow for peer effects.

There is a vast econometric literature on identification of models of social interactions where choices of peers affect preferences but not the choice sets (see Blume, Brock, Durlauf and Ioannides, 2011, Bramoullé, Djebbari and Fortin, 2020, De Paula, 2017, and Graham, 2015 for comprehensive reviews of this literature). We depart from this literature in that in our framework, the direct interdependence between choices (endogenous effects) is captured by consideration sets in addition to preferences. Interestingly, we show that the two mechanisms can be set apart in practice. We view our work as complementing the existing results on the peer effect in preferences by adding a second mechanism that might be particularly 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, Brock, Durlauf and Jayaraman (2015), Bonaldi, Hortaçsu and Kastl (2015), De Paula, Rasul and Souza (2023), Lewbel, Qu and Tang (2023), and Manresa (2013). In the context of discrete choice, Chambers, Cuhadaroglu and Masatlioglu (2023) also identifies the network structure. However, in this paper, peers do not affect consideration sets but directly change preferences (among other differences).

Two other (theoretical) papers that incorporate peer effects in the formation of consideration sets are Borah and Kops (2018) and Lazzati (2020). Borah and Kops (2018) do so in a static framework and rely on the variation of menus for identification. Lazzati (2020) considers a dynamic model, but the time is discrete, and she focuses on two binary options that can be acquired together.

The rest of the paper is organized as follows. Section 2 presents the model, the main assumptions, and some key insights of our approach. Section 3 studies the empirical content of the model. Section 4 extends the initial model in several dimensions. Section 5 applies our model to expansion decisions by firms in the coffee market. Section 6 concludes. Appendix A provides the regularity condition for the identification. All the proofs are collected in Appendix B.

2. The Model

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

2.1. Network, Consideration Sets, and Preferences

Network and Choice Configuration There is a finite set of agents $\mathcal{A} = \{1, 2, \dots, A\}$, $A \geq 2$, and a finite set (menu) of alternatives $\mathcal{Y} = \{0, 1, 2, \dots, Y\}$, $Y \geq 1$, 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.[9]

The agents are connected through a network. We allow agents to interact with others in different ways. Specifically, the choices of peers can affect the set of alternatives the agent ends up considering, the preferences over the alternatives, or both. Thus, the network is described by two sets of edges between agents 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 Agent $a \in \mathcal{A}$

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

The full reference group is $\mathcal{N}_a = \mathcal{NC}_a \bigcup \mathcal{NR}_a$. We follow the convention and assume that $a \notin \mathcal{N}_a$. We show in Section 4.2 that all results hold if the past choices of Agent a affect either her consideration sets or preferences. Since we allow for the possibility that some peers affect both considerations and preferences, $\mathcal{NCR}_a = \mathcal{NC}_a \bigcap \mathcal{NR}_a$ can be nonempty.

Choice Revision Process We model the revision process of alternatives as a standard continuous-time Markov process on the space of choice configurations. We assume that 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.[10] When this happens, the agent forms a consideration set and then selects an alternative among the ones she is actually considering. 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 observation has useful implications for identification. (We discuss a model with perfectly correlated clocks in Section 4.5.)

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

[9]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.

[10]That is, each Agent a is endowed with a collection of random variables $\{\tau_n^a\}_{n=1}^{\infty}$ such that each difference $\tau_n^a - \tau_{n-1}^a$ is exponentially distributed with mean $1/\lambda_a$. These differences are independent across people and time.

moment of revising decisions. We indicate by $\mathrm{Q}_a(v \mid \mathbf{y}, \mathcal{NC}_a)$ the 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 $a \in \mathcal{A}$, $\mathbf{y} \in \mathcal{Y}^A$, the probability of facing consideration set $\mathcal{C}$, which is a subset of menu $\mathcal{Y}$, is

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

Since the consideration set cannot be empty, we assume that the default alternative is always considered. That is, $\mathrm{Q}_a(0 \mid \mathbf{y}, \mathcal{NC}_a) = 1$ 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). Leaving aside peer effects, this process of formation of consideration sets is analogous to the one studied by Manski (1977) and Manzini and Mariotti (2014).

We next offer a simple example.

Example 1. Assume that the attention given to alternative v is determined by its popularity among peers. In particular, its inclusion into the consideration set could be modeled as an indicator function $\mathbb{1}(c_{v,a}(\mathbf{y}, \mathcal{NC}_a) \geq \varepsilon_v)$, where $c_{v,a}$ measures the mean attention of Agent a to alternative v as a function of the choices of her peers; and ε_v is an independent of $\mathbf{y}$ random attention shock. That is, v is considered if and only if $c_{v,a}(\mathbf{y}, \mathcal{NC}_a) \geq \varepsilon_v$. Then, the probability of considering v is

$$\mathrm{Q}_a(v \mid \mathbf{y}, \mathcal{NC}_a) = F_\varepsilon(c_{v,a}(\mathbf{y}, \mathcal{NC}_a)),$$

where F_ε denotes the cumulative distribution function (c.d.f.) of ε. □

Peer Effect in Preferences After the consideration set is formed, the agent selects an alternative among the ones she is actually considering according to some choice rule. We allow preferences to be random. We do not need to specify a particular form of utility function since our object of interest is the choice rule —the distribution over elements of a given consideration set. In practice, one can identify and estimate the underlying preferences from the choice rule. Therefore, we focus on the choice rule and leave the association between the choice rule and preferences to be flexible. Formally, given consideration set $\mathcal{C}$, we let the choice rule $\mathrm{R}_a(\cdot \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C})$ be a distribution function supported on $\mathcal{C}$. That is, $\mathrm{R}_a(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C}) \geq 0$ for all v and

$$\sum_{v \in \mathcal{C}} \mathrm{R}_a(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C}) = 1.$$

The choice rule tells us what the probability of picking an alternative is in a given consideration set. Choice rules summarize the decision process after the consideration set is formed. The choices of agents may be random from the researcher's perspective after conditioning on consideration sets if there are latent preference shocks across choice instances (see Example 2) or if agents randomize when indifferent.

We incorporate the peer effect in preferences by allowing the choice rule of each agent to depend on the configuration of choices and her preference reference group $\mathcal{NR}_a$.

We next offer a simple example.

Example 2. Suppose the utility that Agent a gets from alternative v if it is considered in set $\mathcal{C}$ is given by $u_{a,v,\mathcal{C}}(\mathbf{y}, \mathcal{NR}_a) + \xi_v$, where $u_{a,v,\mathcal{C}}$ captures the mean utility from the alternative when it is in the given consideration set and the ξ_vs are independent and identically distributed (i.i.d.) taste shocks that are distributed according to the standard Type I extreme value distribution. As a result, for $v \in \mathcal{C}$,

$$\mathrm{R}_a(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C}) = \frac{\exp(u_{a,v,\mathcal{C}}(\mathbf{y}, \mathcal{NR}_a))}{\sum_{v' \in \mathcal{C}} \exp(u_{a,v',\mathcal{C}}(\mathbf{y}, \mathcal{NR}_a))}.$$

Note that consideration sets can directly affect utilities from alternatives. In particular, we allow for violations of the independence of irrelevant alternatives since, for $v \neq v'$,

$$\frac{\mathrm{R}_a(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C})}{\mathrm{R}_a(v' \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C})} = \frac{\exp(u_{a,v,\mathcal{C}}(\mathbf{y}, \mathcal{NR}_a))}{\exp(u_{a,v',\mathcal{C}}(\mathbf{y}, \mathcal{NR}_a))}$$

can vary with $\mathcal{C}$. □

By combining preferences and random consideration sets, the probability that Agent a selects (at the moment of choosing) alternative v given choice configuration $\mathbf{y}$, $\mathrm{P}_a(v \mid \mathbf{y})$, has to satisfy

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

Altogether, the elements just described characterize our initial model of peer effects in choices. This model leads to a sequence of decisions of every agent that evolves through time according to a Markov random process.

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

Remark 1. Our identification arguments *only* use variation in the choices of peers. That is, we do not use exogenous variation in observable characteristics (i.e., covariates) or menus. Thus, to simplify the exposition, at this stage, we do not include observable covariates in the model. We can interpret our setting as if we were conditioned on them.

Remark 2. We show in our application that covariates in the data can be easily incorporated into the model for estimation purposes. They could also offer extra sources of identification. In particular, alternative-specific covariates can serve as further exclusion restrictions for the consideration or preference. For example, the product-specific level of advertisement might only affect attention to such a specific product (Goeree, 2008).

Remark 3. The dynamic interaction process we model assumes that each agent best responds to the observed choices of others, and does not attempt to anticipate their actions in the future or how her choice could affect them. Allowing an agent to explicitly think how her choice could affect the decisions of others 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 discuss this possibility in Section 6 as a relevant follow-up with an interesting application in mind.

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 reference groups of Agent

a who select option v in choice configuration $\mathbf{y}$. 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 $\mathrm{NR}_a^{\mathcal{S}}(\mathbf{y}) = (\mathrm{NR}_a^v(\mathbf{y}))_{v \in \mathcal{S}\setminus\{0\}}$ and $\mathrm{NC}_a^{\mathcal{S}}(\mathbf{y}) = (\mathrm{NC}_a^v(\mathbf{y}))_{v \in \mathcal{S}\setminus\{0\}}$ for any $\mathcal{S} \subseteq \mathcal{Y}$. We write nc^v for a possible value of $\mathrm{NC}_a^v(\mathbf{y})$. The first 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) $\mathrm{Q}_a(v \mid \mathbf{y}, \mathcal{NC}_a) > 0$;

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

(iii) $\mathrm{Q}_a(v \mid n+1) / \mathrm{Q}_a(v \mid n)$ is different from both 1 and $\mathrm{Q}_a(v \mid n+2) / \mathrm{Q}_a(v \mid n+1)$ for $n = 0$.

Assumption 2(i) states that the probability of considering each option is strictly positive regardless of how many peers have selected that option. This assumption captures the idea that people 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). Note that we allow alternatives to be considered with probability 1. Hence, capturing a form of persistence in consideration sets (i.e., an alternative can be considered with probability 1 several consecutive time periods).[11] 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 mild shape restriction that is satisfied if the consideration probabilities are not constant or exponential functions (i.e., $\ln \mathrm{Q}_a(v \mid \cdot)$ is nonlinear) of the number of peers. Instead of $n = 0$, this restriction can be imposed at any 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.). This third restriction is a variability requirement stating that choices of consideration peers effectively have an effect on consideration probabilities. We do not need the full power of Assumption 2(iii) for all our results. For instance, the identification

[11] In Section 4.2, we allow even richer forms of the evolution of consideration sets by introducing history dependence to the model.

of $\mathcal{N}_a$ and $\mathcal{NR}_a$ would only require the consideration probabilities to vary with the choices of peers. However, the identification of $\mathcal{NC}_a$ requires nonexponential consideration probabilities.

Example 1 (continued)**.** Suppose that the mean attention of Agent a, $c_{v,a}$, is such that

$$c_{v,a}(\mathbf{y},\mathcal{NC}_a) = \mathbb{1}\left(\mathrm{NC}_a^v(\mathbf{y}) > 0\right).$$

As a result,

$$\mathrm{Q}_a(v|\mathbf{y},\mathcal{NC}_a) = F_\varepsilon\left(\mathbb{1}\left(\mathrm{NC}_a^v\left(\mathbf{y}\right) > 0\right)\right),$$

and, if $F_\varepsilon(1) > F_\varepsilon(0) > 0$, then Assumption 2 is satisfied. □

Let $(\mathbf{0})_v^1$ denote the vector obtained by replacing the v-th component of the zero vector by 1. The second 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 $v \in \mathcal{C} \setminus \{0\}$, we have that

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

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

(iii) $\mathrm{R}_a\left(v \mid (\mathbf{0})_v^1,\mathcal{C}\right) - \mathrm{R}_a\left(v \mid \mathbf{0},\mathcal{C}\right) \neq 0$ 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 is observed with a positive probability. 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 peer selecting it 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. 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 since. Indeed, the signs of all these effects can be recovered from the data.

Example 2 (continued)**.** Suppose that the mean utility of Agent a from alternative v given the consideration set $\mathcal{C}$, $u_{a,v,\mathcal{C}}$, is such that

$$u_{a,v,\mathcal{C}}\left(\mathbf{y},\mathcal{NR}_a\right) = \bar{u}_{a,v,\mathcal{C}}\left(\mathrm{NR}_a^v\left(\mathbf{y}\right)\right),$$

where $\bar{u}_{a,v,\mathcal{C}}(\cdot)$ is a function that satisfies $\bar{u}_{a,v,\mathcal{C}}(0) \neq \bar{u}_{a,v,\mathcal{C}}(1)$. Then Assumptions 3 is satisfied. □

We extend the model in Section 4.2 to allow the dependence of Q_a and R_a on the current or past choices of Agent a (e.g., a Markov process with memory), thus capturing environments with inertia by setting the probability of considering the current choice with probability 1. We write the assumption in a stricter way here only to simplify the exposition.

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

Assumption 4 (Network)**.** For each $a \in \mathcal{A}$, 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 simultaneously affects consideration and preferences, then the agent also has at least another peer that affects either only consideration or only preferences. The choice of such peer only enters either the consideration probability or the choice rule and so provides an exclusion restriction. Note that we do not need the two sets of agents to be nonempty. Indeed, if only one of the peer-effect mechanisms operates in practice, our results would allow us to state whether the interdependencies in choices of agents are due to the peer effect in preferences or in consideration. Moreover, this assumption does not rule out agents with no peers at all. (This happens in our empirical application. We will state there how to proceed when some agents have no links.)

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 at each moment of 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. The advantage of this fact for model identification is that there are fewer terms to recover. Blevins (2017, 2018) offer a nice discussion of this feature and its advantage over discrete time models. (In Section 4.5, we extend the idea to coordinated clocks.) Formally, the transition rate from choice configuration $\mathbf{y}$ to any different one $\mathbf{y}'$ is as follows

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

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 simply build from these other values as follows

$$\mathrm{m}\left(\mathbf{y} \mid \mathbf{y}\right) = -\sum\nolimits_{\mathbf{y}' \in \mathcal{Y}^A \setminus \{\mathbf{y}\}} \mathrm{m}\left(\mathbf{y}' \mid \mathbf{y}\right).$$

We indicate by $\mathcal{M}$ the transition rate matrix. In our model, the number of possible choice configurations is $(Y+1)^A$. Thus, $\mathcal{M}$ is a $(Y+1)^A \times (Y+1)^A$ matrix. There are many different ways of ordering the choice configurations and thereby writing the transition rate matrix. To avoid any sort of ambiguity in the exposition, we let the choice configurations be ordered according to the lexicographic order. Constructed in this way the first element of $\mathcal{M}$ is $\mathcal{M}_{11} = \mathrm{m}\left((0,0,\ldots,0)' \mid (0,0,\ldots,0)'\right)$. 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,

$$\mathcal{M}_{\iota(\mathbf{y})\iota(\mathbf{y}')} = \mathrm{m}\left(\mathbf{y}' \mid \mathbf{y}\right).$$

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 $\mathcal{M}$. It indicates the likelihood of each choice configuration $\mathbf{y}$ in the long run, i.e., the stationary distribution. This equilibrium behavior relates to the transition rate matrix in a linear fashion

$$\mu \mathcal{M} = \mathbf{0}.$$

The next proposition establishes equilibrium existence and uniqueness for our model.

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

Proof. Given the simplicity of the result, we prove it here. For an irreducible, finite-state, continuous Markov chain, the equilibrium μ exists, and it is unique. Thus, we only need to prove that Assumptions 2(i) and 3(i) imply that the Markov chain induced by our model is irreducible. Note that

$$\mathrm{P}_a(v \mid \mathbf{y}) = \sum\nolimits_{\mathcal{C}\subseteq\mathcal{Y}} \mathrm{R}_a(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C})\, \mathrm{C}_a(\mathcal{C} \mid \mathbf{y}, \mathcal{NC}_a, \mathcal{Y}).$$

Assumption 2(i) implies that given any $\mathbf{y}$, any v is always considered with a positive probability by any Agent a. Assumption 3(i) then implies that any option is picked with a positive probability if considered. Thus, $0 < \mathrm{P}_a(v \mid \mathbf{y}) < 1$ for all a and $\mathbf{y}$, and we can go from one configuration to the other one in less than A steps with a positive probability. ■

It is important to remark that under our restrictions the equilibrium distribution has full support on $\mathcal{Y}^A$. The full support feature allows identification of all parts of the model. If this result fails, then identification of the model may still be achieved, but it requires some extra restrictions.

3. Empirical Content of the Model

This section shows that under specific restrictions, 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 in $\mathcal{A}$.

We separate the identification analysis in two parts. Let $\mathrm{P} = (\mathrm{P}_a)_{a\in\mathcal{A}}$ be the profile of conditional choice probabilities (CCPs) of agents in the network. Each $\mathrm{P}_a(v \mid \mathbf{y}) : \mathcal{Y} \times \mathcal{Y}^A \to (0, 1)$ specifies the (ex-ante) probability that Agent a selects option v when the choice configuration is $\mathbf{y}$. Under our main assumptions, we have that

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

where

$$\mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}_a^{\mathcal{Y}}(\mathbf{y}), \mathcal{Y}\right) = \prod\nolimits_{v \in \mathcal{C}} \mathrm{Q}_a\left(v \mid \mathrm{NC}_a^{v}(\mathbf{y})\right) \prod\nolimits_{v \in \mathcal{Y} \setminus \mathcal{C}} \left(1 - \mathrm{Q}_a\left(v \mid \mathrm{NC}_a^{v}(\mathbf{y})\right)\right).$$

We first show how to recover the set of connections, choice rules, and consideration probabilities from each set of CCPs P. We then build identification of the CCPs P from a sequence of observed choices.

3.1. Identification of the Model From P

The identification strategy we offer is constructive. We start by recovering the network structure, which is achieved in three stages. First, we recover the reference group of every agent. Second, we recover 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 the consideration probabilities and the random preferences.

Network Note that, under Assumptions 1 - 3, changes in the choices of peers induce variation in the CCPs. To see this, note that P_a can be rewritten as

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

In words, the observed probability that v is picked is equal to the product of the probability that it is considered and the probability that it is picked when considered. Moreover, the first term $\mathrm{Q}_a\left(v \mid \mathrm{NC}_a^{v}(\mathbf{y})\right)$ only depends on $\mathrm{NC}_a^{v}(\mathbf{y})$, and $\mathrm{NC}_a^{v}(\mathbf{y})$ does not affect the second term. These two observations allow us to use certain changes in logarithms of P_a to identify the network.

Let $\Delta_{a'}^{v}$ be a linear operator that indicates the increment of a given function when the action of

Agent a changes to v in the action configuration $\mathbf{y}$. Formally, given $f : \mathcal{Y}^A \to \mathbb{R}$, let

$$\Delta_{a'}^{v} f(\mathbf{y}) = f(\mathbf{y}_{a'}^{v}) - f(\mathbf{y}),$$

where $\mathbf{y}_{a'}^{v}$ 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, by computing the implied difference in the logarithms of P_a, we get that

$$\Delta_{a'}^{v} \ln \mathrm{P}_a (v \mid \mathbf{0}) \equiv \ln \mathrm{P}_a (v \mid \mathbf{0}_{a'}^{v}) - \ln \mathrm{P}_a (v \mid \mathbf{0}) =$$
$$\Delta_{a'}^{v} \ln \mathrm{Q}_a (v \mid \mathrm{NC}_a^{v}(\mathbf{0})) + \Delta_{a'}^{v} \ln \sum\nolimits_{\mathcal{C} \subseteq \mathcal{Y} \setminus \{v\}} \mathrm{R}_a \left(v \mid \mathrm{NR}_a^{\mathcal{C} \cup \{v\}}(\mathbf{0}), \mathcal{C} \cup \{v\}\right) \mathrm{C}_a \left(\mathcal{C} \mid \mathrm{NC}_a^{\mathcal{Y} \setminus \{v\}}(\mathbf{0}), \mathcal{Y} \setminus \{v\}\right), \tag{2}$$

where $\mathbf{0} = (0, 0, \ldots, 0)'$ denotes the configuration where everyone picks the default. Each term in Equation (2) relates to one (and only 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, $\Delta_{a'}^{v} \ln \mathrm{P}_a (v \mid \mathbf{0}) \neq 0$ 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 of this result requires a "regularity condition" (Assumption 6) that we discuss in details in Appendix A. This extra condition rules out the possibility that peer effects in consideration and preferences be of opposite signs *and* of equal magnitude. Under these conditions, it follows that the reference groups (even if they are empty) can be recovered from the CCPs.

Proposition 3.1. *Suppose Assumptions 1, 2, 3, and 6 hold. Then, for any $a \in \mathcal{A}$, $\mathcal{N}_a$ is identified.*

We now proceed to identify whether Agent a' in $\mathcal{N}_a$ affects Agent a's preferences or consideration only. Note that by analyzing the differences of logarithms of P_a, we can identify the reference group of each agent, 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 (2) 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 the double difference as follows

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

Specifically, we have that

$$\begin{aligned}\Delta_{a''}^{w}\Delta_{a'}^{v}\ln \mathrm{P}_a(v \mid \mathbf{0}) =& \Delta_{a''}^{w}\Delta_{a'}^{v}\ln \mathrm{Q}_a(v \mid \mathrm{NC}_a^{v}(\mathbf{0})) +\\ & \Delta_{a''}^{w}\Delta_{a'}^{v}\ln \sum\nolimits_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}_a^{\mathcal{C}\cup\{v\}}(\mathbf{0}), \mathcal{C}\cup\{v\}\right)\mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}_a^{\mathcal{Y}\setminus\{v\}}(\mathbf{0}), \mathcal{Y}\setminus\{v\}\right).\end{aligned}$$

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

$$\Delta_{a'}^{v}\ln \sum\nolimits_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}_a^{\mathcal{C}\cup\{v\}}(\mathbf{0}), \mathcal{C}\cup\{v\}\right)\mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}_a^{\mathcal{Y}\setminus\{v\}}(\mathbf{0}), \mathcal{Y}\setminus\{v\}\right) = 0.$$

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 preferences, then the second summand in Equation (2) will not disappear after switching Agent a'' from v to w. Indeed, under Assumptions 2, 3, and the regularity condition, $a' \in \mathcal{NR}_a$ if and only if

$$\Delta_{a''}^{w}\Delta_{a'}^{v}\ln \sum\nolimits_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}_a^{\mathcal{C}\cup\{v\}}(\mathbf{0}), \mathcal{C}\cup\{v\}\right)\mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}_a^{\mathcal{Y}\setminus\{v\}}(\mathbf{0}), \mathcal{Y}\setminus\{v\}\right) \neq 0.$$

Since the first summand in Equation (2) is 0, then this is the same as to say $\Delta_{a''}^{w}\Delta_{a'}^{v}\ln \mathrm{P}_a(v \mid \mathbf{0}) \neq 0$. 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 (which may or may not affect consideration).

As a final step, we note that the group of preference peers can be further separated into two sets of agents 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

$$\Delta^v_{a'} \ln \mathrm{P}_a (v \mid \mathbf{0}) \neq \Delta^v_{a''} \ln \mathrm{P}_a (v \mid \mathbf{0}).$$

This allows us to separate the preference peers into two groups that we define as $\mathcal{M}'$ and $\mathcal{M}''$ (each of these sets is allowed to be empty). While we know that one of these sets is $\mathcal{NCR}_a$ (without further restrictions), we cannot tell which. We address this last question below.

Proposition 3.2. *Suppose Assumptions 1, 2, 3, and 6 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. Moreover, $\mathcal{NCR}_a \in \{\mathcal{M}', \mathcal{M}''\}$.*

Our last step in recovering the network structure is to identify the set of consideration-preference peers (i.e., $\mathcal{NCR}_a$) from the group of peers that affect preferences. We discuss the 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 consideration-only or preferences-only. 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

$$\Delta^v_{a''}\Delta^v_{a'} \ln \mathrm{P}_a (v \mid \mathbf{0}) = \Delta^v_{a''}\Delta^v_{a'} \ln \mathrm{Q}_a (v \mid \mathrm{NC}^v_a(\mathbf{0})).$$

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

Suppose next that there is no consideration-only peer. We can still implement a similar idea by replicating the consideration-only peer behavior with two peers, one that affects consideration and preferences and the other one that affects only preferences. Notice that these two peers can be identified by Proposition 3.2. Pick some Agent $a' \in \mathcal{M}'$ and Agent $a'' \in \mathcal{M}''$. Next, take $\mathbf{y}$ such that $y_a = 0$ for all $a \neq a'$ and $y_{a'} = v$. Note that

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

Thus, this operation with two peers, one from $\mathcal{NCR}_a$ and the other one from $\mathcal{NR}_a \setminus \mathcal{NC}_a$, is equivalent (up to sign) to switching one peer from $\mathcal{NC}_a \setminus \mathcal{NR}_a$. Finally, take another Agent a''' in either $\mathcal{M}'$ or $\mathcal{M}''$ and proceed as we did before —when there was a peer in the consideration-only group. In doing so, we identify whether Agent a''' is a consideration-preference or preference-only peer. (Note that this procedure requires to have at least three peers in $\mathcal{N}_a$.) This information allows us to know whether $\mathcal{NCR}_a = \mathcal{M}'$ or $\mathcal{NCR}_a = \mathcal{M}''$.

Proposition 3.3. *Suppose Assumptions 1, 2, 3, 4, and 6 hold. Suppose also that $\mathcal{NC}_a \setminus \mathcal{NR}_a$ is identified (or known) and $|\mathcal{N}_a| \geq 3 - |\mathcal{NC}_a \setminus \mathcal{NR}_a|$. Then, $\mathcal{NC}_a$ and $\mathcal{NR}_a$ are identified.*

The last proposition offers final conditions for all the parts of the network structure to be identified. It takes as an input the consideration-only set $\mathcal{NC}_a \setminus \mathcal{NR}_a$, which is allowed to be empty. It states that if we know or identify the consideration-only peers, the full network structure is identified when there are enough peers. Moreover, the result holds even for binary settings (i.e., $Y = 1$). However, we still need $Y \geq 2$ to identify $\mathcal{NC}_a \setminus \mathcal{NR}_a$ using Proposition 3.2. We discuss the binary case in detail in Section 4.1.

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

Consideration mechanisms and Random Preferences We first state that if we know the network structure, and each agent has at least one consideration-only peer —or such peer can be constructed from consideration-preference and preference-only peers, as we do above— then we can recover ratios of probabilities of considering alternatives. To show this claim, let $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$. Then, it is easy to see that 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

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

Thus, we can identify

$$\mathrm{Q}_a\left(v \mid 1\right) / \, \mathrm{Q}_a\left(v \mid 0\right).$$

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 $a' \in \mathcal{NR}_a \setminus \mathcal{NC}_a$ and $a'' \in \mathcal{NCR}_a$. Then, if we switch a' from v to the default and a'' from the default to v, then the changes in P_a are driven by changes in consideration probabilities only. That is, for $\mathbf{y}$ such that $y_{\tilde{a}} = 0$ for all $\tilde{a} \neq a'$ and $y_{a'} = v$,

$$\ln \mathrm{P}_a\left(v \mid (\mathbf{y}^v_{a''})^0_{a'}\right) - \ln \mathrm{P}_a\left(v \mid \mathbf{y}\right) = \ln \mathrm{Q}_a(v \mid 1) - \ln \mathrm{Q}_a(v \mid 0),$$

so the ratio of consideration probabilities is identified. By applying the same ideas to different initial configurations, we can identify ratios of consideration probabilities as we formally state next.

Proposition 3.4. *Suppose $\mathcal{NC}_a$ and $\mathcal{NR}_a$ are known and Assumptions 1, 2, 3, and 4 hold. Then*

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

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 4. 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 completely dropped. Indeed, by definition, we have that

$$\mathrm{P}_a(v \mid \mathbf{y}) = \mathrm{Q}_a(v \mid \mathrm{NC}^v_a(\mathbf{y}))\,\mathrm{Pr}_a(v \mid \mathbf{y}, v \text{ is considered}),$$

where the second term is the conditional probability that v is picked by Agent a conditional on being considered. Thus, the variation in choices of $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$ 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 in general does not convey information about the probability of several items being considered simultaneously.

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

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

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 $\mathcal{Y} \setminus \mathcal{Z}$. That is, P_a^* 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 probability is the observed probability of Agent a choosing alternative 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 $\mathrm{P}_a^*(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\})$, we obtain that

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

where $t_{v'} = \mathrm{Q}_a\left(v' \mid 1\right) / \mathrm{Q}_a\left(v' \mid 0\right) \neq 1$ can be identified using Proposition 3.4. It follows that we can recover the counterfactual CCP $\mathrm{P}_a^*(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\})$ 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. In other words, we can exclude any set of nondefault alternatives as long as its cardinality is less than or equal to $|\mathcal{NC}_a \setminus \mathcal{NR}_a|$.

The next result formalizes and extends this argument.

Proposition 3.5. *Suppose $\mathcal{NC}_a$ and $\mathcal{NR}_a$ are known, and Assumptions 1, 2, and 3 are satisfied. Then $\mathrm{P}_a^*\left(v \mid \mathbf{y}, \mathcal{Y} \setminus \mathcal{Z}\right)$ 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 allows us to address an important counterfactual prediction. Namely, what would happen if some alternatives are removed or become unavailable. Notice that the identification of these counterfactual CCPs does not require identifying either Q_a or R_a. Indeed, we only use the identification of the ratios of Q_as.

In our setting, the variation in the choices of consideration-only peers is equivalent to menu variation in the stochastic choice literature. In particular, if one has enough of those peers, then one can identify the counterfactual CCPs for binary menus

$$\mathrm{P}_a^*(v \mid \mathbf{y}, \{0, v\}) = \mathrm{Q}_a\left(v \mid \mathrm{NC}_a^v(\mathbf{y})\right) \mathrm{R}_a\left(v \mid \mathrm{NR}_a^v(\mathbf{y}), \{0, v\}\right).$$

Hence, if either $\mathrm{Q}_a\left(v \mid \mathrm{NC}_a^v(\mathbf{y})\right)$ or $\mathrm{R}_a\left(v \mid \mathrm{NR}_a^v(\mathbf{y}), \{0, v\}\right)$ is known, we can recover $\mathrm{Q}_a(v \mid \cdot)$ (by using Proposition 3.4) and then $\mathrm{R}_a\left(v \mid \mathrm{NR}_a^v(\mathbf{y}), \{0, v\}\right)$ from our recent ideas. Applying the same argument for menus of size three, we can identify R_a for sets of size three, and so on.

Proposition 3.6. *Suppose that 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* $\mathrm{Q}_a(v \mid n_1)$ *or* $\mathrm{R}_a(v \mid n_2, \{0, v\})$ *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 $\mathrm{Q}_a(v \mid n_1)$ or $\mathrm{R}_a(v \mid n_2, \{0, v\})$ 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., $\mathrm{R}_a(v \mid |\mathcal{NR}_a|, \{0, v\}) = 1$). Another example is when the alternative is considered with probability 1 if enough (or all) consideration peers pick the alternative (i.e., $\mathrm{Q}_a(v \mid |\mathcal{NC}_a|) = 1$).

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. These two datasets coincide in that they contain long sequences of choices from agents in the network. They differ in the timing at which the researcher observes these choices (including the default). In Dataset 1, agents' choices are observed in real-time. This allows the researcher to record the precise moment at which an agent revises her strategy and the configuration of choices at that time. In Dataset 2, the researcher simply observes the joint 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 $\mathcal{P}(\Delta)$.[12] The connection between $\mathcal{P}(\Delta)$ and the transition rate matrix $\mathcal{M}$ described

[12]Here again, we assume that the choice configurations are ordered according to the lexicographic order when we

in Section 2.3 is given by

$$\mathcal{P}(\Delta) = e^{(\Delta\mathcal{M})},$$

where $e^{(\Delta\mathcal{M})}$ is the matrix exponential of $\Delta\mathcal{M}$. The two datasets we consider differ regarding Δ: in Dataset 1, we let the time interval be very small. This is an ideal dataset that registers agents' choices at the exact time at which any given agent revises her choice. As we mentioned earlier, with the proliferation of online platforms and scanners, this sort of data might indeed be available for some applications. In Dataset 2, we allow the time interval to be of arbitrary size. The next statements formally describe what we can identify regarding the transition probabilities from Datasets 1 and 2. That is,

Dataset 1 The researcher knows $\lim_{\Delta\to 0}\mathcal{P}(\Delta)$,

Dataset 2 The researcher knows $\mathcal{P}(\Delta)$.

In both cases, the identification question is whether (or under what extra restrictions) it is possible to uniquely recover $\mathcal{M}$ from the transition probabilities in $\mathcal{P}(\Delta)$, which are identified and estimated from the data directly. The first identification result concerns Dataset 1 and is as follows.

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 relies on the fact that when the time interval between the observations goes to zero, then we can recover $\mathcal{M}$.

There are at least two well-known cases that produce the same identification result without assuming $\Delta \to 0$. One of them requires the length of the interval Δ to be below a threshold $\overline{\Delta}$. The main difficulty of this identification 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 next proposition, based on Theorem 1 in Blevins (2018), states that the transition rate

construct $\mathcal{P}(\Delta)$.

matrix can be identified from agents' choices even if these choices are observed at the endpoints of discrete time intervals.

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

The restriction on eigenvalues of $\mathcal{M}$ is a regularity condition that is generically satisfied.[13] The key element in proving Proposition 3.8 is that the transition rate matrix in our model is rather parsimonious. To see why, recall that, at any given time, only one person revises her selection with a nonzero probability. This feature of the model translates into a transition rate matrix $\mathcal{M}$ that has many zeros in known locations.

4. Extensions

In this section, we provide several extensions of our baseline model.

4.1. Binary Choice

The identification of consideration-only peers —Proposition 3.2— requires the existence of at least two nondefault alternatives (i.e., $Y \geq 2$). This condition cannot be relaxed without extra assumptions. Indeed, if $\mathcal{Y} = \{0, 1\}$, the choices of Agent a are completely described by one equation:

$$\mathrm{P}_a(1 \mid \mathbf{y}) = \mathrm{Q}_a(1 \mid \mathrm{NC}_a^1(\mathbf{y}))\, \mathrm{R}_a(1 \mid \mathrm{NR}_a^1(\mathbf{y}), \{0, 1\}).$$

Thus, although we still can recover the set $\mathcal{N}_a$, we cannot implement double switches across alternatives to distinguish whether a peer affects only consideration or preferences. To identify the different types of peers we need to add some extra restrictions.

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

The main result is based on a stronger version of Assumption 4.

Assumption 4′. For each $a \in \mathcal{A}$, if $|\mathcal{NCR}_a| \geq 1$, then $|\mathcal{NC}_a \setminus \mathcal{NR}_a|\,|\mathcal{NR}_a \setminus \mathcal{NC}_a| \geq 1$.

This assumption requires that if there are consideration-preference peers, then there must be consideration-only and preference-only peers. Note that as with Assumption 4, if $\mathcal{NCR}_a$ is empty, then no restrictions are imposed.

Proposition 4.1. *Suppose that $Y = 1$ and Assumptions 1, 2, 3, 4′, and 6 hold. If $\mathcal{NC}_a$ or $\mathcal{NR}_a$ is known, then the network structure is identified.*

Remark 5. The same result can be obtained if, instead of restricting the network structure, we add a fourth condition to the assumption on preferences —i.e., Assumption 3: (iv) For $\mathcal{Y} = \{0, 1\}$, $\mathrm{R}(1 \mid 1, \{0, 1\})/\,\mathrm{R}(1 \mid 0, \{0, 1\})$ is different from $\mathrm{R}(1 \mid 2, \{0, 1\})/\,\mathrm{R}(1 \mid 1, \{0, 1\})$. The extra condition is the analog of Assumption 2(iii) —on consideration— to the choice rule.

Remark 6. The assumption that the researcher knows either $\mathcal{NC}_a$ or $\mathcal{NR}_a$ in Proposition 4.1 can be substituted by a sign restriction. Throughout the paper, we sustained the assumption that the researcher does not know the signs of the peer effects in preferences and consideration. Indeed, we offered conditions under which these signs can be recovered from the data. But, in some cases, it might be reasonable to think that the signs of these effects are known. If this were the case, and peer effects in consideration and preferences were of opposite signs, then we could dispense with the assumption that the researcher knows either $\mathcal{NC}_a$ or $\mathcal{NR}_a$ in Proposition 4.1. An interesting example, where this sign restriction could be used, is the case of vaccines. One could argue that a person becomes aware of a particular vaccine if more of her friends are getting shots. Thus, the peer effect in consideration is positive. Moreover, we could also argue that the peer effect in preferences is negative. The reason is that if more of her peers get vaccinated, then the chances of getting sick reduce, and this reduces the payoff of getting the vaccine. Indeed, in a different model, a similar idea has been used by Agranov, Elliott and Ortoleva (2021) to explain some data on COVID-19 vaccine uptake.

After the network structure is identified, the identification of the consideration mechanisms and preferences follows directly from Proposition 3.6.

4.2. History Dependence and Own Past Choices

Along the paper, we assume that (at the moment of choosing) agents take into account only aggregate information about the choices of peers and ignore their own past choices. In this section, we relax these modelling restrictions.

Suppose that both the consideration and preferences of a given agent depend on the history of her own choices and those of her peers. Formally, let $\{t_k\}_{k=1}^{+\infty}$ be a random (increasing) sequence of 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). Given the history of choice configurations $h_t = (\mathbf{y}_{t_k})_{t_k<t}$, the probability that alternative v is picked by Agent a at time t would be

$$\begin{aligned}\mathrm{P}_a\left(v \mid \mathbf{y}_t, h_t\right) = & \sum\nolimits_{\mathcal{C}\subseteq\mathcal{Y}} \mathrm{R}_a\left(v \mid \mathbf{y}_t, h_t, \mathcal{NR}_a, \mathcal{C}\right) \\ & \prod\nolimits_{v'\in\mathcal{C}} \mathrm{Q}_a\left(v' \mid \mathbf{y}_t, h_t, \mathcal{NC}_a\right) \prod\nolimits_{v'\in\mathcal{Y}\setminus\mathcal{C}} \left(1-\mathrm{Q}_a\left(v' \mid \mathbf{y}_t, h_t, \mathcal{NC}_a\right)\right).\end{aligned}$$

These CCPs depend on the past choices but not on the exact times at which these choices were made. Also, note that none of our previous results use variation beyond the choices of peers at the moment of making a decision. Hence, if we condition on the choice of 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 by using our previous ideas —thus, we omit the proof of the next result.

Proposition 4.2. *Suppose that Assumptions 1, 2, 3, 4, and 6 are satisfied conditional on y_{at} and the history h_t for all possible y_{at} and h_t. Also, extend the definition of P_a^* to allow for dependence on y_{at} and the history h_t. Then, the conclusions of all propositions from Section 3.1 are still valid.*

Since we allow consideration probabilities to be equal to 1, the dependence on past choices allows nontrivial dynamics in consideration sets. For example, the consideration sets may not change much over long periods of time. Proposition 4.2 takes as input the CCPs that (now) depend on the histories of choices of everyone in the network. That is, it is assumed that P_a is identified. Since we only observe choices of 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. This can be done by assuming that there exists finite $K \geq 1$ such that for any $k > K$, $(\mathbf{y}_{t_{k'}})_{k'=1,\ldots,k}$, v, and a

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

If there is enough variation in choices, then P can be recovered from Dataset 1.

We finally note that under additional restrictions on how choices made many periods ago affect current choices, the CCPs could be identified even if consideration probabilities and choice rules depend on the whole history of choice. These restrictions usually imply that the impact of the remote past is decaying sufficiently fast with time (see, Härdle, Lütkepohl and Chen, 1997 for examples). Moreover, the past history often enters via an index. For example, in our empirical application, we assume that history enters only via the total number of times an agent picked the alternative in the past. In cases like this, P can be identified using the results in Bierens (1996).

4.3. Short Panel Dataset

The main identification results in the paper require the researcher to observe the choice of agents in the same network many times. As a result, the identification of each agent's CCPs is achieved by exploiting the across-time variation. While such data allows us to work with fully heterogeneous agents (i.e. every agent has her own choice rule and consideration probabilities), it may not be available in some applications. In this subsection, we discuss several identification arguments that can be used with short panel datasets to identify CCPs by pooling information across many networks.

Similar to Lewbel et al. (2023), suppose that there are L fixed networks (i.e., groups or classes). Each network has its own network structure. The researcher observes the choices of every agent in every network for a short time period, but does not know the network structure of each group. We assume that there is an unknown distribution of latent network structure, from which each group's network is drawn independently from other groups (i.e., the networks are i.i.d.).

Let us assume the researcher observes agents' choices at time intervals of length Δ and can estimate the transition probability $\Pr\left(\mathbf{y}^{t+\Delta} = \mathbf{y}' \mid \mathbf{y}^t = \mathbf{y}\right)$ for each pair $\mathbf{y}', \mathbf{y} \in \mathcal{Y}^A$ from pooling data across networks. The transition probability can be consistently estimated when the number of networks L is growing to infinity while the size of the networks is bounded by some universal constant. In contrast to the baseline framework, where the network is fixed, these transition probabilities are generated from a finite mixture of unobserved network structures Γ. Specifically, the observed transition probability can be expressed as

$$\Pr\left(\mathbf{y}^{t+\Delta} = \mathbf{y}' \mid \mathbf{y}^t = \mathbf{y}\right) = \sum_{\Gamma} \Pr\left(\mathbf{y}^{t+\Delta} = \mathbf{y}' \mid \mathbf{y}^t = \mathbf{y}, \Gamma\right) \Pr(\Gamma \mid \mathbf{y}^t = \mathbf{y}).$$

Therefore, we have to first identify the network-specific transition probability $\Pr\left(\mathbf{y}^{t+\Delta} = \mathbf{y}' \mid \mathbf{y}^t = \mathbf{y}, \Gamma\right)$. We do not need to know the identity of Γ, we only need to know that a given network-specific transition probability corresponds to some Γ. When these network specific transition probabilities are identified, one can use the results from Section 3.2 to identify the network-specific CCPs P. Then our main identification results can be used to recover the unknown network structure Γ, the consideration mechanism, and preferences.

To identify $\Pr\left(\mathbf{y}^{t+\Delta} = \mathbf{y}' \mid \mathbf{y}^t = \mathbf{y}, \Gamma\right)$, one can follow the identification results in the finite mixture/measurement error literature (see, for instance, Hall and Zhou, 2003, Hu, 2008, Kasahara and Shimotsu, 2009, Bonhomme, Jochmans and Robin, 2016, Kitamura and Laage, 2018, and references therein). For example, once can exploit the joint distribution for the choices over time, i.e., $\Pr\left(\mathbf{y}^{t+\Delta+\Delta} = \mathbf{y}'', \mathbf{y}^{t+\Delta} = \mathbf{y}', \mathbf{y}^t = \mathbf{y}\right)$, or the presence of excluded covariates that shift the weights of the network, but not the network-specific transition probability. We leave the detailed analysis of this case for future research.

4.4. Nonobservable Default

In many settings, the decision to choose the default alternative often is 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 the empirical application at the end of the paper, 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 (in our empirical application, on average, firms make a decision every day). 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.

4.5. Perfectly Correlated Clocks

In some applications, one may need to model a situation where two or more connected agents have perfectly synchronized clocks. In this case, they make decisions simultaneously. Given the boundedly rational nature of limited consideration, we assume that in these cases, the agents are unaware that their clocks are perfectly correlated and make a decision without taking into account the strategic consideration. This allows us to avoid complications caused by a potential multiplicity of equilibria. As a result, one would need to treat synchronized agents as one agent and redefine their choice set to consist of all pairs of alternatives from the original choice set.

We assume that each agent behaves as if she is unaware of the synchronized clocks. Therefore, the identification of the CCPs with Dataset 1, the network structure, the consideration probabilities, and the choice rules are the same as the baseline result. If the choice of default is not observed in the data, one then needs to normalize the arrival rate for all agents, i.e., $\lambda_a = 1$.

4.6. Bundles

Introducing small modifications, our framework can be extended to cover some bundle models.[14] Suppose that when agents face a consideration set $\mathcal{C}$, they are allowed to pick more than one option

[14] See, for instance, Gentzkow (2007), Dunker, Hoderlein and Kaido (2017), Fox and Lazzati (2017), Iaria and Wang (2020), Allen and Rehbeck (2022), Kashaev (2023), and Wang (2023).

from it. Define

$$\mathcal{B}(\mathcal{C}) = \{b \subseteq \mathcal{C} \setminus \{0\} \,:\, b \neq \emptyset\} \cup \{0\}.$$

That is, $\mathcal{B}(\mathcal{C})$ is the collection of all possible bundles from consideration set $\mathcal{C}$. Note that the default cannot be bundled with other options. In this case, we only need to extend the definition of R_a from $\mathcal{C}$ to $\mathcal{B}(\mathcal{C})$. A bundle (i.e., a representative element of $\mathcal{B}(\mathcal{C})$) will be denoted by b. Let $\mathrm{R}_a\left(b \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{B}(\mathcal{C})\right)$ denote the probability that bundle b is picked from $\mathcal{B}(\mathcal{C})$. We also need to slightly modify Assumption 3 and the regularity condition (Assumption 6 in Appendix A).

Assumption 3′. For each $a \in \mathcal{A}$, $\mathbf{y} \in \mathcal{B}(\mathcal{Y})^A$, $\mathcal{C} \subseteq \mathcal{Y}$, and $b \in \mathcal{B}(\mathcal{C}) \setminus \{0\}$, we have that

(i) $\mathrm{R}_a\left(b \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{B}(\mathcal{C}^*)\right) > 0$ for some $\mathcal{C}^*$ such that $\mathrm{C}_a(\mathcal{C}^* | \mathbf{y}, \mathcal{NC}_a) > 0$;

(ii) $\mathrm{R}_a\left(b \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{B}(\mathcal{C})\right) \equiv \mathrm{R}_a\left(b \mid \mathrm{NR}_a^{\mathcal{B}(\mathcal{C})}\left(\mathbf{y}\right), \mathcal{B}(\mathcal{C})\right)$; and

(iii) $\mathrm{R}_a\left(v \mid (\mathbf{0})_v^1, \mathcal{B}(\mathcal{C})\right) - \mathrm{R}_a\left(v \mid \mathbf{0}, \mathcal{B}(\mathcal{C})\right) \neq 0$ and its sign does not depend on $\mathcal{C}$.

Assumption 3′ and the modified regularity condition (Assumption 6′) only differ from the original conditions in terms of the domain: the new conditions are defined on the set of all possible bundles $\mathcal{B}(\mathcal{Y})$. The meaning of each restriction is exactly as before.

The following proposition establishes the validity of some of the previous results under the above modifications. Its proof is omitted since it directly follows from the proofs of Propositions 2.1-3.4.

Proposition 4.3. *The conclusions of Propositions 2.1-3.4 are valid for the bundles model if Assumptions 3 and 6 are replaced by Assumptions 3′ and 6′.*

The analogs of Propositions 3.5 and 3.6 for bundles can be established with a small modification of the definition of P_a^*. Define

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

for each $\mathcal{Z} \subseteq \mathcal{Y} \setminus \{0\}$. That is, $\mathrm{P}_a^*\left(b \mid \mathbf{y}, \mathcal{B}(\mathcal{Y} \setminus \mathcal{Z})\right)$ is the counterfactual probability of selecting bundle b under choice configuration $\mathbf{y} \in \mathcal{B}(\mathcal{Y})^A$ when we restrict the set of available options or the menu from $\mathcal{Y}$ to $\mathcal{Y} \setminus \mathcal{Z}$. (The set of available bundles changes from $\mathcal{B}(\mathcal{Y})$ to $\mathcal{B}(\mathcal{Y} \setminus \mathcal{C})$.)

Proposition 4.4. *Suppose $\mathcal{NC}_a$ and $\mathcal{NR}_a$ are known, and Assumptions 1, 2, and 3′ are satisfied. Then $\mathrm{P}^*_a(b \mid \mathbf{y}, \mathcal{B}(\mathcal{Y} \setminus \mathcal{Z}))$ is identified from P_a for each $\mathcal{Z} \subseteq \mathcal{Y} \setminus \{0\}$ 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 peers in $\mathcal{NC}_a \setminus \mathcal{NR}_a$ are 0.*

Proposition 4.5. *Suppose that the assumptions of Proposition 4.4 are satisfied. If, in addition, we have that $|\mathcal{NC}_a \setminus \mathcal{NR}_a| \geq Y - 1$ and either $\mathrm{Q}_a(v \mid n_1)$ or $\mathrm{R}_a(v \mid n_2, \{0, v\})$ is known for some n_1 and n_2 in the support and for each $v \in \mathcal{Y} \setminus \{0\}$, then Q_a and R_a are identified from P_a.*

The proofs of these two results follow directly from the proofs of Propositions 3.5 and 3.6 and are thereby omitted.

4.7. More General Consideration Mechanism

The identification results we presented do not use any exogenous variation in observed covariates. These results only rely on the variation of choices of different types of peers to recover the different parts of the model. In this section, we show that when covariates (with large support) that affect only preferences are available in the data, then (under minimal restrictions on R_a) we can use them to identify a very general model of consideration. In particular, we can fully drop Assumption 1.

Assume that, in addition to the variation of choice of peers, we observe a vector of covariates w, supported on W, that only affects preferences. In particular, assume that given a consideration set, the choices of agents are consistent with the additive random utility model. That is, for all v and $\mathcal{C} \neq \emptyset$ such that $v \in \mathcal{C}$

$$\mathrm{R}_a(v \mid w, \mathbf{y}, \mathcal{NR}_a, \mathcal{C}) = \int \prod_{v' \in \mathcal{C} \setminus \{v\}} \mathbb{1}\left(U_a(v|w, \mathbf{y}, \mathcal{NR}_a) + \varepsilon_v \geq U_a(v'|w, \mathbf{y}, \mathcal{NR}_a) + \varepsilon_{v'} \right) dF_{a,\varepsilon}(\varepsilon \mid \mathbf{y}),$$

where $F_{a,\varepsilon}$ is the agent specific distribution of shocks $\varepsilon = (\varepsilon_v)_{v \in \mathcal{Y}}$ and U_a is the mean utility function.

Assumption 5. For every $a \in \mathcal{A}$ and $v \in \mathcal{Y} \setminus \{0\}$, there exist a known covariate $w_{a,v}$ and a function of it $f_{a,v}(\cdot)$ such that (i) $\mathrm{U}_a(v'|\mathbf{y}, w)$ does not depend on $w_{a,v}$ for all $v' \neq v$; and (ii) the closure of the conditional support of $w_{a,v}$ conditional on all other covariates contains a point $\overline{w}_{a,v}$ such that

$$\lim_{w_{a,v}\to\overline{w}_{a,v}} \mathrm{U}_a(v|w,\mathbf{y},\mathcal{NR}_a) = +\infty,$$
$$\lim_{w_{a,v}\to\overline{w}_{a,v}} \frac{\mathrm{U}_a(v|w,\mathbf{y},\mathcal{NR}_a)}{f_{a,v}(w_{a,v})} = O(1).$$

This assumption requires the existence of an alternative specific covariate that can make this nondefault alternative to be picked every time it is considered. (This covariate could or could not be agent-specific as well.) The presence of such covariate essentially serves as an exclusion restriction in the utility function. Moreover, the local behavior of the utility function in the neighborhood of the extreme point $\overline{w}^p_{a,v}$ is known and is captured by $f_{a,v}(\cdot)$. Assumption 5 is satisfied if, for instance,

$$\mathrm{U}_a(v|w,\mathbf{y},\mathcal{NR}_a) = f_{a,v}(w_{a,v}) + g_{a,v}(\mathbf{y},\mathcal{NR}_a)$$

with $\{(f_{a,v}(w_{a,v}))_{v\in\mathcal{Y}\setminus\{0\}} : w \in W\}$ is equal to $\mathbb{R}^Y_+$. Note that $f_{a,v}$ does not have to vary over the whole Euclidean space.

The next result establishes nonparametric identification of C_a without assuming that consideration sets are formed by independent draws from $\mathcal{Y}$.

Proposition 4.6. *If Assumption 5 is satisfied and* C_a *does not depend on* w*, then* C_a *is identified for all* $a \in \mathcal{A}$*.*

The proof of Proposition 4.6 is based on the idea that we can send the mean utilities of alternatives to positive infinity at different rates. As a result, the observed distribution over choices would correspond to choices of an agent with deterministic preferences (i.e., $\mathrm{R}_a(v \mid w,\mathbf{y},\mathcal{NR}_a,\mathcal{C}) \in \{0,1\}$). In other words, particular limits of observed P_a correspond to a model with random consideration only. For instance, if given deterministic preferences are such that alternative v is picked only when nothing else (except the default) is considered (i.e., v is the worst alternative after the default), then

$$\mathrm{P}_a(v \mid w,\mathbf{y}) = \mathrm{C}_a(\{0,v\} \mid \mathbf{y}),$$

and we identify $\mathrm{C}_a(\{0,v\} \mid \mathbf{y})$. Since we can use different rates for different alternatives, we can identify $\mathrm{C}_a(0, v)$ for all v. Repeating the above argument for consideration sets of bigger sizes, we can recursively identify $\mathrm{C}_a(\mathcal{C})$ for all $\mathcal{C}$.

Next, note that

$$\mathrm{P}_a\left(v \mid w, \mathbf{y}\right) = \textstyle\sum_{\mathcal{C} \subseteq 2^{\mathcal{Y}}} \mathrm{R}_a\left(v \mid w, \mathbf{y}, \mathcal{NR}_a, \mathcal{C}\right) \mathrm{C}_a\left(\mathcal{C} \mid \mathbf{y}, \mathcal{NC}_a\right),$$

where $\mathrm{C}_a\left(\mathcal{C} \mid \mathbf{y}, \mathcal{NC}_a\right)$ has been already identified. Thus, if there are peers that only affect consideration, we can use the known variation in consideration probabilities C_a to build a system of linear equations in which the unknown parameters are $\{\mathrm{R}_a\left(v \mid w, \mathbf{y}, \mathcal{NR}_a, \mathcal{C}\right)\}_{\mathcal{C} \subseteq \mathcal{Y}, v \in \mathcal{C}}$. Hence, if Agent a has enough peers in $\mathcal{NC}_a \setminus \mathcal{NR}_a$ to generate variation, then $\{\mathrm{R}_a\left(v \mid \mathbf{y}, \mathcal{NR}_a, \mathcal{C}\right)\}_{\mathcal{C} \subseteq \mathcal{Y}, v \in \mathcal{C}}$ can be identified as a solution to a system of linear equations. Formally, for a given $v \in \mathcal{Y} \setminus \{0\}$, let $\{\mathcal{C}_k\}_{k=1}^{2^{Y-1}}$ be the collection of all subsets of $\mathcal{Y}$ that contain v and 0. Also, let $\{n^l\}$ be a set of all nonnegative integer-valued vectors of length Y such that $\sum_{v' \in \mathcal{Y} \setminus \{0, v\}} n_{v'}^l \leq |\mathcal{NC}_a \setminus \mathcal{NR}_a|$ for each l. Then, under Assumption 2, for fixed $a \in \mathcal{A}$, $y_a \in \mathcal{Y}$, and $(\mathrm{NR}_a^v)_{v' \in \mathcal{Y}}$, let matrix B be such that the (l, k)-th element of it is

$$B_{l,k} = \mathrm{C}_a(\mathcal{C}_k \mid n^l).$$

If B has full column rank, then R_a is identified.

5. Application

We finally investigate the effect of limited consideration on the expansion decisions of the two dominant coffee chains in China. The goal of this application is three-fold. 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 understanding regarding the profitability of different markets and, thus, firm behavior. In particular, under limited consideration, the fact that some stores are not opened is not always due to the lack of profitability in those markets but rather limited consideration. In this situation, any government subsidy program that aims to foster competition in a monopolistic market will be less efficient since the increased profitability of the market might be ignored by a boundedly rational firm that does not consider that market. Third, we quantify the direct effect of limited consideration and peer effect in consideration on the dynamics of market structure by conducting two counterfactual exercises. In these exercises we either make firms fully attentive or shut the peer effect in consideration (but not in preferences) down. These types of counterfactual analysis are crucial for evaluating consumer welfare and, thus, informed policy recommendations. Specifically, given the presence of spillover effects across markets, the government might achieve desirable outcomes by subsidizing only a few important markets.

5.1. Coffee Industry Background and Data

China is quickly turning into one of the fastest-growing coffee markets worldwide, with sales increasing from 47 billion yuan in 2015 to 82 billion yuan in 2020 and projected to be 219 billion yuan in 2025.[15] Starbucks, the dominant coffee chain in China, opened its first store in Beijing in 1999 and its $6,000$th store in September 2022, also the $1,000$th store in Shanghai, the first city in the world to pass the milestone.

To avoid complications related to the COVID-19 pandemic, we restrict our sample until the end of 2019. By then, Starbucks had expanded into 176 markets with 4042 stores. The second key player in the Chinese market is Luckin. It registered its first store in Beijing in September 2017 but quickly expanded to 54 markets with 3588 stores by the end of 2019. This allowed Luckin to reach a size similar to that of Starbucks and to become Starbucks's main competitor. Costa and Pacific, two other coffee chains, expand much slower and operate on a much lower scale in the Chinese market (see Figure 1).[16] Thus, we will only focus on the entry and expansion decisions of Starbucks and Luckin.

Starbucks has not accepted franchisee relationships in mainland China since September 2017, so all Starbucks stores in China are directly owned by the company. Luckin focuses mainly on

[15] https://www.statista.com/statistics/1171765/china-coffee-market-size/.

[16] Specifically, Costa and Pacific registered, respectively, their first stores in Shanghai in September 2006 and January 2011. Costa had a presence in 36 markets with 458 stores, while Pacific had expanded into 29 markets with 238 stores by the end of 2019.

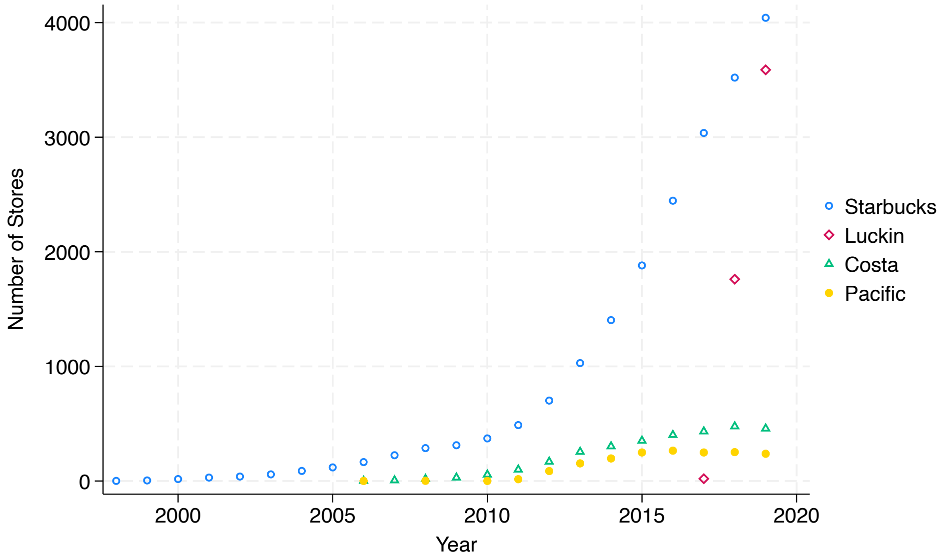


Figure 1 – **Total number of stores of each coffee chain over time.**

the self-operating model. It started a partnership operation in September 2019, but not for its coffee stores. The franchised stores are only allowed for its tea brand, which was separated from the operation of its coffee stores. Therefore, all stores in our data period are operated by Luckin.[17]

5.2. Empirical Model

In this subsection, we describe the model of firm expansion decision, depending on the focal and neighboring market's characteristics, and introduce the specifications for the consideration and payoff. All unknown parameters of the model below will be collected by a vector θ. After we set the model, we formally list all of them.

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 every pair $(f, m) \in \mathcal{F} \times \mathcal{M}$ an agent. Thus, $\mathcal{A} = \mathcal{F} \times \mathcal{M}$ and $\mathcal{Y} = \{0, 1\}$. At the moment of deciding whether to open a new store, the attention that firm f pays to market m depends not only on her own and her competitor's past behavior in market m, but on the past openings choices at "neighboring markets." Formally, $\mathcal{N}_a$ is the set of pairs (f', m') that influence the decision of firm f

[17] https://www.hehuoren.cn/news/dongtai/1/160.html.

in market m ($a = (f, m)$). Similarly, $\mathcal{NC}_a$ and $\mathcal{NR}_a$ are the sets of pairs of firms and markets that affect consideration and preferences, respectively, of firm f in market m.

Since $Y = 1$ in this setting, we cannot identify all the components of the network structure without additional assumptions (see Section 4.1 for a detailed discussion). We rely on the competition feature to first recover the peers that only affect payoffs in the market. Specifically, we follow the literature (e.g., Arcidiacono, Bayer, Blevins and Ellickson, 2016) and assume that the marginal profit of firm f in market m from opening a new store is only affected by her own and her competitors' choices in market m. That is, only the competitors in the same markets are the peers affecting payoffs (or preferences). Formally,

$$(f', m') \in \mathcal{NR}_{(f,m)} \iff m = m'.$$

One potential concern is that spillover effects across markets on profits might be through shipping cost savings from the distribution chain (see, for instance, Jia, 2008, Holmes, 2011, Zheng, 2016, Houde, Newberry and Seim, 2023). However, coffee chains in China rely on third-party logistics for shipping instead of building their own distribution centers. Moreover, there is evidence that storage and shipping account for only approximately 5% of the total cost expenditure.[18] Therefore, we believe the assumption of no spillover in profit is reasonable.

Assuming that $\mathcal{NR}_a$ is known allows us to use Proposition 4.1 to identify all the rest of the components of the network structure. In particular, we can focus on recovering the consideration network or the scope of neighborhood markets. No restrictions on $\mathcal{NC}_a$ are needed for identification. We will only impose assumptions on $\mathcal{NC}_a$ to facilitate the estimation.

Observable Characteristics Every market m, which is part of $a = (f, m)$ for some given f, 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 $N_t = (N_{at})_{a \in \mathcal{A}}$.

Market Consideration In our application, there are 152 markets where firms can open a new

[18] https://report.iresearch.cn/report_pdf.aspx?id=4032

store. Given how complex the decision of opening a new store is, we allow firms to consider only a subset of markets at the moment of making a decision. Firm f's consideration set is defined as the set of the markets that firm f considers. Market m is considered by firm f based on the following attention index:

$$\tilde{\pi}_{at} = \bar{\tilde{\pi}}_{at}(S_t, N_t; \theta) - \tilde{\varepsilon}_{at},$$

where the $\tilde{\varepsilon}_{at}$s are i.i.d. across a and t with a known c.d.f. $F_{\tilde{\varepsilon}}$ (in our application $F_{\tilde{\varepsilon}}$ is the Logistic c.d.f.). We allow the current market features of market m's (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. Specifically, we parameterize the mean attention index of Agent a for market m as[19]

$$\begin{aligned}\bar{\tilde{\pi}}_{at}(S_t, N_t; \theta) =& S'_{mt}\tilde{\beta}_f + \sum_{f'} \left[\ln\left(1 + N_{(f',m)t}\right)\tilde{\alpha}_{f,f'} + \ln^2\left(1 + N_{(f',m)t}\right)\tilde{\gamma}_{f,f'} \right] + \\ &+ \sum_{f'} \left[\ln\left(1 + \sum_{a'' \in \mathcal{NC}_a : f'' = f'} N_{a''t}\right)\tilde{\delta}_{f,f'} + \ln^2\left(1 + \sum_{a'' \in \mathcal{NC}_a : f'' = f'} N_{a't}\right)\tilde{\eta}_{f,f'} \right].\end{aligned}$$

The mean attention consists of three parts. The first one captures the impact of the observable market characteristics. The second one captures the impact of the history of previous choices of all firms in market m. The third part captures the peer effect in consideration from markets different from m, where we allow the peer effect to be firm-specific. That is, for the peer effect in consideration, we allow the number of stores of firm f in her neighborhood market to affect her mean attention differently from her competitor's number of stores in her neighborhood market.

The firm pays attention to a market if its attention index is above 0, i.e., $\tilde{\pi}_{at} \geq 0$. As a result, the probability that firm f considers opening a new store in market m at time t is

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

[19] We take logarithms of the number of stores to compensate for a rapid increase in the number of stores in some markets. The results for $N_{a't}$ are qualitatively the same.

where h_t is the history of choice configurations up to time t. The vector of parameters θ contains the consideration parameter $\tilde{\alpha}_{f,f'}$, $\tilde{\beta}_{f,f'}$, $\tilde{\gamma}_{f,f'}$, $\tilde{\delta}_{f,f'}$, and $\tilde{\eta}_{f,f'}$, where $f, f' \in \mathcal{F}$, and the network structure $\mathcal{NC}_a$, $a \in \mathcal{A}$.

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. We assume that the marginal profit of opening an extra store in market m at time t by firm f is

$$\pi_{at} = \bar{\pi}_{at}(S_t, N_t; \theta) - \varepsilon_{at},$$

where the ε_{at}s are i.i.d. across a and t with a known c.d.f. F_ε (in our application F_ε is the Logistic c.d.f.); and

$$\begin{aligned}\bar{\pi}_{at}(S_t, N_t; \theta) =& S'_{mt}\beta_{f'} + \left[\ln(1 + N_{at})\alpha_f + \ln^2(1 + N_{at})\gamma_f\right] + \\ &+ \left[\ln\left(1 + \sum_{a' \in \mathcal{NR}_a} N_{a't}\right)\alpha_{f'} + \ln^2\left(1 + \sum_{a' \in \mathcal{NR}_a} N_{a't}\right)\gamma_{f'}\right]\end{aligned}$$

is the mean marginal profit from one extra store.[20] Hence, the probability of opening a new store in market m by firm f at time t conditional on it being considered is

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

It follows that the probability that a new store is opened in market m by firm f at time t is

$$\mathrm{P}_a(1 \mid h_t, S_t) = F_{\tilde{\varepsilon}}\left(\bar{\bar{\pi}}_{at}(S_t, N_t; \theta)\right) F_\varepsilon\left(\bar{\pi}_{at}(S_t, N_t; \theta)\right), \tag{3}$$

which completely characterizes the probability of observing a new store in a given market by a

[20] 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 availability of the data and to simplify the analysis.

given store conditional on the past history and the market characteristics.

As a result, the vector of parameters θ contains the consideration parameters $\tilde{\alpha}_{f,f'}$, $\tilde{\beta}_{f,f'}$, $\tilde{\gamma}_{f,f'}$, where $f, f' \in \mathcal{F}$, the consideration network structure $\mathcal{NC}_a$, $a \in \mathcal{A}$, and the preference parameters $\alpha_{f'}$, $\beta_{f'}$, and $\gamma_{f'}$, $f' \in \mathcal{F}$. Note that $\mathcal{NR}_a = \{(f', m') \,:\, f' \neq f, m = m'\}$ is assumed to be known and, thus, is not the part of θ.

5.3. Estimation and Inference

The data we have consist of three objects: (i) the exact date of new store openings $\{t_k\}_{k=1}^{K}$; (ii) the state of the market structure $\{N_{at_k}\}_{a\in\mathcal{A},k=1,\ldots,K}$ 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 $\{S_{m,t_k}\}_{a\in\mathcal{A},k=1,\ldots,K}$. Given the data on the number of stores, we can construct the state vector $r_{t_k} = (r_{at_k})_{a\in\mathcal{A}}$, where r_{at_k} captures whether there was a change in the number of stores of firm f in market m at time t_k. That is,

$$r_{at_k} = \mathbb{1}\left(N_{at_{k+1}} > N_{at_k}\right).$$

Moreover, the probability of observing r_{t_k}, given the data and model parameters θ conditional on an alarm clock going off is

$$\begin{aligned} p(r_{t_k}, S_{t_k}, N_{t_k}; \theta) = & \prod_{a:r_{at_k}=1} F_{\tilde{\varepsilon}}\left(\bar{\tilde{\pi}}_{at}(S_{t_k}, N_{t_k}; \theta)\right) F_{\varepsilon}\left(\bar{\pi}_{at_k}(S_{t_k}, N_{t_k}; \theta)\right) \\ & \times \prod_{a:r_{at_k}=0} \left[1 - F_{\tilde{\varepsilon}}\left(\bar{\tilde{\pi}}_{at}(S_{t_k}, N_{t_k}; \theta)\right) F_{\varepsilon}\left(\bar{\pi}_{at_k}(S_{t_k}, N_{t_k}; \theta)\right)\right]. \end{aligned}$$

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

$$p_0(S_{t_k}, N_{t_k}; \theta) = \prod_{a\in\mathcal{A}} \left[1 - F_{\tilde{\varepsilon}}\left(\bar{\tilde{\pi}}_{at}(S_{t_k}, N_{t_k}; \theta)\right) F_{\varepsilon}\left(\bar{\pi}_{at_k}(S_{t_k}, N_{t_k}; \theta)\right)\right].$$

Finally, given that the arrival process is exponential, the log-likelihood of observing the data given θ and normalizing $\lambda_a = 1$ to be equal to 1 (the action of "doing nothing" is not observed; see Section 4.4 for details on this normalization) is

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

and we can define the maximum likelihood estimator of θ, $\hat{\theta}$, as the maximizer of $\hat{L}$ over a parameter space Θ.

Parameter Space Restrictions The vector of parameters θ consists of two very different parts: the parameters corresponding to the network structure (i.e., $\mathcal{NC}_a$, $a \in \mathcal{A}$), and the parameters corresponding to the attention index and marginal profits. The second set of parameters is standard and does not pose any challenge in estimation. Maximizing the likelihood value by searching θ in its parameter space, in theory, could be done in an inner and outer loop fashion. Specifically, in the inner loop, fixing the network structure, one can maximize over the consideration and payoff parameters using the profiled likelihood estimation. The outer loop would search for the network structure that leads to the highest likelihood. Unfortunately, checking all possible network structures is often computationally prohibitive without restrictions. For example, in our application, without any restrictions, the parameter space for $\mathcal{NC}_a$, $a \in \mathcal{A}$, consists of $2^{2\times 152\times(152-1)} = 2^{45904} > 10^{10000}$ possible network structures (i.e., there are $2 \times 152 \times (152 - 1)$ binary variables). If we assume that every firm has the same reference groups in every market and that links are not directed, then the size of the parameter space drops to $2^{152\times(152-1)/2} = 2^{11476} > 10^{1000}$. To simplify the estimation, we will use spatial information about markets. In particular, we assume that if market m' is in the neighborhood of market m, then at least one of the following three conditions holds:

(i) m' and m are in the same province;

(ii) the prefectures where m' and m are located share a border; and/or

(iii) m' is at least 5-th closest (in terms of geographical distance) market to market m.

With these additional constraints, the number of binary parameters describing the network structure is 1582. Searching through all possible network structures given such a simpler initial one is still computationally infeasible, i.e., 2^{1582}. To further facilitate the estimation, instead of searching every possible network, we start the search from the initial/largest possible network and then shut down one link at a time to find the best improvement of the likelihood. We repeat this procedure until no link shutdown leads to any improvement.[21] This method is only guaranteed to converge to a local optimum, which may not be global. However, we believe it provides an informative approximation of the solution.

Estimation of Full Consideration Model We also estimate the preference parameters, assuming that every market is considered by every firm. This allows us to compare the probability of opening a new store under the full consideration and limited consideration models. We expect the full consideration model would predict smaller probabilities of opening a new store. This is due to the fact that the full consideration model would attribute negative marginal profits to a situation in which the market is not considered.

Inference To construct confidence sets for the payoffs and consideration parameters (or their functions), we assume that the estimated consideration network is the true network. In this case, the asymptotic variance matrix of the profits and consideration parameters is a standard inverse of the Fisher information matrix. We suspect that estimation error in the network structure should not affect the asymptotic behavior of the profit and consideration parameters since the cardinality of the parameter space for the network is finite.[22]

5.4. Estimation Results

Network Structure The estimated consideration network has 960 directed links (i.e., the adjacency matrix is not symmetric and has 960 nonzero elements). The initial restrictions we impose allow up to 1582 links, so we use the likelihood value to close down around 600 links. The network is dense in a few regions, especially around some cities. Beijing, Guangzhou, and Shanghai are among those

[21]This heuristic algorithm is a variation of a greedy optimization algorithm. See Kitagawa and Wang (2023) for a recent application in the context of treatment allocation in sequential network games.

[22]We leave a formal analysis of this claim for future research.

cities (see Figure 2).

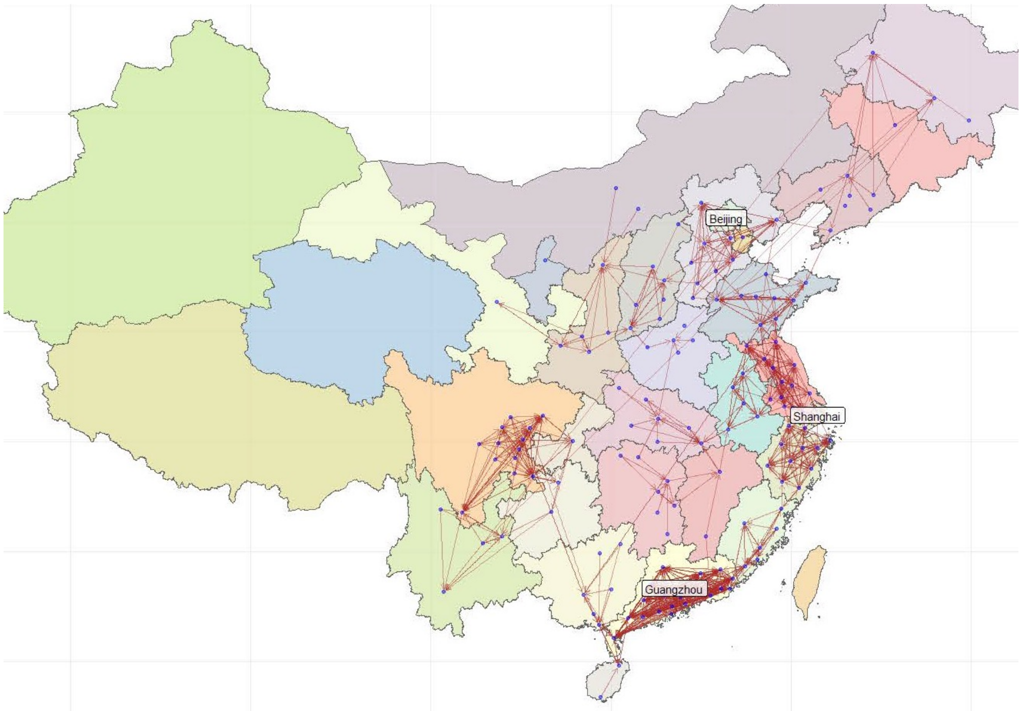


Figure 2 – Consideration network structure.

We next analyze two centrality measures of the estimated directed network: in-degree and out-degree centrality. For a given market, the in-degree (out-degree) centrality measures the number of markets that affect (affected by) the market. The average (median) number of the in-degree and out-degree centrality across markets are 6.3 (4.6) and 6.3 (5), respectively. While consideration of around 30% of markets is not affected by other markets, almost all markets affect consideration of some other markets (see Figure 3). Note that even though the fraction of markets whose consideration is unaffected by other markets is nonnegligible, we can still identify and estimate the preference and consideration parameters since we assume they do not vary across markets.

Consideration First, using the estimated parameters, we computed the consideration probabilities for every market for both firms at the initial condition (i.e., at the moment in which Luckin started operating). Figure 4 depicts the fraction of markets as a function of consideration probability (i.e., a normalized histogram of consideration probabilities). Starbucks displays close to full consideration in almost all markets. Luckin's consideration probabilities, in contrast, are very heterogeneous with

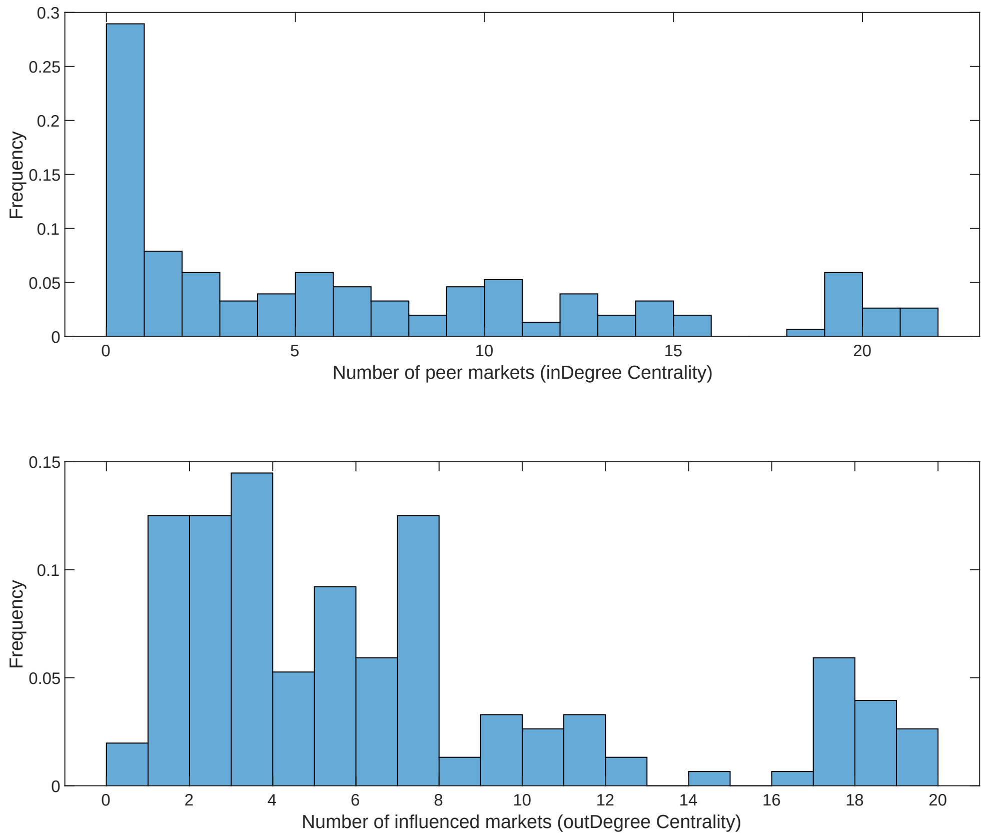


Figure 3 – **The distribution of centrality of the estimated network.** The top and bottom panels correspond to in-degree and out-degree centralities, respectively.

many markets close to being not considered at all.[23]

Next, to demonstrate that the number of stores in neighboring markets affects consideration probabilities, we analyze the relationship between the number of stores and the consideration probabilities using Shanghai as an illustrative example. Shanghai, by the last date of measurements, is the biggest city in terms of number of stores. There are 740 and 419 Starbucks and Luckin stores, respectively. The overall number of stores in Shanghai's neighborhood markets for Starbucks and Luckins are 458 and 336, respectively. To illustrate the impact of different factors on consideration, we change the value of a single factor at a time. We keep all other variables fixed at their values in the last time period. Figure 5 depicts the probability of considering Shanghai (at the moment of deciding whether to open a new store) as a function of the number of stores that belong to

[23]The averages (the standard deviations) across markets of the consideration probabilities are 0.9992 (0.0075) and 0.5511 (0.3634) for Starbucks and Luckin, respectively.

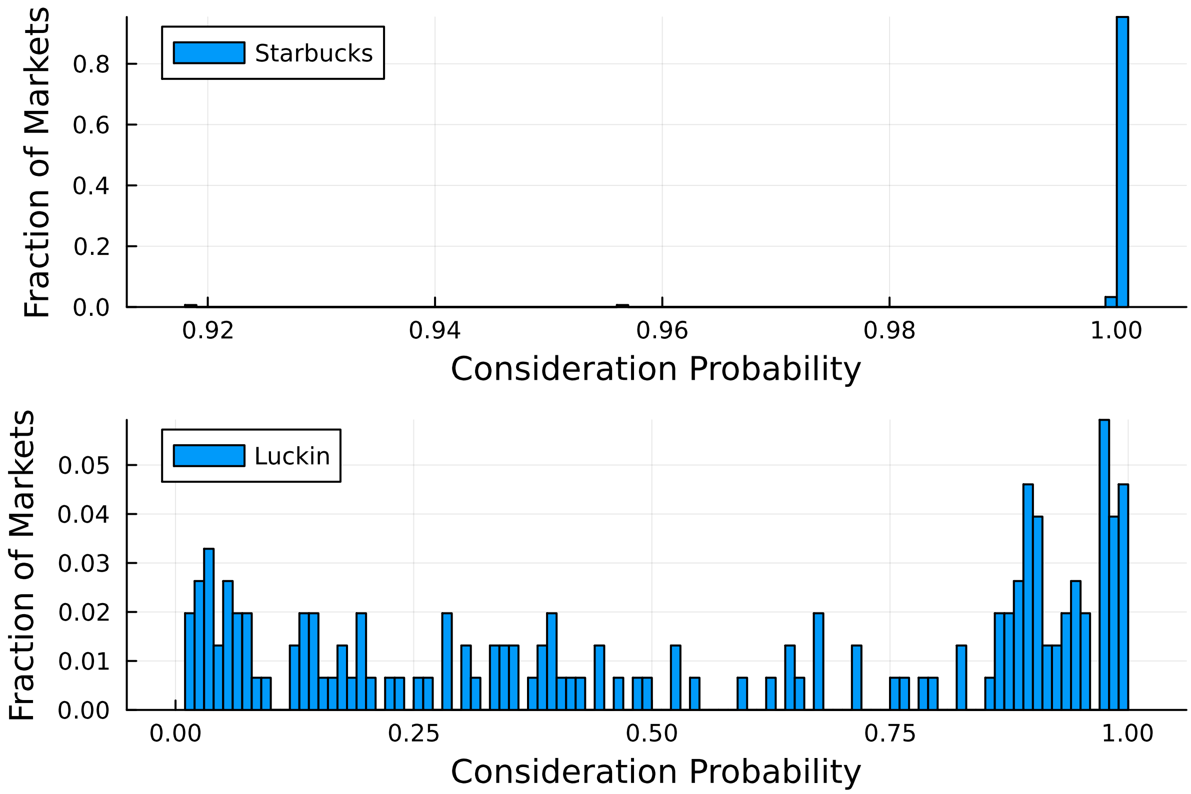
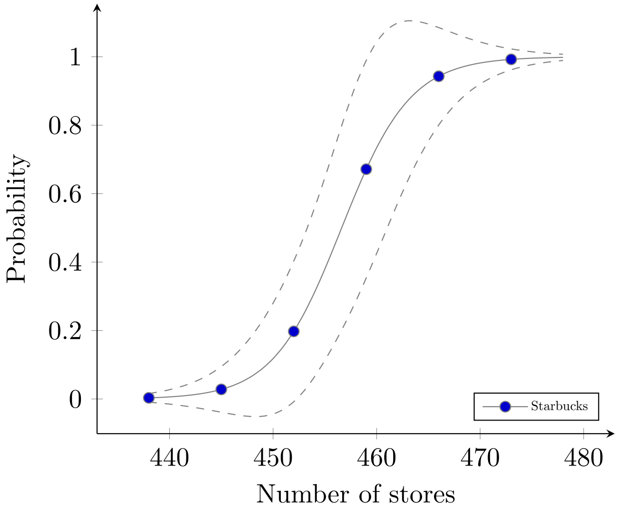


Figure 4 – **Normalized histogram of consideration probabilities for both firms.** The consideration probabilities are computed at the beginning of the data.

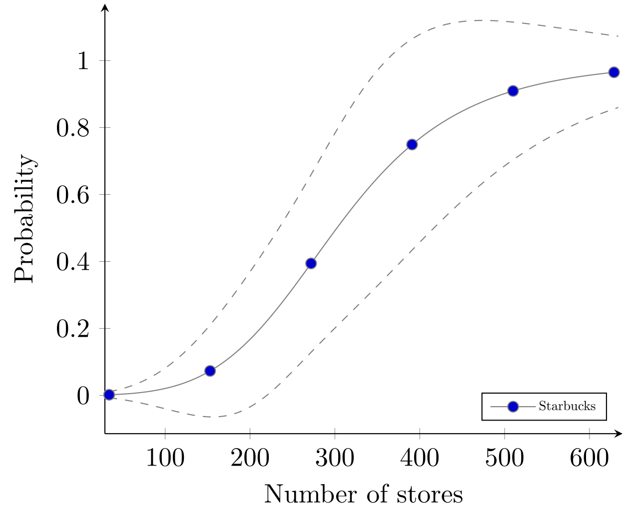
the firm/competitor in the neighborhood together with pointwise 95% confidence bands.[24] A few features are worth noting. First of all, the consideration probabilities vary with the number of stores in the neighbourhood markets, which validates our identification condition. Second, there is difference in how firms change their consideration with increasing own or competitors presence in Shanghai's neighborhood —they move in the opposite direction as the number of stores increases.

Marginal Profits Next, we analyze the probability of opening a new store in Shanghai (conditional on Shanghai being considered) as a function of the number of stores in and around the market. 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 (i.e., we assume that $F_{\tilde{\varepsilon}}(\tilde{\varepsilon}) = 1$ for all $\tilde{\varepsilon}$ and reestimated the marginal profit parameters). 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 6. We note that the estimated expansion probabilities under

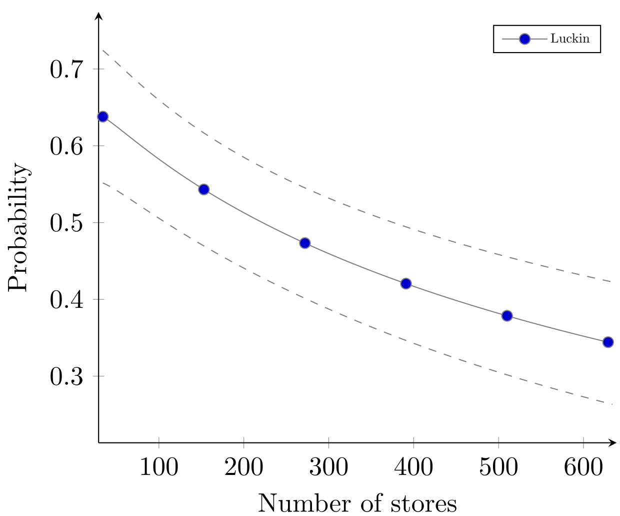
[24]The range in each figure is centered around the number of stores the firm had at the last time period.



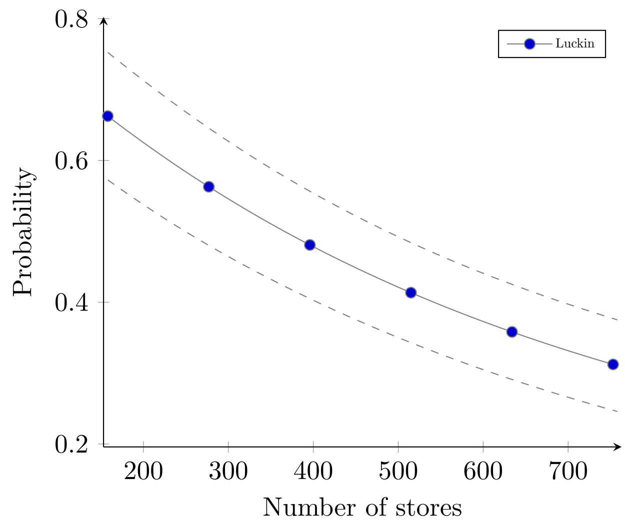
(a) Starbucks. Own stores in the neighbourhood



(b) Starbucks. Opponent's stores in the neighbourhood



(c) Luckin. Own stores in the neighbourhood



(d) Luckin. Opponent's stores in the neighbourhood

Figure 5 – **Consideration probabilities for Starbucks and Luckin as a function of the number of own or competitor's stores in Shanghai's neighborhood.** Dashed lines depict pointwise 95% confidence bands. The fixed variables (e.g., market characteristics) are taken from the last time period.

limited consideration are uniformly bigger than those under full consideration. Importantly, the difference between these probabilities is substantial for Luckin. 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 (a new store is not opened because the market is not considered). The Starbucks limited consideration expansion probabilities are not that different from the full consideration ones. This result is unsurprising given the estimated consideration probabilities–Starbucks considers almost all markets.

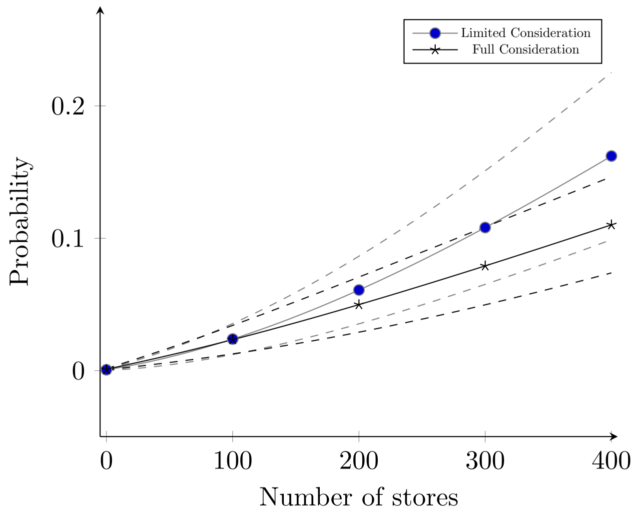
5.5. Counterfactuals

We aim to evaluate the effect of limited consideration and spillover effects in consideration across markets on market structure. We will do so by comparing the fraction of monopolistic, duopolistic, and markets that are not served across time between our full model and different counterfactual scenarios. Note that we do not re-estimate the preference parameters for all the analyses. In Scenario 1, we force both firms to consider all 152 markets. That is, we assume that $F_{\tilde{\varepsilon}}(\tilde{\varepsilon}) = 1$ for all $\tilde{\varepsilon}$.[25] In Scenario 2, we remove connections across markets via consideration. In particular, we assume that the consideration probabilities are only affected by the number of own and opponent stores in the focal market but are not influenced by the number of stores in the neighbouring markets. In Scenario 3, we switch the initial conditions between the two firms, so that Starbucks becomes a newcomer and Luckin is the incumbent.

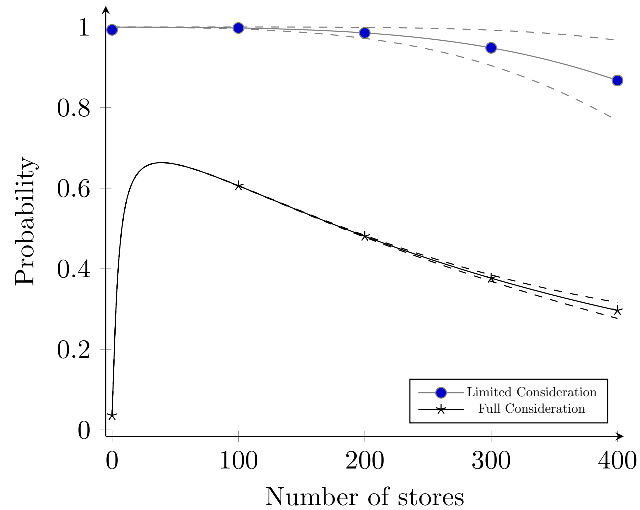
In Scenarios 1 and 2, we first assign to Starbucks and Luckin the number of stores that each of them had in each market at the first time period (i.e., Luckin started with zero stores), and then simulate the expansion decisions for the length of the actual data (about 900 days). In Scenario 3, we switch the initial conditions for firms —that is, we assign to Luckin the number of stores that Starbucks had when the firm started its operations.

Effects of Limited Consideration and Peer Effects in Limited Consideration Figure 7 depicts the fraction of duopolistic, monopolistic, and markets that are not served by any firm as a function of time for Scenarios 1 and 2 and the original model. The dynamics of the fraction of the markets that are not served (i.e., the market penetration) does not change much when one makes both firms fully aware of all markets or when markets are not connected. This result can be explained by Starbucks' behavior. Specifically, since Starbucks is already close to being a full consideration firm, eliminating limited consideration does not change its behavior much. Secondly, because of being close to full consideration, the peer effect in consideration is also very weak for

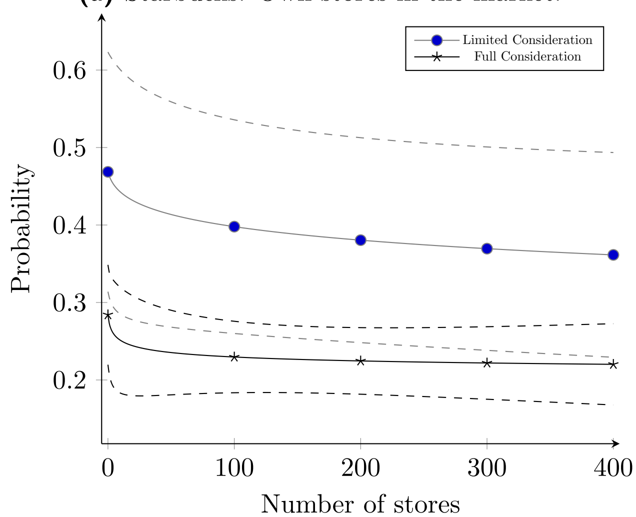
[25] The marginal profit parameters are not reestimated in this scenario.



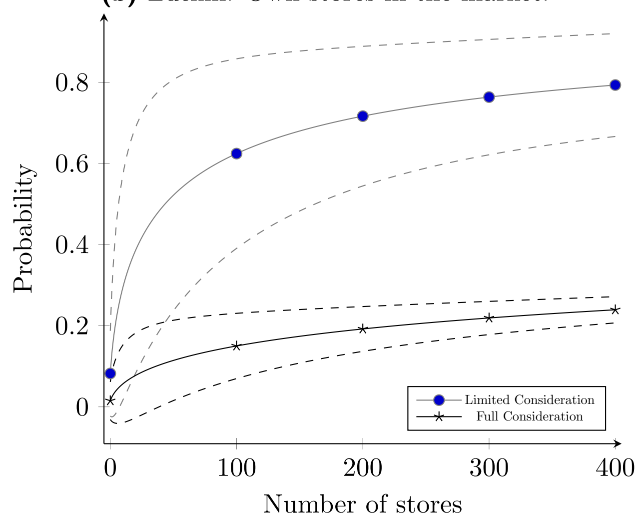
(a) Starbucks. Own stores in the market.



(b) Luckin. Own stores in the market.



(c) Starbucks. Competitor's stores in the market



(d) Luckin. Competitor's stores in the market

Figure 6 – **Expansion probabilities for both coffee chains conditional on considering Shanghai as a function of the number of own or competitor's stores in Shanghai.** Dashed lines depict pointwise 95% confidence bands. The fixed variables (e.g., market characteristics) are taken from the last time period.

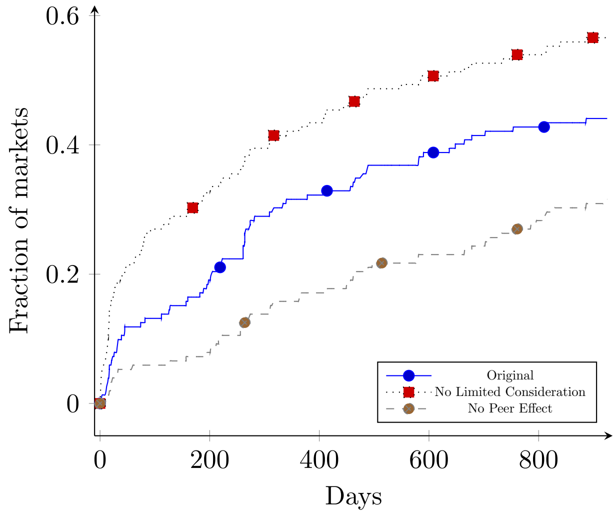
Starbucks. Therefore, the market penetration rate is similar under all three scenarios.

The results for duopolies and monopolies are strikingly different. For example, the full consideration model achieves 44% of duopoly markets about 17 months before the limited consideration model. Similarly, the fraction of 48% of monopolistic markets is achieved by the full consideration model about 29 months faster. Since the market penetration is not affected by limited consideration, the faster rate in occurrence of duopolies in Scenario 1 is purely driven by increased competition. That is, limited consideration substantially slows down competition. Contrary to this result, peer effects in consideration across markets have a positive effect on competition. For example, the final fraction of duopolies and monopolies under Scenario 2 is achieved by our model in about 18 and 30 months, respectively.

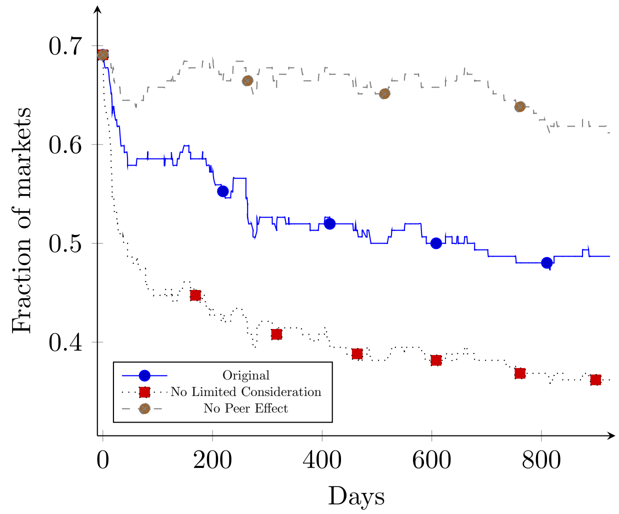
The Effect of the Initial Condition Figure 8 depicts the fraction of monopolistic, duopolistic, and markets that are not served by any firm as a function of time for Scenario 3 and the original model. Switching the initial conditions of firms does not have much effect on market penetration. As before the result is driven by the fact that when either firm has enough stores in many markets, then the firm behaves as a full consideration firm. However, switching of the initial condition has a substantial effect (comparable in magnitude to the effect of limited consideration) on the fraction of duopolies. This indicates that there is an important heterogeneity between firms beyond the differences in initial conditions. Starbucks expands substantially faster than Luckin when firms are switched.

6. Final Remarks

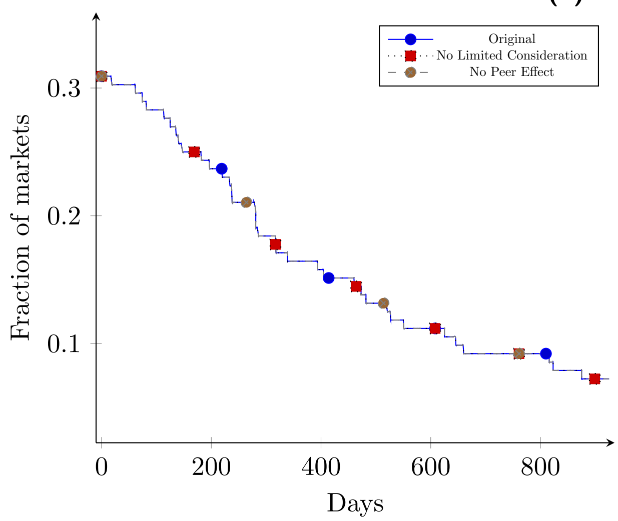
This paper offers a rich model of social interactions where the choices of peers can affect the decision of a given agent in different ways. In particular, the choices of peers might affect the set of options that the agent ends up considering, the preferences over these options, or both. We show that these peer effect mechanisms have different behavioral implications in the data. This allow us to recover the full network structure. In particular, we nonparametrically identify not just the set of



(a) Duopolies.



(b) Monopolies.

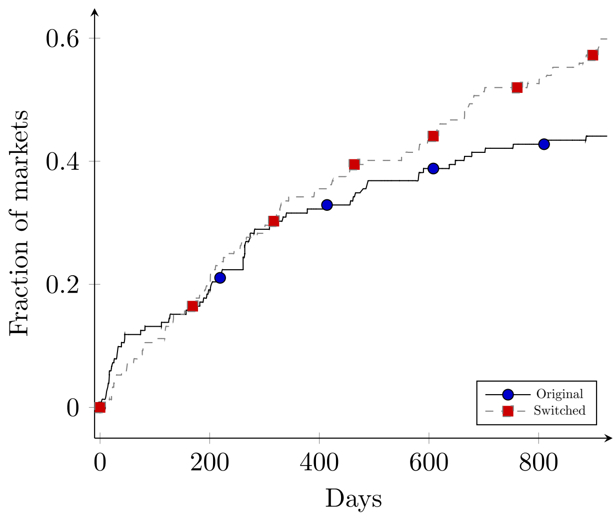


(c) Not served.

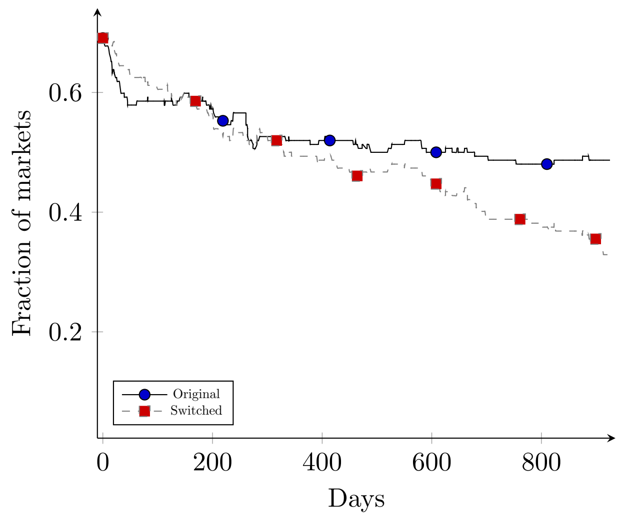
Figure 7 – **Fraction of duopolistic or monopolistic markets or markets that are not served over time.**

connections between the agents, but the type of interactions between them. We then use variation in the choices of peers as the main tool to recover the consideration probabilities and the random preferences. The identification strategy allows very general forms of heterogeneity across people. We propose a consistent estimator of the model parameters and apply it to data on coffee chains expansions in China. The empirical application adds to the literature on boundedly rational firms.

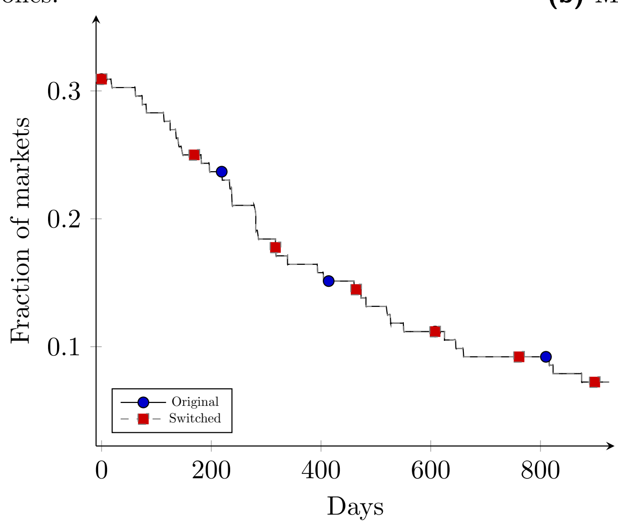
While studying a rather general model, we leave a few interesting variants of our approach for future research. Among them, the possibility of strategic behavior where each agent ends up making decisions with the purpose of affecting the consideration set of others. We believe this set



(a) Duopolies.



(b) Monopolies.



(c) Not served.

Figure 8 – **Fraction of duopolistic or monopolistic markets or markets that are not served over time.**

up could lead to a new model of endogenous social norms or rules within a group of people.

References

Abaluck, Jason and Abi 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, Marina, Matt Elliott, and Pietro Ortoleva (2021) "The importance of Social Norms against Strategic Effects: The case of Covid-19 vaccine uptake," *Economics Letters*, 206, 109979.

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

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

Aguiar, Victor H, Maria Jose Boccardi, Nail Kashaev, and Jeongbin Kim (2023) "Random utility and limited consideration," *Quantitative Economics*, 14 (1), 71–116.

Aguiar, Victor H ⓡ Nail Kashaev (2021) "Identification and Estimation of Discrete Choice Models with Unobserved Choice Sets."

Allen, Roy and John Rehbeck (2022) "Latent complementarity in bundles models," *Journal of Econometrics*, 228 (2), 322–341.

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

Arcidiacono, Peter, Patrick Bayer, Jason R Blevins, and Paul 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, Mark and Steffen Huck (2010) "Behavioral economics as applied to firms: A primer."

Barseghyan, Levon, Maura Coughlin, Francesca Molinari, and Joshua C Teitelbaum (2021a) "Heterogeneous choice sets and preferences," *Econometrica*, 89 (5), 2015–2048.

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

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

Blevins, Jason 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.

Blume, Lawrence 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, Lawrence E, William A Brock, Steven N Durlauf, and Yannis M Ioannides (2011) "Identification of social interactions," in *Handbook of social economics*, 1, 853–964: Elsevier.

Blume, Lawrence E, William A Brock, Steven N Durlauf, and Rajshri Jayaraman (2015) "Linear social interactions models," *Journal of Political Economy*, 123 (2), 444–496.

Bonaldi, Pietro, Ali Hortaçsu, and Jakub Kastl (2015) "An empirical analysis of funding costs spillovers in the euro-zone with application to systemic risk,"Technical report, National Bureau of Economic Research.

Bonhomme, Stéphane, Koen Jochmans, and Jean-Marc Robin (2016) "Non-parametric estimation of finite mixtures from repeated measurements," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78 (1), 211–229.

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

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

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

Caeyers, Bet (2014) "Peer effects in development programme awareness of vulnerable groups in rural Tanzania,"Technical report, Centre for the Study of African Economies, University of Oxford.

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

Cattaneo, Matias D, Xinwei Ma, Yusufcan Masatlioglu, and Elchin Suleymanov (2020) "A random attention model," *Journal of Political Economy*, 128 (7), 2796–2836.

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

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

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

Dardanoni, Valentino, Paola Manzini, Marco Mariotti, and Christopher J Tyson (2020) "Inferring cognitive heterogeneity from aggregate choices," *Econometrica*, 88 (3), 1269–1296.

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

Draganska, Michaela and Daniel 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, Esther and Emmanuel 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.

Dunker, Fabian, Stefan Hoderlein, and Hiroaki Kaido (2017) "Nonparametric identification of random coefficients in endogenous and heterogeneous aggregate demand models,"Technical report, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

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

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

Fox, Jeremy T and Natalia Lazzati (2017) "A note on identification of discrete choice models for bundles and binary games," *Quantitative Economics*, 8 (3), 1021–1036.

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

Gentzkow, Matthew (2007) "Valuing new goods in a model with complementarity: Online newspapers," *American Economic Review*, 97 (3), 713–744.

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

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

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

Hall, Peter and Xiao-Hua Zhou (2003) "Nonparametric estimation of component distributions in a multivariate mixture," *The annals of statistics*, 31 (1), 201–224.

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

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

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

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

Houde, Jean-FranÃ§ois, Peter Newberry, and Katja 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.

Hu, Yingyao (2008) "Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution," *Journal of Econometrics*, 144 (1), 27–61.

Iaria, Alessandro and Ao Wang (2020) "Identification and estimation of demand for bundles."

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

Kasahara, Hiroyuki and Katsumi Shimotsu (2009) "Nonparametric identification of finite mixture models of dynamic discrete choices," *Econometrica*, 77 (1), 135–175.

Kashaev, Nail (2023) "Identification and estimation of multinomial choice models with latent special covariates," *Journal of Business & Economic Statistics*, 41 (3), 695–707.

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

Kitagawa, Toru and Guanyi Wang (2023) "Individualized Treatment Allocation in Sequential Network Games."

Kitamura, Yuichi and Louise Laage (2018) "Nonparametric analysis of finite mixtures," *arXiv preprint arXiv:1811.02727*.

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

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

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

Lleras, Juan Sebastian, Yusufcan Masatlioglu, Daisuke Nakajima, and Erkut Y Ozbay (2017) "When more is less: Limited consideration," *Journal of Economic Theory*, 170, 70–85.

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

Manski, Charles 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, Paola and Marco Mariotti (2014) "Stochastic choice and consideration sets," *Econometrica*, 82 (3), 1153–1176.

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

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan (2003) "Price uncertainty and consumer search: A structural model of consideration set formation," *Marketing science*, 22 (1), 58–84.

Qiu, Liangfei, Zhan Shi, and Andrew 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, John H and James M Lattin (1991) "Development and testing of a model of consideration set composition," *Journal of Marketing Research*, 28 (4), 429–440.

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

——— (1955) "A behavioral model of rational choice," *The quarterly journal of economics*, 99–118.

Wang, Rui (2023) "Testing and Identifying Substitution and Complementarity Patterns," *arXiv preprint arXiv:2304.00678*.

Yu, Xiaoyu Fisher (2023) "Bundle Choice with Limited Consideration: An Application to Yogurt Markets."

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

A. The Regularity Condition

In this appendix, we formally state and discuss the regularity condition needed for the identification of the network in Section 3.

For a given $a \in \mathcal{A}$, define the set of all possible values that $\mathrm{NR}_a^{\mathcal{Y}}(\mathbf{y})$ and $\mathrm{NC}_a^{\mathcal{Y}}(\mathbf{y})$ can take:

$$\mathrm{Nrc}_a = \left\{ \left(\mathrm{NR}_a^{\mathcal{Y}}(\mathbf{y}), \mathrm{NC}_a^{\mathcal{Y}}(\mathbf{y}) \right) \;:\; \mathbf{y} \in \mathcal{Y}^A \right\}.$$

Also define $\bar{\mathrm{P}}_a \left(v \mid \mathbf{nr}, \mathbf{nc} \right)$ as the probability that Agent a picks option $v \neq 0$ conditional on $(\mathbf{nr}, \mathbf{nc}) \in \mathrm{Nrc}_a$, where $nr^{v'}$ and $nc^{v'}$ denote the number of peers that affect preference only and consideration only, respectively, picking alternative v', $v' \in \mathcal{Y} \setminus \{0\}$. That is,

$$\bar{\mathrm{P}}_a \left(v \mid \mathbf{nr}, \mathbf{nc} \right) = \sum_{\mathcal{C} \subseteq \mathcal{Y}} \mathrm{R}_a \left(v \mid nr^{\mathcal{C}}, \mathcal{C} \right) \mathrm{C}_a \left(\mathcal{C} \mid \mathbf{nc}, \mathcal{Y} \right),$$

where $nr^{\mathcal{C}} = \left(nr^{v'} \right)_{v' \in \mathcal{C} \setminus \{0\}}$ and

$$\mathrm{C}_a \left(\mathcal{C} \mid \mathbf{nc}, \mathcal{Y} \right) = \prod_{v' \in \mathcal{C}} \mathrm{Q}_a(v \mid nc^{v'}) \prod_{v' \in \mathcal{Y} \setminus \mathcal{C}} (1 - \mathrm{Q}_a(v \mid nc^{v'})).$$

Let $\Delta_{v,v'} f(\mathbf{x}, \mathbf{y})$ denote an operator that computes the increment of a given function when the v-th component of $\mathbf{x}$ and the v'-th component of $\mathbf{y}$ are increased by 1, respectively. We use a convention that if $v = 0$ ($v' = 0$), then $\mathbf{x}$ ($\mathbf{y}$) remains unchanged.

Assumption 6. For any $a \in \mathcal{A}$,

(i) there exist an alternative $v \in \mathcal{Y} \setminus \{0\}$ and a vector of aggregate peers' choices $(\mathbf{nr}, \mathbf{nc}) \in \mathrm{Nrc}_a$ such that

$$\Delta_{v,v} \ln \bar{\mathrm{P}}_a(v \mid \mathbf{nr}, \mathbf{nc}) \neq 0;$$

(ii) there exist three sets of alternative pairs and aggregate peers' choices, i.e., $\{v_i, w_i, \mathbf{nr}_i, \mathbf{nc}_i\}$, where $v_i, w_i \in \mathcal{Y} \setminus \{0\}$, $v_i \neq w_i$, $(\mathbf{nr}_i, \mathbf{nc}_i) \in \mathrm{Nrc}_a$, and $i = 1, 2, 3$, such that

$$\Delta_{w_1,0}\Delta_{v_1,0} \ln \bar{\mathrm{P}}_a(v_1 \mid \mathbf{nr}_1, \mathbf{nc}_1) \neq 0,$$
$$\Delta_{0,w_2}\Delta_{v_2,0} \ln \bar{\mathrm{P}}_a(v_2 \mid \mathbf{nr}_2, \mathbf{nc}_2) \neq 0,$$
$$\Delta_{w_3,w_3}\Delta_{v_3,0} \ln \bar{\mathrm{P}}_a(v_3 \mid \mathbf{nr}_3, \mathbf{nc}_3) \neq 0.$$

Assumption 6(i) guarantees that the peer effects in consideration and preferences do not cancel each other. That is, peers that affect both consideration and preferences are distinguishable from those who are not in one's reference group. Assumption 6(ii) is needed for distinguishing peers who affect consideration only from those who affect preference. Specifically, for the peers who affects consideration only, the double shift described above always equal zero, while the double shift in the observed CCPs is guaranteed to be nonzero for peers who affect preference for at least three scenarios by Assumption 6(ii).

It is worth emphasizing that the above inequality is only required to hold for one selection of the actions and peers' configuration. Additionally, Assumption 6(ii) allows $v_1 = v_2 = v_3$, $w_1 = w_2 = w_3$, and $(\mathbf{nr}_1, \mathbf{nc}_1) = (\mathbf{nr}_2, \mathbf{nc}_2) = (\mathbf{nr}_3, \mathbf{nc}_3)$. Furthermore, as the number of peers and/or the size of the menu is growing, it gets harder to violate Assumption 6. Therefore, Assumption 6 is a mild functional form restriction that is usually generically satisfied.

The following example clarifies the scope of Assumption 6.

Example 3. Suppose that

$$\mathrm{R}_a\left(v \mid t, \mathcal{C}\right) = \frac{u_v(t_v)}{\sum_{v' \in \mathcal{C}} u_{v'}(t_{v'})},$$

where $u_0(t_0) = 1$ and $u_v(\cdot)$, $v \in \mathcal{Y} \setminus \{0\}$, are strictly monotone positive functions. That is, after the consideration set is formed, Agent a picks alternatives according to a logit-type rule.

Then, for the binary choice case, i.e., $Y = 1$ and $v = 1$, we have that

$$\bar{\mathrm{P}}_a(v \mid nr^v, nc^v) = \mathrm{Q}_a\left(v \mid nc^v\right) \frac{u_1(nr^v)}{1 + u_1(nr^v)}.$$

Note that Assumption 6(i) is violated if, the following equality holds:

$$\frac{\mathrm{Q}_a\left(v \mid nc^v + 1\right)}{\mathrm{Q}_a\left(v \mid nc^v\right)} = \frac{u_1(nr^v)}{1 + u_1(nr^v)} \frac{1 + u_1(nr^v + 1)}{u_1(nr^v + 1)},$$

for all admissible values of nr^v and nc^v. Evidently, the larger the peer group is, the harder this equality to hold for all values.

We also illustrate that the same conclusion holds with a larger menu choice. Specifically, if we add one more alternative $v' = 2$, then

$$\begin{aligned}
\bar{\mathrm{P}}_a(v \mid nr^v, nr^{v'}, nc^v, nc^{v'}) &= \mathrm{Q}_a\left(v \mid nc^v\right) \mathrm{Q}_a\left(v' \mid nc^{v'}\right) \frac{u_1(nr^v)}{1 + u_1(nr^v) + u_2(nr^{v'})} \\
&\quad + \mathrm{Q}_a\left(v \mid nc^v\right) (1 - \mathrm{Q}_a\left(v' \mid nc^{v'}\right)) \frac{u_1(nr^v)}{1 + u_1(nr^v)} \\
&= \mathrm{Q}_a\left(v \mid nc^v\right) \frac{u_1(nr^v)}{1 + u_1(nr^v)} \frac{1 + u_1(nr^v) + \left[1 - \mathrm{Q}_a\left(v' \mid nc^{v'}\right)\right] u_2(nr^{v'})}{1 + u_1(nr^v) + u_2(nr^{v'})}.
\end{aligned}$$

So Assumption 6(i) is violated only if for $v \in \{1, 2\}$, $\Delta_{v,v} \ln \bar{\mathrm{P}}_a(v \mid \mathbf{nr}, \mathbf{nc}) = 0$ for all admissible values of $\mathbf{nr}$ and $\mathbf{nc}$. Specifically, it is violated at $\mathbf{nr} = \mathbf{nc} = \mathbf{0}$ indicates that

$$\begin{aligned}
&\ln \frac{\mathrm{Q}_a\left(v \mid 1\right)}{\mathrm{Q}_a\left(v \mid 0\right)} + \ln \left[\frac{u_1(1)}{1 + u_1(1)} \frac{1 + u_1(0)}{u_1(0)}\right] \\
&+ \ln \frac{1 + u_1(0) + u_2(0)}{1 + u_1(1) + u_2(0)} - \ln \frac{1 + u_1(0) + \left[1 - \mathrm{Q}_a\left(v' \mid 0\right)\right] u_2(0)}{1 + u_1(1) + \left[1 - \mathrm{Q}_a\left(v' \mid 0\right)\right] u_2(0)} = 0
\end{aligned}$$

for $v \in \{1, 2\}$. Note that if Nrc_a is rich enough to allow for switch from $\mathrm{Q}_a\left(v' \mid 0\right)$ to $\mathrm{Q}_a\left(v' \mid 1\right)$

without changing other parameters, then if Assumption 6(i) is violated at any point of the support, then

$$\ln \frac{\mathrm{Q}_a(v \mid 1)}{\mathrm{Q}_a(v \mid 0)} + \ln \left[\frac{u_1(1)}{1+u_1(1)} \frac{1+u_1(0)}{u_1(0)} \right]$$
$$+ \ln \frac{1+u_1(0)+u_2(0)}{1+u_1(1)+u_2(0)} - \ln \frac{1+u_1(0)+[1-\mathrm{Q}_a(v' \mid 1)]\, u_2(0)}{1+u_1(1)+[1-\mathrm{Q}_a(v' \mid 1)]\, u_2(0)} = 0.$$

Thus, it has to be the case that

$$\frac{1+u_1(0)+[1-\mathrm{Q}_a(v' \mid 1)]\, u_2(0)}{1+u_1(1)+[1-\mathrm{Q}_a(v' \mid 1)]\, u_2(0)} = \frac{1+u_1(0)+[1-\mathrm{Q}_a(v' \mid 0)]\, u_2(0)}{1+u_1(1)+[1-\mathrm{Q}_a(v' \mid 0)]\, u_2(0)}.$$

The latter is possible if and only if $\mathrm{Q}_a(v' \mid 0) = \mathrm{Q}_a(v' \mid 1)$ which violates Assumption 2(iii).

Note that Assumption 2(iii) also establishes that

$$\Delta_{0,v'} \Delta_{v,0} \ln \bar{\mathrm{P}}_a(v \mid \mathbf{0}, \mathbf{0}) \neq 0,$$

which guarantees Assumption 6(ii) with $i = 2$. However, Assumption 6(ii) is more general. It only requires the nonzero to hold for one configuration and one action. We illustrate the restrictions of Assumption 6(ii) below. Specifically, violations of Assumption 6(iii) for case $i = 1$ means that the following equations hold

$$\Delta_{v',0} \Delta_{v,0} \ln \bar{\mathrm{P}}_a(v \mid \mathbf{nr}, \mathbf{nc}) = 0,$$
$$\Delta_{v,0} \Delta_{v',0} \ln \bar{\mathrm{P}}_a(v' \mid \mathbf{nr}, \mathbf{nc}) = 0,$$

for all combinations of the choice configuration $\mathbf{nr}$ and $\mathbf{nc}$, including $\mathbf{nr} = \mathbf{0}$ and $\mathbf{nc} = \mathbf{0}$. It is harder for all equations to hold when the peer group size is getting larger or if the size of the alternatives becomes larger. Similar logic carries to cases $i = 1$ and $i = 3$.

A.1. The case of Bundles

In this appendix, we discuss how the regularity condition needs to be modified to handle the extension of our model to bundles in Section 4.6. Note that with bundles the menu becomes $\mathcal{B}(\mathcal{Y})$. Thus, we need to redefine Nrc_a and $\bar{\mathrm{P}}_a$. Let

$$\mathrm{Nrc}_a^{\mathcal{B}(\mathcal{Y})} = \left\{ \left(\mathrm{NR}_a^{\mathcal{B}(\mathcal{Y})}(\mathbf{y}), \mathrm{NC}_a^{\mathcal{B}(\mathcal{Y})}(\mathbf{y}) \right) \, : \, \mathbf{y} \in \mathcal{B}(\mathcal{Y})^A \right\}.$$

Also define $\bar{\mathrm{P}}_a \left(b \mid \mathbf{nr}, \mathbf{nc} \right)$, with $(\mathbf{nr}, \mathbf{nc}) \in \mathrm{Nrc}_a^{\mathcal{B}(\mathcal{Y})}$, as the probability that Agent a picks bundle $b \neq 0$ conditional on $nr^{v'}$ and $nc^{v'}$ peers that affect preference only and consideration only, respectively, picking bundles v', $v' \in \mathcal{B}(\mathcal{Y}) \setminus \{0\}$. That is,

$$\bar{\mathrm{P}}_a \left(b \mid \mathbf{nr}, \mathbf{nc} \right) = \sum_{\mathcal{C} \subseteq \mathcal{Y}} \mathrm{R}_a \left(v \mid nr^{\mathcal{B}(\mathcal{C})}, \mathcal{B}(\mathcal{C}) \right) \mathrm{C}_a \left(\mathcal{C} \mid \mathbf{nc}, \mathcal{Y} \right),$$

where $nr^{\mathcal{C}} = \left(nr^{b'} \right)_{b' \in \mathcal{B}(\mathcal{Y}) \setminus \{0\}}$ and

$$\mathrm{C}_a \left(\mathcal{C} \mid \mathbf{nc}, \mathcal{Y} \right) = \prod_{v' \in \mathcal{C}} \mathrm{Q}_a(v \mid nc^{v'}) \prod_{v' \in \mathcal{Y} \setminus \mathcal{C}} (1 - \mathrm{Q}_a(v \mid nc^{v'})).$$

The modified version of Assumption 6 then becomes

Assumption 6′. For any $a \in \mathcal{A}$

(i) there exist $b \in \mathcal{B}(\mathcal{Y}) \setminus \{0\}$ and $(\mathbf{nr}, \mathbf{nc}) \in \mathrm{Nrc}_a^{\mathcal{B}(\mathcal{Y})}$ such that

$$\Delta_{v,v} \ln \bar{\mathrm{P}}_a(b \mid \mathbf{nr}, \mathbf{nc}) \neq 0;$$

(ii) there exist $b_i, d_i \in \mathcal{B}(\mathcal{Y}) \setminus \{0\}$, and $(\mathbf{nr}_i, \mathbf{nc}_i) \in \mathrm{Nrc}_a$, $i = 1, 2, 3$, such that $b_i \neq d_i$, $i = 1, 2, 3$, and

$$\Delta_{d_1,0} \Delta_{b_1,0} \ln \bar{\mathrm{P}}_a(b_1 \mid \mathbf{nr}_1, \mathbf{nc}_1) \neq 0,$$

$$\Delta_{0,d_2} \Delta_{b_2,0} \ln \bar{\mathrm{P}}_a(b_2 \mid \mathbf{nr}_2, \mathbf{nc}_2) \neq 0,$$

$$\Delta_{d_3,d_3}\Delta_{b_3,0}\ln\bar{\mathrm{P}}_a(b_3\mid \mathbf{nr}_3,\mathbf{nc}_3)\neq 0.$$

B. Proofs

B.1. Proof of Proposition 3.1

Fix some $a\in\mathcal{A}$. We will prove that

$$a'\notin\mathcal{N}_a \iff \frac{\mathrm{P}_a(v\mid\mathbf{y})}{\mathrm{P}_a(v\mid\mathbf{y}')}=1 \text{ for all } v, \text{ and } \mathbf{y},\mathbf{y}' \text{ that are different in the } a'\text{th component only.}$$

The "only if" part is straightforward. To show the "if" part, assume, towards a contradiction, that

$$\frac{\mathrm{P}_a(v\mid\mathbf{y})}{\mathrm{P}_a(v\mid\mathbf{y}')}=1 \text{ for all } \mathbf{y},\mathbf{y}' \text{ that are different in the } a'\text{th component only,}$$

and $a'\in\mathcal{N}_a$. Let $\mathbf{y}_z^v$ denote the vector in which the z-th component of $\mathbf{y}$ is replaced by v.

Note that the observed CCP can be expressed as

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

where the first component only depends on the number of consideration peers selecting alternative v, and the second components depends on the whole vector of the number of preference peers' choices.

If $a'\in\mathcal{NC}_a\setminus\mathcal{NR}_a$, then

$$\frac{\mathrm{P}_a(v\mid\mathbf{0}_{a'}^v)}{\mathrm{P}_a(v\mid\mathbf{0})}=\frac{\mathrm{Q}_a(v\mid 1)}{\mathrm{Q}_a(v\mid 0)}\neq 1,$$

where the first equality holds by Assumption 2(ii) and the fact that $\mathrm{NC}_a^v(\mathbf{0})=0$ and $\mathrm{NC}_a^v(\mathbf{0}_{a'}^v)=1$. (It also follows as the change of Agent a''s choice does not affect Agent a's preference towards any alternative.) The last inequality follows from Assumption 2(iii).

Similarly, if $a' \in \mathcal{NR}_a \setminus \mathcal{NC}_a$, then

$$\begin{aligned}
\frac{\mathrm{P}_a\left(v \mid \mathbf{0}^v_{a'}\right)}{\mathrm{P}_a\left(v \mid \mathbf{0}\right)} &= \frac{\sum_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}_a^{\mathcal{C}\cup\{v\}}\left(\mathbf{0}^v_{a'}\right), \mathcal{C}\cup\{v\}\right) \mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}_a^{\mathcal{Y}\setminus\{v\}}(\mathbf{0}^v_{a'}), \mathcal{Y}\setminus\{v\}\right)}{\sum_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}_a^{\mathcal{C}\cup\{v\}}\left(\mathbf{0}\right), \mathcal{C}\cup\{v\}\right) \mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}_a^{\mathcal{Y}\setminus\{v\}}(\mathbf{0}), \mathcal{Y}\setminus\{v\}\right)} \\
&= \frac{\sum_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid 1, \mathcal{C}\cup\{v\}\right) \mathrm{C}_a\left(\mathcal{C} \mid 0, \mathcal{Y}\setminus\{v\}\right)}{\sum_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid 0, \mathcal{C}\cup\{v\}\right) \mathrm{C}_a\left(\mathcal{C} \mid 0, \mathcal{Y}\setminus\{v\}\right)} \\
&\neq 1,
\end{aligned}$$

where the first equality holds because the probability of considering v does not change when switching Agent a''s choice from 0 to v. The second equality holds by Assumption 2(ii) and 3(ii), and the last equality holds from Assumption 3(iii).

Hence, the only left possibility is $a' \in \mathcal{NCR}_a$. But the latter contradicts Assumption 6(i), since $a' \in \mathcal{NCR}_a$ would imply that the consideration peer effect offsets the preference peer effect *everywhere* over the support. The contradiction completes the proof.

B.2. Proof of Proposition 3.2

Note that $\mathcal{N}_a$ is identified by Proposition 3.1. Take any two distinct agents $a', a'' \in \mathcal{N}_a$. We will show that $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$ if and only if

$$\frac{\mathrm{P}_a\left(v \mid \mathbf{y}^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}\right)} = \frac{\mathrm{P}_a\left(v \mid (\mathbf{y}^v_{a'})^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}^v_{a'}\right)}, \tag{4}$$

for all $v \in \mathcal{Y} \setminus \{0\}$, all $w \notin \{0, v\}$, and all $\mathbf{y}$ with $y_{a'} = y_{a''} = 0$. Thus, $\mathcal{NC}_a \setminus \mathcal{NR}_a$ is identified from P_a.

To prove the "only if" part note that if $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$ and $\mathbf{y}$ is such that $y_{a'} = y_{a''} = 0$, then

$$\frac{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}^w_{a''}\right)+1\right)}{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}\right)+1\right)} = \frac{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}\right)+1\right)}{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}\right)+1\right)} = 1 = \frac{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}\right)\right)}{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}\right)\right)} = \frac{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}^w_{a''}\right)\right)}{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}\right)\right)},$$

where the first and the last equalities follow from the fact that $w \neq v$. Hence, since $(\mathbf{y}^v_{a'})^w_{a''} = (\mathbf{y}^w_{a''})^v_{a'}$

we have that

$$
\begin{aligned}
\frac{\mathrm{P}_a\left(v \mid (\mathbf{y}^v_{a'})^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}^v_{a'}\right)} = \frac{\mathrm{P}_a\left(v \mid (\mathbf{y}^w_{a''})^v_{a'}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}^v_{a'}\right)} &= \frac{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}^w_{a''}\right)+1\right)}{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}\right)+1\right)} \\
&\times \frac{\sum_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}^{\mathcal{C}\cup\{v\}}_a\left(\mathbf{y}^w_{a''}\right), \mathcal{C}\cup\{v\}\right) \mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}^{\mathcal{Y}\setminus\{v\}}_a(\mathbf{y}^w_{a''}), \mathcal{Y}\setminus\{v\}\right)}{\sum_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}^{\mathcal{C}\cup\{v\}}_a\left(\mathbf{y}\right), \mathcal{C}\cup\{v\}\right) \mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}^{\mathcal{Y}\setminus\{v\}}_a(\mathbf{y}), \mathcal{Y}\setminus\{v\}\right)} \\
&= \frac{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}^w_{a''}\right)\right)}{\mathrm{Q}_a\left(v \mid \mathrm{NC}^v_a\left(\mathbf{y}\right)\right)} \\
&\times \frac{\sum_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}^{\mathcal{C}\cup\{v\}}_a\left(\mathbf{y}^w_{a''}\right), \mathcal{C}\cup\{v\}\right) \mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}^{\mathcal{Y}\setminus\{v\}}_a(\mathbf{y}^w_{a''}), \mathcal{Y}\setminus\{v\}\right)}{\sum_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \mathrm{R}_a\left(v \mid \mathrm{NR}^{\mathcal{C}\cup\{v\}}_a\left(\mathbf{y}\right), \mathcal{C}\cup\{v\}\right) \mathrm{C}_a\left(\mathcal{C} \mid \mathrm{NC}^{\mathcal{Y}\setminus\{v\}}_a(\mathbf{y}), \mathcal{Y}\setminus\{v\}\right)} \\
&= \frac{\mathrm{P}_a\left(v \mid \mathbf{y}^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}\right)}.
\end{aligned}
$$

To prove the "if" part, note that it is equivalent to the statement that if $a' \in \mathcal{NR}_a$, then there exist $a'' \in \mathcal{N}_a$, v, w, and $\mathbf{y}$ with $y_{a'} \neq v$ and $y_{a''} \notin \{v, w\}$ such that

$$
\frac{\mathrm{P}_a\left(v \mid \mathbf{y}^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}\right)} \neq \frac{\mathrm{P}_a\left(v \mid (\mathbf{y}^v_{a'})^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}^v_{a'}\right)}.
$$

or equivalently

$$
\Delta^v_{a'}\Delta^w_{a''} \log \mathrm{P}_a\left(v \mid \mathbf{y}\right) \neq 0.
$$

If $a'' \in \mathcal{NR}_a \setminus \mathcal{NC}_a$, then let $i = 1$. If $a'' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$, then let $i = 2$. Finally, if $a'' \in \mathcal{NCR}_a$, then let $i = 3$. Take $v = v_i$, $w = w_i$, and $\mathbf{y}$ such that $y_{a'} = y_{a''} = 0$, $\mathrm{NR}^{\mathcal{Y}}_a(\mathbf{y}) = \mathbf{nr}_i$, and $\mathrm{NC}^{\mathcal{Y}}_a(\mathbf{y}) = \mathbf{nc}_i$ from Assumption 6(ii). Then

$$
\Delta^v_{a'}\Delta^w_{a''} \ln \mathrm{P}_a\left(v \mid \mathbf{y}\right) = \Delta^w_{a''}\Delta^v_{a'} \log \mathrm{P}_a\left(v \mid \mathbf{y}\right) = \begin{cases} \Delta_{w_1,0}\Delta_{v_1,0} \ln \bar{\mathrm{P}}_a(v_1 \mid \mathbf{nr}_1, \mathbf{nc}_1) \neq 0 \text{ if } i = 1, \\ \Delta_{0,w_2}\Delta_{v_2,0} \ln \bar{\mathrm{P}}_a(v_2 \mid \mathbf{nr}_2, \mathbf{nc}_2) \neq 0 \text{ if } i = 2, \\ \Delta_{w_3,w_3}\Delta_{v_3,0} \ln \bar{\mathrm{P}}_a(v_3 \mid \mathbf{nr}_3, \mathbf{nc}_3) \neq 0 \text{ if } i = 3, \end{cases}
$$

where the first equality follows from exchangeability of the difference operator, the second equality follows from the definition of $\bar{\mathrm{P}}_a$, and the last inequality follows from Assumption 6(ii). So in all possible cases, Assumption 6(ii) implies that if $a' \in \mathcal{NR}_a$, then there exist v, w, and $\mathbf{y}$ with

$y_{a'} = y_{a''} = 0$ such that

$$\frac{\mathrm{P}_a\left(v \mid \mathbf{y}^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}\right)} \neq \frac{\mathrm{P}_a\left(v \mid (\mathbf{y}^v_{a'})^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}^v_{a'}\right)}.$$

B.3. Proof of Proposition 3.3

Note that we know $\mathcal{N}_a$ and $\mathcal{NR}_a$ (or $\mathcal{NC}_a \setminus \mathcal{NR}_a$). To identify the rest of the network structure ($\mathcal{NR}_a \setminus \mathcal{NC}_a$ and $\mathcal{NCR}_a$), suppose that $\mathcal{NC}_a \setminus \mathcal{NR}_a \neq \emptyset$. Take $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$. First, note that

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

Thus, for any $a'' \in \mathcal{NR}_a$,

$$\Delta^v_{a''}\Delta^v_{a'} \ln \mathrm{P}_a(v \mid \mathbf{0}) \neq 0 \iff a'' \in \mathcal{NCR}_a.$$

Hence, $\mathcal{NCR}_a$ is identified from P_a.

Next, suppose that $\mathcal{NC}_a \setminus \mathcal{NR}_a = \emptyset$. Then, by Assumption 4, either $\mathcal{N}_a = \mathcal{NR}_a \setminus \mathcal{NC}_a$ or both $\mathcal{NR}_a \setminus \mathcal{NC}_a$ and $\mathcal{NCR}_a$ are nonempty. Since the consideration effect is nonzero, the effects of preference-only peers and preference-and-consideration peers have to be different. As a result, we can identify the partition of $\mathcal{NR}_a$, $\mathcal{M}'$ and $\mathcal{M}''$, such that one of its elements is $\mathcal{NCR}_a$. Since $|\mathcal{N}_a| \geq 3 - |\mathcal{NC}_a \setminus \mathcal{NR}_a| = 3$, we can take $a' \in \mathcal{M}'$ and $a'' \in \mathcal{M}''$. Next, take $\mathbf{y}$ such that $y_a = 0$ for all $a \neq a'$ and $y_{a'} = v$. Next note that

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

Finally, take another $a''' \notin \{a', a''\}$ in either $\mathcal{M}'$ or $\mathcal{M}''$. Without loss of generality, assume that $a''' \in \mathcal{M}'$. Note that, by Assumption 2,

$$\Delta^v_{a'''} \ln \mathrm{P}_a\left(v \mid \mathbf{y}\right) - \Delta^v_{a'''} \ln \mathrm{P}_a\left(v \mid \left(\mathbf{y}^0_{a'}\right)^v_{a''}\right) = 0 \iff a''' \in \mathcal{NR}_a \setminus \mathcal{NC}_a.$$

Thus, we identify $\mathcal{NR}_a \setminus \mathcal{NC}_a$ and $\mathcal{NCR}_a$.

B.4. Proof of Proposition 3.4

Fix $a \in \mathcal{A}$ and $v \in \mathcal{Y} \setminus \{0\}$. Assume first that $|\mathcal{NC}_a \setminus \mathcal{NR}_a| \geq 1$. Under this situation, the relative consideration probability is identified via switching the choice of just one consideration-only peer from alternative v to the default while keeping the configuration of others fixed. Specifically, take $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$ and $\mathbf{y}$ such that every peer in $\mathcal{NC}_a$ picks v. Then

$$\frac{\mathrm{P}_a(v|\mathbf{y})}{\mathrm{P}_a(v|\mathbf{y}^0_{a'})} = \frac{\mathrm{Q}_a(v|\, |\mathcal{NC}_a|)}{\mathrm{Q}_a(v|\, |\mathcal{NC}_a| - 1)}.$$

Next, redefine $\mathbf{y}$ as before except that we let one of the peers from $\mathcal{NCR}_a$ to pick 0. As a result,

$$\frac{\mathrm{P}_a(v|\mathbf{y})}{\mathrm{P}_a(v|\mathbf{y}^0_{a'})} = \frac{\mathrm{Q}_a(v|\, |\mathcal{NC}_a| - 1)}{\mathrm{Q}_a(v|\, |\mathcal{NC}_a| - 2)}.$$

Repeating this procedure, we identify

$$\mathrm{Q}_a(v \mid n_1) / \mathrm{Q}_a(v \mid n_1 - 1) \text{ for all } n_1 \in \{|\mathcal{NC}_a| - |\mathcal{NCR}_a|, \ldots, |\mathcal{NC}_a|\}.$$

Next, we take $\mathbf{y}$ such that all peers in $\mathcal{NCR}_a$ and one of the peers in $\mathcal{NC}_a \setminus \mathcal{NR}_a$ different from a' are picking 0 and the rest of peers in $\mathcal{NC}_a \setminus \mathcal{NR}_a$ are picking v. Switching one by one all peers in $\mathcal{NC}_a \setminus \mathcal{NR}_a$ we identify $\mathrm{Q}_a(v \mid n_1) / \mathrm{Q}_a(v \mid n_1 - 1)$ for all n_1.

We next show that the relative consideration probability can be identified even if the consideration-only group is empty. Specifically, assume that $|\mathcal{NC}_a \setminus \mathcal{NR}_a| = 0$, so we have $|\mathcal{NR}_a \setminus \mathcal{NC}_a| \geq 1$ by Assumption 4. Then the relative consideration probability can be identified by switching one preference-only peer from v to the default and one preference-and-consideration peer from the default to alternative v. Specifically, take $a' \in \mathcal{NR}_a \setminus \mathcal{NC}_a$, $a'' \in \mathcal{NCR}_a$, and $\mathbf{y}$ such that every peer in $\mathcal{NCR}_a$ picks v and a' picks 0. Then, comparing Agent a's probability of choosing alternative v between configuration $\mathbf{y}$ and a configuration of switching Agent a' from 0 to alternative v and Agent a'' from alternative v to 0, which does not change the choice probability given consideration because the number of peers affecting preference is the same in both scenario, so we have

$$\frac{\mathrm{P}_a(v|\mathbf{y})}{\mathrm{P}_a\left(v|\,(\mathbf{y}^v_{a'})^0_{a''}\right)} = \frac{\mathrm{Q}_a(v|\,|\mathcal{NC}_a|)}{\mathrm{Q}_a(v|\,|\mathcal{NC}_a|-1)}.$$

Next, redefine $\mathbf{y}$ as before except that we let one of the peers from $\mathcal{NCR}_a$ different from a'' to pick 0. As a result,

$$\frac{\mathrm{P}_a(v|\mathbf{y})}{\mathrm{P}_a\left(v|\,(\mathbf{y}^v_{a'})^0_{a''}\right)} = \frac{\mathrm{Q}_a(v|\,|\mathcal{NC}_a|-1)}{\mathrm{Q}_a(v|\,|\mathcal{NC}_a|-2)}.$$

Repeating this procedure finitely many times we identify $\mathrm{Q}_a(v \mid n_1)/\,\mathrm{Q}_a(v \mid n_1 - 1)$ for all $n_1 \in \{1, \ldots, |\mathcal{NC}_a|\}$.

B.5. Proof of Proposition 3.5

Fix some $a \in \mathcal{A}$ and $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$. Moreover, take any distinct $v, v' \in \mathcal{Y} \setminus \{0\}$. Take any $\mathbf{y}$ such that no one picks v'. Since we will only use the variation in choices of Agent a', we drop the choices of everyone else from the notation. For example, $\mathrm{P}_a(v|v')$ is equal to $\mathrm{P}_a(v|\mathbf{y})$, where $y_{a'} = v'$. We use $t_{v'}$ to denote the ratio between the probability that Agent a picks v' conditional on Agent a' choosing v' and the default 0:

$$t_{v'} \equiv \frac{\mathrm{P}_a(v'|v')}{\mathrm{P}_a(v'|0)} = \frac{\mathrm{Q}_a(v'|1)}{\mathrm{Q}_a(v'|0)} \neq 1,$$

where the second equality holds because we can cancel out the choice probability conditional on considering v', and the inequality follows by Assumption 2(iii). Note that $t_{v'}$ is identified from the data.

Moreover,

$$\mathrm{P}_a(v|0) = \mathrm{Q}_a(v'|0)\,\{\mathrm{R}^*_a(v|v') - \mathrm{P}^*_a\,(v \mid \mathcal{Y} \setminus \{v'\})\} + \mathrm{P}^*_a(v \mid \mathcal{Y} \setminus \{v'\}),$$

$$\mathrm{P}_a(v|v') = \mathrm{Q}_a(v'|1)\,\{\mathrm{R}^*_a(v|v') - \mathrm{P}^*_a(v \mid \mathcal{Y} \setminus \{v'\})\} + \mathrm{P}^*_a(v \mid \mathcal{Y} \setminus \{v'\}),$$

where $\mathrm{R}^*_a(v|v')$ denotes the probabilities that Agent a picks v conditional on considering v'. Since,

$\mathrm{Q}_a(v'|0)t_{v'} = \mathrm{Q}_a(v'|1)$, we obtain from the above two equations that

$$\mathrm{P}^*_a(v \mid \mathcal{Y} \setminus \{v'\}) = \frac{\mathrm{P}_a(v|v') - t_{v'}\,\mathrm{P}_a(v|0)}{1 - t_{v'}}.$$

Since the choice of v, v', a, a', and choices of everyone else was arbitrary, we can identify $\mathrm{P}^*_a(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v'\})$ for all $a \in \mathcal{A}$, $v' \neq v$, $v' \neq 0$, and $\mathbf{y}$ such that (i) $y_{a'} \neq v'$ for all $a' \in \mathcal{N}_a$ and (ii) $y_{a'} = 0$ for some $a' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$.

Applying the above argument to $\mathrm{P}^*_a(\cdot|\cdot, \mathcal{Y} \setminus \{v'\})$, we can identify $\mathrm{P}^*_a(v \mid \mathbf{y}, \mathcal{Y} \setminus \{v', v''\})$ for all $a \in \mathcal{A}$, $v'' \neq v$, $v'' \neq v'$, $v'' \neq 0$, and $\mathbf{y}$ such that (i) $y_{a'} \notin \{v', v''\}$ for all $a' \in \mathcal{N}_a$ and (ii) $y_{a'} = y_{a''} = 0$ for some $a', a'' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$, $a' \neq a''$.

Repeating the above argument $|\mathcal{NC}_a \setminus \mathcal{NR}_a|$ times, we can identify $\mathrm{P}^*_a(\cdot \mid \mathbf{y}, \mathcal{Y} \setminus \mathcal{Z})$ for all $\mathcal{Z} \subseteq \mathcal{Y} \setminus \{0\}$ and $\mathbf{y}$ with the following two properties. First, $y_{a'} \notin \mathcal{Z}$ for all $a' \in \mathcal{N}_a$. Second, if we take any different $|\mathcal{Z}|$ components of $\mathbf{y}$ that correspond to peers from $\mathcal{NC}_a \setminus \mathcal{NR}_a$, then these components have to be equal to 0 since we switched these $|\mathcal{Z}|$ peers to 0.

B.6. Proof of Proposition 3.6

Fix some $v \neq 0$. If $\mathrm{Q}_a(v \mid n_1)$ is known for some n_1 in the support, by Proposition 3.4, we identify $\mathrm{Q}_a(v \mid \cdot)$. If, instead, we know $\mathrm{R}_a(v \mid n_2, \{0, v\})$, then, since $|\mathcal{NC}_a \setminus \mathcal{NR}_a| \geq Y$, by Proposition 3.5, we identify

$$\mathrm{P}^*_a(v \mid \mathbf{y}, \{0, v\}) = \mathrm{Q}_a\left(v \mid \mathcal{NC}^v_a(\mathbf{y})\right) \mathrm{R}_a\left(v \mid \mathrm{NR}^v_a(\mathbf{y}), \{0, v\}\right)$$

for some $\mathbf{y}$ such that $\mathrm{NR}^v_a(\mathbf{y}) = n_2$. Hence, we identify $\mathrm{Q}_a(v \mid \mathcal{NC}^v_a(\mathbf{y}))$ and, by Proposition 3.4, we also identify $\mathrm{Q}_a(v \mid \cdot)$. Since, the choice of v was arbitrary, we identify Q_a.

By Proposition 3.5, we now can identify $\mathrm{R}_a(v \mid n_2, \{0, v\})$ for all $v \neq 0$ and n_2 in the support. Next, consider

$$\begin{aligned}\mathrm{P}^*_a(v \mid \mathbf{y}, \{0, v, v'\}) = {} & \mathrm{Q}_a\left(v \mid \mathcal{NC}^v_a(\mathbf{y})\right) \mathrm{Q}_a\left(v' \mid \mathcal{NC}^{v'}_a(\mathbf{y})\right) \mathrm{R}_a\left(v \mid \mathrm{NR}^v_a(\mathbf{y}), \{0, v\}\right) + \\ & + \mathrm{Q}_a\left(v \mid \mathcal{NC}^v_a(\mathbf{y})\right)\left(1 - \mathrm{Q}_a\left(v' \mid \mathcal{NC}^{v'}_a(\mathbf{y})\right)\right) \mathrm{R}_a\left(v \mid \mathrm{NR}^v_a(\mathbf{y}), \mathrm{NR}^{v'}_a(\mathbf{y}), \{0, v, v'\}\right).\end{aligned}$$

Since Q_a and R_a for binary consideration sets are identified, we identify R_a for all possible sets of size 3. Repeating the above argument, we identify R_a for all possible sets of size 4. Applying this argument finitely many times, we can identify R_a for all possible sets.

B.7. Proof of Proposition 3.7

Since $\lim_{\Delta\to 0}\mathcal{P}(\Delta)=\mathcal{M}$, we can recover the transition rate matrix $\mathcal{M}$ from the data. Recall that each element in the transition rate matrix is defined as

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

Thus, $\lambda_a\,\mathrm{P}_a(y'_a\mid\mathbf{y})=\mathrm{m}(y'_a,\mathbf{y}_{-a}\mid\mathbf{y})$. It follows that we can recover $\lambda_a\,\mathrm{P}_a(v\mid\mathbf{y})$ for each $v\in\mathcal{Y}$, $\mathbf{y}\in\mathcal{Y}^A$, and $a\in\mathcal{A}$. Note that, for each $\mathbf{y}\in\mathcal{Y}^A$,

$$\sum_{v\in\mathcal{Y}}\lambda_a\,\mathrm{P}_a(v\mid\mathbf{y})=\lambda_a\sum_{v\in\mathcal{Y}}\mathrm{P}_a(v\mid\mathbf{y})=\lambda_a.$$

Then we can also recover λ_a for each $a\in\mathcal{A}$.

B.8. Proof of Proposition 3.8

This proof builds on Theorem 1 of Blevins (2017) and Theorem 3 of Blevins (2018). For the present case, it follows from these two theorems, that the transition rate matrix $\mathcal{M}$ is generically identified if, in addition to the conditions in Proposition 3.8, we have that

$$(Y+1)^A-AY-1\geq\frac{1}{2}.$$

This condition is always satisfied if $A>1$. Thus, the identification of $\mathcal{M}$ follows because $A\geq 2$. We can then uniquely recover $(\mathrm{P}_a)_{a\in\mathcal{A}}$ and $(\lambda_a)_{a\in\mathcal{A}}$ from $\mathcal{M}$ as in the proof of Proposition 3.7.

B.9. Proof of Proposition 4.1

Under the assumptions of the propositions, by Proposition 3.2, we identify $\mathcal{N}_a$.

Suppose we know $\mathcal{NR}_a$. If it is empty, then $\mathcal{NR}_a \setminus \mathcal{NC}_a$ is empty. Hence, by Assumption 4′, $\mathcal{NCR}_a$ is empty and we identify the network. Suppose $\mathcal{NR}_a$ is nonempty and equal to $\mathcal{N}_a$. Then $\mathcal{NC}_a \setminus \mathcal{NR}_a$ is empty and Assumption 4′ implies that $\mathcal{NCR}_a$ is empty and we identify the network. Finally, suppose $\mathcal{NR}_a$ is nonempty and is not equal to $\mathcal{N}_a$. Hence, we identify a consideration-only peer and can use her to separate $\mathcal{NCR}_a$ from $\mathcal{NR}_a \setminus \mathcal{NC}_a$ by using double differences of $\ln \mathrm{P}_a$.

Suppose we know $\mathcal{NC}_a$. Similarly, to the above, the cases when $\mathcal{NC}_a$ is empty or is equal to $\mathcal{N}_a$, Assumption 4′ imply the identification of the network. Suppose $\mathcal{NC}_a$ is nonempty and is different from $\mathcal{N}_a$. Hence, we identify $\mathcal{NR}_a \setminus \mathcal{NC}_a$ and the sign of the peer effect in preferences. Next, we look at the set

$$\{\Delta^1_{a'} \ln \mathrm{P}_a(1 \mid \mathbf{0}) \,:\, a' \in \mathcal{NC}_a\}.$$

If this set has cardinality 2, then since we know the sign of the preference effect we can identify the sign of the consideration effect and thus identify $\mathcal{NCR}_a$. If the cardinality of the set is 1,then $\mathcal{NC}_a = \mathcal{NC}_a \setminus \mathcal{NR}_a$ by Assumption 4′ and the network structure is identified.

B.10. Proof of Remark 5

Suppose we know $\mathcal{NR}_a$. Then $\mathcal{NC}_a \setminus \mathcal{NR}_a$ is identified. If the latter is nonempty, then, by using the double difference involving a consideration-only peer and any peer from $\mathcal{NR}_a$, we can identify if the peer from $\mathcal{NR}_a$ is in $\mathcal{NCR}_a$ or not. Similarly, if we know $\mathcal{NC}_a$, we identify the network structure if $\mathcal{NR}_a \setminus \mathcal{NC}_a$ is not empty —to this end, we use Assumption 3(iv) to show that the double difference involving one preference-only peer and one peer from $\mathcal{NCR}_a$ is different from 0. Hence, the only two cases that are left are $\mathcal{NR}_a = \mathcal{N}_a$ and $\mathcal{NC}_a = \mathcal{N}_a$. But these cases are excluded by assumption.

B.11. Proof of Proposition 4.6

We fix any a and $\mathbf{y}$ and drop them from the notation to simplify the exposition. First note that for any linear order $\succ$ on $\mathcal{Y}$ that ranks the default the worst, by manipulating how fast the excluded covariates converge to the extreme point, we can always construct a sequence

$$\left\{w_{\succ,k} = \left(w_{a,v,k}\right)_{v\in\mathcal{Y}\setminus\{0\}}\right\}_{k=1}^{\infty}$$

with $\lim_{k\to+\infty} w_{a,v,k} = \overline{w}_{a,v}$ such that

$$\lim_{k\to\infty}\frac{\mathrm{U}_a(v'|w_{\succ,k})}{\mathrm{U}_a(v''|, w_{\succ,k})} = \lim_{k\to\infty}\frac{f_{a,v'}(w_{a,v'})}{f_{a,v''}(w_{a,v''})}\begin{cases}0, \text{ if } v''\succ v' \\ +\infty, \text{ otherwise.}\end{cases}$$

Next take $\succ$ such that v is the second worst according to $\succ$ and default is the worst. Then

$$\lim_{k\to\infty} \mathrm{P}_a\left(0 \mid w_{\succ,k}\right) = \mathrm{C}_a\left(\{0\}\right)$$

identifies the probability of considering the default only and

$$\lim_{k\to\infty} \mathrm{P}_a\left(v \mid w_{\succ,k}\right) = \mathrm{C}_a\left(\{0, v\}\right)$$

identifies $\mathrm{C}_a\left(\{0, v\}\right)$. Since we can do it for any v we identify the consideration probabilities for all sets of cardinality 2. Next let $\succ$ be such that $v'\succ v\succ 0$ and v' is the third worst. Then

$$\lim_{k\to\infty} \mathrm{P}_a\left(v' \mid w_{\succ,k}\right) = \mathrm{C}_a\left(\{0, v'\}\right) + \mathrm{C}_a\left(\{0, v, v'\}\right).$$

Since $\mathrm{C}_a\left(\{0, v'\}\right)$ was identified in the previous step and the choice of v and v' was arbitrary, we can identify the consideration probabilities for all sets of cardinality 3. Recursively, we can identify C_a for sets of all sizes. The fact that the choice of a and $\mathbf{y}$ was arbitrary completes the proof.